Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 55, No. 1 **September, 1974**

Printed in Japan by International Academic Printing Co., Ltd., Tokyo, Japan

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

CONTINUOUS SPECTRA OF A SINGULAR SYMMETRIC DIFFERENTIAL OPERATOR ON A HILBERT SPACE OF VECTOR-VALUED FUNCTIONS

ROBERT ANDERSON

Let H be the Hilbert space of complex vector-valued functions $f: [a, \infty) \to C^2$ such that f is Lebesgue measurable on $[a, \infty)$ and $\int_a^\infty f^*(s)f(s)ds < \infty$. Consider the formally self adjoint expression $\iota(y) = -y'' + Py$ on $[a, \infty)$, where y is a 2-vector and P is a 2×2 symmetric matrix of continuous real valued functions on $[a, \infty)$. Let D be the linear manifold in H defined by

$$D = \{f \varepsilon H: f, f' \text{ are absolutely continuous on compact subintervals of } [a, \infty),\ f \text{ has compact support interior to } [a, \infty) \text{ and } \iota(f) \varepsilon H\}.$$

Then the operator L defined by $L(y) = \iota(y)$, $y \varepsilon D$, is a real symmetric operator on D. Let L_0 be the minimal closed extension of L. For this class of minimal closed symmetric operators this paper determines sufficient conditions for the continuous spectrum of self adjoint extensions to be the entire real axis. Since the domain, D_0, of L_0 is dense in H, self adjoint extensions of L_0 do exist.

A general background for the theory of the operators discussed here is found in [1], [3], and [5]. The theorems in this paper are motivated by the theorems of Hinton [4] and Eastham and El-Deberky [2]. In [4], Hinton gives conditions on the coefficients in the scalar case to guarantee that the continuous spectrum of self adjoint extensions covers the entire real axis. Eastham and El-Deberky [2] study the general even order scalar operator.

DEFINITION 1. Let $\tilde{L}$ denote a self adjoint extension of L_0. Then we define the *continuous spectrum*, $C(\tilde{L})$, of $\tilde{L}$ to be the set of all λ for which there exists a sequence $\langle f_n \rangle$ in $D_{\tilde{L}}$, the domain of $\tilde{L}$, with the properties:

(i) $\|f_n\| = 1$ for all n,

(ii) $\langle f_n \rangle$ contains no convergent subsequence (i.e., is not compact), and

(iii) $\|(\tilde{L} - \lambda)f_n\| \to 0$ as $n \to \infty$.

For the self adjoint operator $\tilde{L}$ we have the following well-known lemma.

LEMMA 1. *The continuous spectrum of $\tilde{L}$ is a subset of the real numbers.*

Proof. Let $\lambda = \alpha + i\beta$ where $\beta \neq 0$. Then for all $f \in D_{\tilde{L}}$ we can see by expanding $\|(\tilde{L} - \lambda)f\|^2$ that

$$\|(\tilde{L} - \lambda)f\|^2 \geqq |\beta|^2 \|f\|^2 ,$$

which implies $\lambda \notin C(\tilde{L})$.

THEOREM 2. *Let $L(y) = y'' + P(t)y$ for $a \leqq t < \infty$, where $P(t) = \begin{bmatrix}\alpha(t) & \gamma(t)\\ \gamma(t) & \beta(t)\end{bmatrix}$ where $\gamma(t)$ is positive and has two continuous derivatives. Let $g(t) > 0$ be one of $\alpha(t)$ or $\beta(t)$, where both α and β are continuous on $[\alpha, \infty)$ and $g(t)$ has a continuous derivative. Then if for some sequence of intervals $\{A_m\}$ where $A_m \subseteqq [a, \infty)$, $A_m = [c_m - a_m, c_m + a_m]$ and $a_m \to \infty$, the following are satisfied:*

(i) $\min_{x \in A_m} \{g(x)\} \to \infty$,

(ii) $\int_{A_m} ((g'(x))^2)/(g(x))\, dx = o(a_m)$,

(iii) $\int_{A_m} g(x)dx = o(a_m^3)$,

(iv) $\int_{A_m} [\gamma(x)]^2 dx = o(a_m)$,

we can conclude that $C(\tilde{L})$ is $(-\infty, \infty)$.

Proof. We will establish the theorem for $g(t) = \alpha(t)$ since the other case follows in exactly the same way.

Note that to prove the theorem then we need only show that for any real number μ there is a sequence $\langle f_m \rangle$ in $D(\tilde{L})$ such that $\|f_m\| = 1$, $f_m \to 0$ a.e., f_m vanishes outside A_m and $\|(\tilde{L} - \mu)f_m\| \to 0$ as $m \to \infty$.

Let $\langle h_m \rangle$ be defined by

$$h_m(t) = \begin{Bmatrix} [1 - \{(t - c_m)/a_m\}^2]^3 & \text{for } |t - c_m| \leqq a_m \\ 0 & \text{for } |t - c_m| > a_m \end{Bmatrix} . \tag{1}$$

Then define $\langle f_m(t) \rangle$ by

$$f_m(t) = h_m(t)\begin{bmatrix} b_{m1}e^{iQ_1(t)} \\ b_{m2}e^{iQ_2(t)} \end{bmatrix} , \tag{2}$$

where Q_1, Q_2 are real functions with two continuous derivatives and b_{m1}, b_{m2} are normalization constants.

To find $|b_m| = \sqrt{b_{m1}^2 + b_{m2}^2}$ we have

$$1 = \|f_m\|^2 = \int_{c_m - a_m}^{c_m + a_m} |b_m|^2 h_m^2(t)dt = |b_m|^2 \int_{-a_m}^{a_m} \left[1 - \left(\frac{x}{a_m}\right)^2\right]^6 dx$$

$$= |b_m|^2 \int_{-1}^{1} a_m[1 - y^2]^6 dy = |b_m|^2(2a_m)\left[1 + \sum_{r=1}^{6} \binom{6}{r}(2r + 1)^{-1}\right] .$$

Hence for some positive constant K

(3) $$|b_m|^2 = K(2a_m)^{-1} ,$$

and

(4) $$|f_m(t)| \leqq |b_m| = \sqrt{K}/\sqrt{2}a_m .$$

Hence

(5) $$f_m \longrightarrow 0 \quad \text{as} \quad m \longrightarrow \infty ,$$

(6) $$|h_m^{(r)}(t)| \leqq K_r(a_m)^{-r} ,$$

where K_r does not depend on t or m.

Since $f_m \in D(\tilde{L})$, we have

$$(\tilde{L} - \mu I)f_m = f_m'' + (P - \mu I)f_m$$
$$= \begin{bmatrix} f_{m1}'' + (\alpha - \mu)f_{m1} + \gamma f_{m2} \\ f_{m2}'' + (\beta - \mu)f_{m2} + \gamma f_{m1} \end{bmatrix}$$
$$(\tilde{L} - \mu I)f_m = \begin{bmatrix} \{-Q_1'^2 + (\alpha - \mu)\}f_{m1} + \gamma f_{m2} + iQ_1''f_{m1} \\ \{-Q_2'^2 + (\beta - \mu)\}f_{m2} + \gamma f_{m1} + iQ_2''f_{m2} \end{bmatrix} + \begin{bmatrix} b_{m1}e^{iQ_1}h_m'' + 2iQ_1'b_{m1}e^{iQ_1}h_m' \\ b_{m2}e^{iQ_2}h_m'' + 2iQ_2'b_{m2}e^{iQ_2}h_m' \end{bmatrix} .$$

Now if Q_1 is chosen so that

$$Q_1'^2 = \alpha - \mu , \qquad Q_1'' = \frac{\alpha'}{2\sqrt{\alpha - \mu}} ,$$

and b_{m2} is chosen to be identically zero we have that

$$(\tilde{L} - \mu I)f_m = \begin{bmatrix} iQ_1''f_{m1} \\ \gamma f_{m1} \end{bmatrix} + \begin{bmatrix} b_{m1}e^{iQ_1}h_m'' + 2iQ_1'b_{m1}e^{iQ_1}h_m' \\ 0 \end{bmatrix} .$$

By the way Q_1 is chosen,

$$\|(\tilde{L} - \mu I)f_m\| \leqq \left\|\left(\frac{\alpha'}{2\sqrt{\alpha - \mu}}\right)f_m\right\| + \|\gamma f_m\| + \|b_m h_m''\| + \|2Q_1'b_m h_m'\| .$$

Now, by (ii)

$$\left\|\frac{\alpha'}{2\sqrt{\alpha - \mu}}f_m\right\| \leqq \left[\frac{K}{a_m}\int_{A_m}\left(\frac{\alpha'}{2\sqrt{\alpha - \mu}}\right)^2\right]^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty .$$

By condition (iv),

$$\|\gamma f_m\| \leqq \left(\frac{K}{a_m}\int_{A_m}|\gamma|^2\right)^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty .$$

Next, by (iii), (3) and (6)

$$\| Q_1' b_m h_m' \| = \left(\int_{A_m} (\alpha - \mu) \frac{K}{2a_m} \cdot \frac{K_1^2}{a_m^2} \right)^{1/2}$$

$$= K_1 K^{1/2} \left(\frac{1}{2a_m^3} \int_{A_m} (\alpha - \mu) \right)^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty .$$

Then, by (3), (6), and the Cauchy-Schwartz Inequality

$$\| b_m h_m'' \| \leqq \left(\int_{A_m} | b_m |^2 \right)^{1/2} \left(\int_{A_m} | h_m'' |^2 \right)^{1/2}$$

$$\leqq \sqrt{K/2} \left(\int_{A_m} (K_r^2 / a_m^2) \right)^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty .$$

Hence it follows that

$$\| (\tilde{L} - \mu I) f_m \| \longrightarrow 0 \quad \text{as} \quad m \longrightarrow \infty ,$$

which is what we were to show.

COROLLARY 3. *If* $P(t) = \begin{bmatrix} at^\sigma & ct^\gamma \\ ct^\gamma & bt^\beta \end{bmatrix}$ *on some half-line* $d \leqq t < \infty$ *in Theorem 2 and*

(i) $a, c > 0$ *with* $\delta < 0$, $0 < \sigma < 2$, *or*

(ii) $b, c > 0$ *with* $\delta < 0$, $0 < \eta < 2$

then $C(\tilde{L}) = (-\infty, \infty)$.

THEOREM 4. *Suppose* $L(y)$ *is as in Theorem 2, where* $\gamma(t)$ *is positive and has two continuous derivatives. If for some sequence of intervals* $\{A_m\}$, *where* $A_m = [c_m - a_m, c_m + a_m]$, $A_m \subseteqq [a, \infty)$ *and* $a_m \to \infty$, *the following are satisfied*:

(i) $\min_{t \in A_m} \{\gamma(t)\} \to \infty$,

(ii) $\int_{A_m} ((\gamma'(t))^2)/(\gamma(t)) \, dt = o(a_m)$,

(iii) $\int_{A_m} \gamma(t) dt = o(a_m^3)$,

(iv) $\int_{A_m} \alpha^2(t) dt$ *and* $\int_{A_m} \beta^2(t) dt$ *are* $o(a_m)$,

then $C(\tilde{L}) = (-\infty, \infty)$.

Proof. In the proof of Theorem 2 choose $Q_1'^2 = Q_2'^2 = \gamma(t) - \mu$, so that $f_{m1} = f_{m2}$. Then $Q_1'' = Q_2'' = (\gamma'(t))/(2\sqrt{\gamma(t) - \mu})$ and applying conditions (i) – (iv) as before where $g(t)$ is replaced by $\gamma(t)$ we get that $\| (\tilde{L} - \mu I) f_m \| \to 0$ as $m \to \infty$.

COROLLARY 5. *Let* $P(t) = \begin{bmatrix} at^\sigma & ct^\delta \\ ct^\delta & bt^\eta \end{bmatrix}$ *in Theorem* 4. *If* $c > 0$, $0 < \delta < 2$ *and* σ, $\eta < 0$ *then* $C(\tilde{L}) = (-\infty, \infty)$.

Let H be the Hilbert space $\tilde{L}_2([a, \infty), w)$ of complex vector-valued functions $f: [a, \infty) \to C^2$ such that $\|f\|^2 = \int_a^\infty w(f^*f) < \infty$, where w is positive and $w \in C^{(2)}[a, \infty)$. Let $l(y) \equiv (1/w)y'' + Py$. Then define L_0 as before and let $\tilde{L}$ be a self adjoint extension of L_0.

THEOREM 6. *Suppose there is a sequence of intervals, $A_m \subseteqq [a, \infty)$, $A_m = [c_m - a_m, c_m + a_m]$ where $a_m \to \infty$ as $m \to \infty$, such that*

(i) $\int_{A_m} (\alpha(w')^2)/w^3 = o(|a_m|)$, $\int_{A_m} \alpha/w = o(|A_m|)^3$, $\min_{t \in A_m} \alpha(t) \to \infty$,

(ii) $\int_{A_m} (w')^4/w^6 = o(|A_m|)$, $\int_{A_m} (w''/w^2)^2 = o(|A_m|)$,

$\int_{A_m} 1/w^2 = o(|A_m|^5)$,

(iii) $\int_{A_m} ([(wa)']^2)/(\alpha w^3) = o(|A_m|)$, *and*

(iv) $\int_{A_m} \gamma^2 = o(|A_m|)$

as $m \to \infty$. Then $C(\tilde{L}) = (-\infty, \infty)$.

Note that (ii) implies that $\int_{A_m} (w'/w^2)^2 = o(|A_m|^3)$ by $(w'/w^2)^2 = (w')^2/w^3 \cdot 1/w$ and Cauchy-Schwartz Inequality.

Proof. As is the previous theorem define

$$f_m = \begin{bmatrix} f_{m1} \\ f_{m2} \end{bmatrix} \quad \text{where} \quad f_{m2} = 0 \quad \text{and} \quad f_{m1} = (b_m e^{iQ} h_m) w^{-1/2} .$$

Then again $b_m^2 = K/a_m$ and $|f_{m1}| \leqq b_m w^{-1/2} = (K/(wa_m))^{1/2}$. Calculating

$$\begin{aligned} f'_{m1} &= w^{-1/2} b_m e^{iQ} h'_m + f_{m1}[iQ' - 1/2 w^{-1} w'] \\ f''_{m1} &= f_{m1}[-(Q')^2 - iQ' w^{-1} w' + 3/4 w^{-2}(w')^2 - 1/2 w^{-1} w'' + iQ''] \\ &\quad + b_m e^{iQ}[2w^{-1/2} iQ' h'_m - w^{-3/2} w' h'_m + w^{-1/2} h''_m] . \end{aligned}$$

Then $(\tilde{L} - \mu I) f_m = (1/w) f''_m + P f_m$, where the top element is

$$\begin{aligned} &\frac{1}{w} f''_{m1} + (\alpha - \mu) f_{m1} = \frac{f_{m1}}{w}[-(Q')^2 + (\alpha - \mu) w] \\ &+ \frac{f_{m1}}{w}\left[-iQ' w^{-1} w' + \frac{3}{4} w^{-2}(w')^2 - \frac{1}{2} w^{-1} w'' + iQ''\right] \\ &+ b_m e^{iQ}[w^{-3/2}][2iQ' h'_m - w^{-1} w' h'_m + h''_m] \\ &= \frac{f_{m1}}{w}[-(Q')^2 + (\alpha - \mu) w] + \frac{f_{m1}}{w^3}\left[-iQ' w w' + \frac{3}{4}(w')^2 - \frac{1}{2} w w'' + w^2 iQ''\right] \\ &+ b_m e^{iQ} w^{-3/2}[2iQ' h'_m - w^{-1} w' h'_m + h''_m] . \end{aligned}$$

Of course, the second element of $(L - \mu I) f_m$ is γf_{m1}. By choosing $(Q')^2 = (\alpha - \mu) w$ we have that by (i)

$$Q' = [(\alpha - \mu)w]^{1/2} = O((\alpha w)^{1/2}) \quad \text{as} \quad t \longrightarrow \infty \, .$$

$$Q'' = O\left(\frac{[\alpha w]'}{\sqrt{\alpha w}}\right) \quad \text{as} \quad t \longrightarrow \infty \, .$$

Then by the calculations above

$$\begin{aligned} \| (\tilde{L} - \mu I) f_m \| \leqq & \left\| \frac{f_{m1}}{w^2} Q' w' \right\| + \frac{3}{4} \left\| \frac{f_{m1}}{w^3} (w')^2 \right\| + \frac{1}{2} \left\| f_{m1} \frac{w''}{w^2} \right\| \\ & + \left\| \frac{f_{m1} Q''}{w} \right\| + 2 \| b_m w^{-3/2} Q' h'_m \| \\ & + \| b_m w^{-5/2} w' h'_m \| \\ & + \| b_m w^{-3/2} h''_m \| + \| \gamma f_{m1} \| \, . \end{aligned} \tag{7}$$

Since $|f_{m1}|^2 \leqq K/(wa_m)$ and $(Q')^2 = (\alpha - \mu)w$,

$$\begin{aligned} & \| f_{m1} w^{-2} Q' w' \| \\ \leqq & \left(\frac{K}{a_m} \int_{A_m} (\alpha - \mu) w^{-3} (w')^2 \right)^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty \quad \text{by (i)} \, . \end{aligned}$$

Similarly,

$$\| f_{m1} w^{-3} (w')^2 \| \leqq \left(\frac{K}{a_m} \int_{A_m} [(w')^2 w^{-3}]^2 \right)^{1/2} = o(1) \quad \text{by (ii)} \, .$$

By the definition of Q and f_{m1},

$$\| f_{m1} w^{-1} Q'' \| = O\left(\int_{A_m} \frac{K[(\alpha w)']^2}{a_m \alpha w^3} \right)^{1/2} = o(1) \quad \text{by (iii)} \, .$$

And by condition (ii),

$$\| f_{m1} w^{-2} w'' \| \leqq \left(\frac{K}{a_m} \int_{Am} [(w'')^2 w^{-4}] \right)^{1/2} = o(1) \, .$$

Since $|b_m|^2 = K/a_m$ and $|h'_m| \leqq K_1/a_m$,

$$\| b_m w^{-3/2} Q' h'_m \| \leqq \left((KK_1^2/a_m^3) \int_{A_m} \left(\frac{\alpha - \mu}{w} \right) \right)^{1/2} = o(1) \quad \text{by (i)} \, .$$

Similarly, by the remark at the end of the theorem,

$$\| b_m w^{-5/2} w' h'_m \| \leqq \left((KK_1^2/a_m^3) \int_{A_m} (w')^2 w^{-4} \right)^{1/2} = o(1) \, .$$

Since $|h''_m| \leqq K_2/a_m^2$,

$$\| b_m w^{-3/2} h''_m \| \leqq \left((KK_2^2/a_m^5) \int_{A_m} w^{-2} \right)^{1/2} = o(1) \quad \text{by (ii)} \, .$$

By (iv),

$$\|\gamma f_{m1}\| \leqq \left((K/a_m)\int_{A_m}\gamma^2\right)^{1/2} = o(1) \quad \text{as} \quad m \longrightarrow \infty \ .$$

Hence, by the above calculations and (7),

$$\|(\tilde{L} - \mu I)f_m\| \longrightarrow 0 \quad \text{as} \quad m \longrightarrow \infty \ .$$

Since this is what we were to show, this conclude the proof.

REFERENCES

1. N. Dunford and J. T. Schwartz, *Linear Operators, Part II: Spectral Theory*, Interscience-Wiley, New York, 1963.
2. M. S. P. Eastham and A. A. El-Deberky, *The Spectrum of differential operators with large coefficients*, J. London Math. Soc., **2** (1970), 257-266.
3. I. M. Glazman, *Direct Methods of Qualitative spectral Analysis of Singular Differential Operators*, Israel Program for Scientific Translations, Jerusalem, 1965.
4. D. Hinton, *Continuous spectra of second-order differential operators*, Pacific J. Math., **33**, No. 3 (1970), 641-643.
5. M. A. Naimark, *Linear Differential Operators, Part II*, Ungar, New York, 1968.

Received February 22, 1974, and in revised form October 8, 1974,

DEERING MILLIKEN, INC.

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

THE ISOMETRIES OF $L^p(X,K)$

MICHAEL CAMBERN

Let (X, Σ, μ) be a finite measure space, and denote by $L^p(X, K)$ the Banach space of measurable functions F defined on X and taking values in a separable Hilbert space K, such that $\|F(x)\|^p$ is integrable. In this article a characterization is given of the linear isometries of $L^p(X, K)$ onto itself, for $1 \leqq p < \infty$, $p \neq 2$. It is shown that T is such an isometry iff T is of the form $(T(F))(x) = U(x)h(x)(\Phi(F))(x)$, where Φ is a set isomorphism of Σ onto itself, U is a weakly measurable operator-valued function such that $U(x)$ is a.e. an isometry of K onto itself, and h is a scalar function which is related to Φ via a formula involving Radon-Nikodym derivatives.

Throughout this paper the letter K will represent a separable Hilbert space which may be either real or complex. We denote by $\langle \cdot, \cdot \rangle$ the inner product in K, and by S the one-dimensional Hilbert space which is the scalar field associated with K.

A function F from X to K will be called measurable if the scalar function $\langle F, e \rangle$ is measurable for each $e \in K$. Then for $1 \leqq p < \infty$, we denote by $L^p(X, K)$ the Banach space of (equivalence classes of) measurable functions F from X to K for which the norm

$$\|F\|_p = \left\{\int \|F(x)\|^p d\mu\right\}^{1/p}, \qquad p < \infty ,$$
$$\|F\|_\infty = \operatorname{ess\,sup} \|F(x)\|$$

is finite. (Here $\|\cdot\|_p$ denotes the norm in $L^p(X, K)$ and $L^p(X, S)$, and $\|\cdot\|$ that in K.) If $F \in L^p(X, K)$, we define the support of F to be the set $\{x \in X: F(x) \neq 0\}$.

Let $\{e_1, e_2, \cdots\}$ be some orthonormal basis for K. For $F \in L^p(X, K)$, we define the measurable coordinate functions f_n by $f_n(x) = \langle F(x), e_n \rangle$. Then almost everywhere we have $\sum_n |f_n(x)|^2 < \infty$, and $F(x) = \sum_n f_n(x)e_n$. Moreover, it is easily seen that each f_n belongs to $L^p(X, S)$.

Here we investigate the isometries of $L^p(X, K)$, for $1 \leqq p < \infty$, $p \neq 2$. For the case in which X is the unit interval, μ Lebesgue measure, and $K = S$, the isometries were determined by Banach in [1, p. 178]. In [4], Lamperti obtained a complete description of the isometries of $L^p(X, S)$ for an arbitrary finite measure space (X, Σ, μ).

Following Lamperti's terminology, we will call a mapping Φ of Σ onto itself, defined modulo null sets, a *regular set isomorphism* if it satisfies the properties

$$\Phi(A') = [\Phi(A)]' ,$$
$$\Phi\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \bigcup_{n=1}^{\infty} \Phi(A_n) ,$$

and

$$\mu[\Phi(A)] = 0 \quad \text{if, and only if ,} \quad \mu(A) = 0 ,$$

for all sets A, A_n in Σ. (Throughout, A' will denote the complement of A.) A regular set isomorphism induces a linear transformation, also denoted by Φ, on the space of measurable scalar functions defined on X, which is characterized by $\Phi(\chi_A) = \chi_{\Phi(A)}$, where χ_A is the characteristic function of the measurable set A. This process is described in [3, pp. 453–454]. The induced transformation, moreover, has the property that it preserves a.e. convergence:

(1) $\quad$ if $\lim_n f_n(x) = f(x)$ a.e., then $\lim_n (\Phi(f_n))(x) = (\Phi(f))(x)$ a.e.

Now given a regular set isomorphism Φ of Σ onto itself, and $F = \sum_n f_n e_n \in L^p(X, K)$, we define $\Phi(F)$ by the equation

$$(\Phi(F))(x) = \sum_n (\Phi(f_n))(x) e_n . \tag{2}$$

For the case in which K is infinite dimensional, one must, of course, verify that the series on the right in (2) is indeed convergent in K for almost all x. But, for all scalar simple functions, we have $(\Phi(|f|^2))(x) = |\Phi(f)|^2(x)$ and hence, by (1), this identity holds for all measurable scalar functions. Thus, as $\|F(x)\|^2 = \sum_n |f_n(x)|^2 = \lim_N \sum_{n=1}^N |f_n(x)|^2$, again using (1), we have

$$\begin{aligned} |\Phi(\|F\|)|^2(x) &= (\Phi(\|F\|^2))(x) = \lim_N \Big(\Phi\Big(\sum_{n=1}^N |f_n|^2\Big)\Big)(x) \\ &= \lim_N \sum_{n=1}^N |(\Phi(f_n))(x)|^2 = \sum_n |(\Phi(f_n))(x)|^2 = \|(\Phi(F))(x)\|^2 . \end{aligned} \tag{3}$$

Moreover, it is readily verified that the definition of $\Phi(F)$ is independent of the choice of orthonormal basis for K.

For the case in which K is one-dimensional, Lamperti has shown that if T is an isometry of $L^p(X, S)$ onto itself, $1 \leqq p < \infty$, $p \neq 2$, then there exists a regular set isomorphism Φ, and a measurable scalar function $h(x)$ such that for $f \in L^p(X, S)$

$$(T(f))(x) = h(x)(\Phi(f))(x) . \tag{4}$$

Moreover, if the measure ν is defined by $\nu(A) = \mu[\Phi^{-1}(A)]$, $A \in \Sigma$, then

$$|h(x)|^p = d\nu/d\mu \quad \text{a.e. on} \quad X . \tag{5}$$

Conversely, given any regular set isomorphism Φ of Σ onto itself, and a function $h(x)$ satisfying (5), the operator T defined by (4) is an isometry of $L^p(X, S)$ onto itself. Here we establish that the isometries of $L^p(X, K)$, for any separable Hilbert space K, closely resemble those of $L^p(X, S)$, except for the emergence of a measurable operator-valued function.

2. **The isometries.** We begin with a lemma whose proof exactly parallels that of Lemma 14, [5, p. 331], with the real numbers ξ and η in that lemma replaced by vectors in K.

Lemma 1. *Let φ and ψ be two elements of K. If $1 \leqq p \leqq 2$, then*

$$\|\varphi + \psi\|^p + \|\varphi - \psi\|^p \leqq 2(\|\varphi\|^p + \|\psi\|^p) ,$$

and if $2 \leqq p < \infty$,

$$\|\varphi + \psi\|^p + \|\varphi - \psi\|^p \geqq 2(\|\varphi\|^p + \|\psi\|^p) .$$

If $p \neq 2$, equality can hold only if φ or ψ is zero.

By integration, we then obtain the following:

Lemma 2. *If $1 \leqq p < \infty$ and $p \neq 2$, and if F and G are in $L^p(X, K)$, then*

$$\|F + G\|_p^p + \|F - G\|_p^p = 2\|F\|_p^p + 2\|G\|_p^p \tag{6}$$

if and only if F and G have a.e. *disjoint supports.*

Throughout the remainder of this article we assume that p is a given real number with $1 \leqq p < \infty$, $p \neq 2$. We define q to be that extended real number such that $1/p + 1/q = 1$. (The usual conventions are in effect.) T will denote a fixed isometry of $L^p(X, K)$ onto itself.

We will repeatedly use the map T^{*-1} defined on $L^q(X, K)$ by

$$\int \langle F(x), (T^{*-1}(G))(x) \rangle d\mu = \int \langle (T^{-1}(F))(x), G(x) \rangle d\mu ,$$

for $F \in L^p(X, K)$, $G \in L^q(X, K)$, which is, almost, the Banach space adjoint of T^{-1}. For the dual space of $L^p(X, K)$ is $L^q(X, K^*)$, where K^* is the dual of K, [2, p. 282]. And if σ is the usual conjugate-linear isometry of K^* onto K, σ induces a conjugate-linear isometric mapping of $L^q(X, K^*)$ onto $L^q(X, K)$, which we shall also denote by σ, and which is determined by $(\sigma(G^*))(x) = \sigma(G^*(x))$, $G^* \in L^q(X, K^*)$. Our map T^{*-1} is then actually $\sigma \circ T^{\#-1} \circ \sigma^{-1}$, where $T^{\#-1}$ is the true Banach space adjoint.

For any element $e \in K$, we denote by $\mathbf{e}$ that element of $L^p(X, K)$ which is constantly equal to e. If $e \neq 0$, it is an easy consequence of (6), and of the fact that T is onto, that the support of $T(\mathbf{e})$ must be equal to X a.e.

LEMMA 3. *Let e be any vector in K. If A is any measurable subset of X, then $T(\chi_A e)$ is equal to $T(\mathbf{e})$ on the support of $T(\chi_A e)$.*

Proof. The functions $\chi_A e$ and $\chi_{A'} e$ have disjoint supports, and thus (6) holds if F and G are replaced, respectively, by $\chi_A e$ and $\chi_{A'} e$. Since T is isometric, it follows that (6) also holds for $T(\chi_A e)$ and $T(\chi_{A'} e)$, and hence that these latter two functions have disjoint supports. Since $T(e) = T(\chi_A e) + T(\chi_{A'} e)$, the desired conclusion follows.

LEMMA 4. *Let e be an element of K with $\| e \| = 1$, and let $F = T(\mathbf{e})$. If E is the vector function defined* a.e. *by $E(x) = F(x)/\| F(x) \|$, then $T^{*-1}(\mathbf{e})$ is that element of $L^q(X, K)$ determined by $(T^{*-1}(\mathbf{e}))(x) = \| F(x) \|^{p-1} E(x)$ for almost all $x \in X$.*

Proof. We have $\| F \|_p = \| \mathbf{e} \|_p = [\mu(X)]^{1/p}$. Moreover, as T^{*-1} is an isometry of $L^q(X, K)$ onto itself, we also have $\| T^{*-1}(\mathbf{e}) \|_q = [\mu(X)]^{1/q}$, this latter equality holding even in the limiting case $q = \infty$, since $\| \mathbf{e} \|_\infty = 1$.

Let $G = T^{*-1}(\mathbf{e})$, and define the vector function H by $H(x) = G(x)/\| G(x) \|$ if x belongs to the support of G, and $H(x) = 0$ otherwise. (If $q = \infty$, we do not yet know that the support of G is equal to X a.e., although this fact can readily be established by a separate argument involving extreme points.) We then have

$$
\begin{aligned}
\mu(X) &= \int \langle \mathbf{e}, \mathbf{e} \rangle d\mu = \int \langle (T(\mathbf{e}))(x), (T^{*-1}(\mathbf{e}))(x) \rangle d\mu \\
&= \int \langle F(x), G(x) \rangle d\mu \\
&= \int \| F(x) \| \, \| G(x) \| \, \langle E(x), H(x) \rangle d\mu \\
&\leqq \int \| F(x) \| \, \| G(x) \| \, d\mu \leqq \| F \|_p \| G \|_q = \mu(X) \, .
\end{aligned} \tag{7}
$$

Hence we must have equality throughout in (7). Thus, by a known result for scalar functions, [5, p. 113], for $p > 1$ the equality $\int \| F(x) \| \, \| G(x) \| \, d\mu = \| F \|_p \| G \|_q$ implies that

$$\| G(x) \|^q = \| G \|_q^q \| F(x) \|^p / \| F \|_p^p = \| F(x) \|^p$$

a.e., so that $\| G(x) \| = \| F(x) \|^{p-1}$ a.e. If $p = 1$, the equality

$\int \| F(x) \| \| G(x) \| d\mu = \mu(X) = \| F \|_1$ implies that $\| G(x) \| = 1 = \| F(x) \|^{p-1}$ a.e. in this case too. Finally, the equality

$$\int \| F(x) \| \| G(x) \| \langle E(x), H(x) \rangle d\mu = \int \| F(x) \| \| G(x) \| d\mu$$

yields the fact that $H(x) = E(x)$ a.e., which completes the proof of the lemma.

LEMMA 5. *Let e and φ be two orthogonal elements of K, each with norm one, and let $F_e = T(\mathbf{e})$ and $F_\varphi = T(\varphi)$. If E_e and E_φ are the vector functions defined* a.e. *by $E_e(x) = F_e(x)/\| F_e(x) \|$ and $E_\varphi(x) = F_\varphi(x)/\| F_\varphi(x) \|$, then $\langle E_e(x), E_\varphi(x) \rangle = 0$* a.e.

Proof. Let A be any measurable subset of X. Then $F_e = \chi_A F_e + \chi_{A'} F_e$, and since the two functions on the right have disjoint supports, (6) holds when F and G are replaced, respectively, by $\chi_A F_e$ and $\chi_{A'} F_e$. Hence (6) also holds for $T^{-1}(\chi_A F_e)$ and $T^{-1}(\chi_{A'} F_e)$, and these latter functions thus have disjoint supports. Since $\mathbf{e} = T^{-1}(\chi_A F_e) + T^{-1}(\chi_{A'} F_e)$, if we let B denote the support of $T^{-1}(\chi_A F_e)$, it follows that $T(\chi_B e) = \chi_A F_e$.

We then have, using Lemma 4,

$$\begin{aligned} 0 = \int \langle \chi_B e, \varphi \rangle d\mu &= \int \langle (T(\chi_B e))(x), (T^{*-1}(\varphi))(x) \rangle d\mu \\ &= \int \langle \chi_A \| F_e(x) \| E_e(x), \| F_\varphi(x) \|^{p-1} E_\varphi(x) \rangle d\mu \\ &= \int_A \| F_e(x) \| \| F_\varphi(x) \|^{p-1} \langle E_e(x), E_\varphi(x) \rangle d\mu \, . \end{aligned}$$

Since $\| F_e(x) \| \| F_\varphi(x) \|^{p-1}$ is an a.e. positive element of $L^1(X, S)$, and A is an arbitrary measurable subset of X, we must have $\langle E_e(x), E_\varphi(x) \rangle = 0$ a.e. on X.

LEMMA 6. *For any element e of K with norm one, let F_e and E_e be defined as in the previous lemma. Then for $f \in L^p(X, S)$, $(T(fe))(x) = \tilde{f}(x) E_e(x)$ for some scalar function $\tilde{f}$, and the mapping $f(x) \to \langle (T(fe))(x), E_e(x) \rangle$ is an isometry of $L^p(X, S)$ onto itself.*

Proof. If A is any measurable subset of X, we know from Lemma 3 that $(T(\chi_A e))(x)$ is equal to $\| F_e(x) \| E_e(x)$ on the support of $T(\chi_A e)$. It thus follows that for any simple function $f \in L^p(X, S)$, $(T(fe))(x) = \tilde{f}(x) E_e(x)$, where $\tilde{f}$ is a function in $L^p(X, S)$ with the same norm as f. For arbitrary $f \in L^p(X, S)$, let $\{f_k\}$ be a sequence of simple functions converging to f in the norm of $L^p(X, S)$. Then

$$\lim_k \int \| (T(f_k e))(x) - (T(fe))(x) \|^p \, d\mu = 0 \, .$$

Hence $\| (T(f_k e))(x) - (T(fe))(x) \|^p$ tends to zero in measure, and so a subsequence tends to zero a.e. That is, $(T(f_{k_j} e))(x)$ tends to $(T(fe))(x)$ almost everywhere.

Now, for almost all x, the elements of K given by $(T(f_{k_j} e))(x)$, $j = 1, 2, \cdots$ lie in the one-dimensional (hence closed) subspace of K spanned by $E_e(x)$, and thus $(T(fe))(x)$ must lie in this subspace. That is, $(T(fe))(x) = \tilde{f}(x)E_e(x)$, for some $\tilde{f} \in L^p(X, S)$ with $\|\tilde{f}\|_p = \|f\|_p$, and the given mapping is an isometry of $L^p(X, S)$ into itself.

It is readily seen that the map is, in fact, onto $L^p(X, S)$. For suppose we are given a function of the form $\tilde{f}(x)E_e(x)$, where $\tilde{f} \in L^p(X, S)$. Incorporate e into an orthonormal basis for K — say $e = e_1$, where $\{e_n : n = 1, 2, \cdots\}$ is such a basis. Let $F(x) = \sum_n f_n(x)e_n$ be the element of $L^p(X, K)$ which maps onto $\tilde{f}(x)E_e(x)$ under T.

Now $F_0(x) = \sum_{n \geqq 2} f_n(x)e_n$ belongs to $L^p(X, \hat{K})$, where $\hat{K}$ is the Hilbert space which is the closed linear span of $\{e_n : n \geqq 2\}$, and vector-valued simple functions of the form $G = \sum_{j=1}^r \chi_{A_j}\varphi_j$, $\varphi_j \in \hat{K}$, are dense in $L^p(X, \hat{K})$. By Lemmas 3 and 5, for all such G, $\langle (T(G))(x), E_e(x) \rangle = 0$ a.e., from which it follows that $\langle (T(F_0))(x), E_e(x) \rangle = 0$ a.e. Thus as $\tilde{f}(x)E_e(x) = (T(f_1 e))(x) + (T(F_0))(x)$, with $(T(f_1 e))(x)$ pointwise a scalar multiple of $E_e(x)$ and $(T(F_0))(x)$ a.e. orthogonal to $E_e(x)$, we conclude that $T(F_0)$, and hence F_0, are both equal to the zero element of $L^p(X, K)$. It follows that the mapping given by the lemma is indeed onto $L^p(X, S)$.

LEMMA 7. *Let* $\{e_n : n = 1, 2, \cdots\}$ *be some fixed orthonormal basis for* K, *and for each* n *define* F_n, E_n *by* $F_n = T(\mathbf{e}_n)$, $E_n(x) = F_n(x)/\|F_n(x)\|$. *Then there exists a regular set isomorphism* Φ *and a fixed scalar function* $h(x)$ *defined on* X *and satisfying* (5), *such that for all* $n = 1, 2, \cdots$ *and for all* $f \in L^p(X, S)$, $(T(fe_n))(x) = h(x)(\Phi(f))(x)E_n(x)$.

Proof. By Lemma 6 and Lamperti's result for scalar functions, we know that if e_m and e_n are two elements of the given orthonormal basis and if $f \in L^p(X, S)$, then $(T(fe_m))(x) = h_m(x)(\Phi_m(f))(x)E_m(x)$ and $(T(fe_n))(x) = h_n(x)(\Phi_n(f))(x)E_n(x)$, where $h_m(x)$ and $h_n(x)$ are scalar functions defined on X, and Φ_m, Φ_n are linear transformations induced by regular set isomorphisms. We wish to show that $h_m = h_n$ and $\Phi_m = \Phi_n$ modulo sets of measure zero.

If A is any measurable subset of X, we have

$$(T(\chi_A e_m))(x) = h_m(x)\chi_{\Phi_m(A)}(x)E_m(x) \, , \tag{8}$$

and

$$(T(\chi_A e_n))(x) = h_n(x)\chi_{\Phi_n(A)}(x)E_n(x) . \tag{9}$$

Consider $\chi_A(e_m + e_n)/\sqrt{2}$. If we let $F_{m,n} = T[(e_m + e_n)/\sqrt{2}]$, and define $E_{m,n}$ by $E_{m,n}(x) = F_{m,n}(x)/\| F_{m,n}(x) \|$, again by using Lemma 6 and Lamperti's result, we conclude that there exists a scalar function $h_{m,n}$ and a regular set isomorphism $\Phi_{m,n}$ such that

$$(T[\chi_A(e_m + e_n)/\sqrt{2}])(x) = h_{m,n}(x)\chi_{\Phi_{m,n}(A)}(x)E_{m,n}(x) . \tag{10}$$

Now, using the linearity of T, we have

$$\begin{aligned} E_{m,n}(x) &= F_{m,n}(x)/\| F_{m,n}(x) \| \\ &= (F_m(x) + F_n(x))/\| F_m(x) + F_n(x) \| \\ &= (\| F_m(x) \| E_m(x) + \| F_n(x) \| E_n(x))/\| F_m(x) + F_n(x) \| . \end{aligned} \tag{11}$$

And, combining (11) with Lemma 4, we have

$$\begin{aligned} &(T^{*-1}[(e_m + e_n)/\sqrt{2}])(x) = \| (F_m(x) + F_n(x))/\sqrt{2} \|^{p-1} E_{m,n}(x) \\ &= \| (F_m(x) + F_n(x))/\sqrt{2} \|^{p-1} (\| F_m(x) \| E_m(x) \\ &\quad + \| F_n(x) \| E_n(x))/\| F_m(x) + F_n(x) \| . \end{aligned} \tag{12}$$

Also, using Lemma 4 and the linearity of T^{*-1}, we find that

$$\begin{aligned} (T^{*-1}[(e_m + e_n)/\sqrt{2}])(x) = \| F_m(x) \|^{p-1} E_m(x)/\sqrt{2} \\ + \| F_n(x) \|^{p-1} E_n(x)/\sqrt{2} . \end{aligned} \tag{13}$$

Since Lemma 5 shows that $E_m(x)$ and $E_n(x)$ are a.e. linearly independent, we conclude from (12) and (13) that

$$2^{(1-p)/2} \| F_m(x) + F_n(x) \|^{p-2} \| F_m(x) \| = \| F_m(x) \|^{p-1}/\sqrt{2} \text{ , a.e.,}$$

from which it follows that $\| F_m(x) + F_n(x) \| = \sqrt{2} \| F_m(x) \|$ a.e. Similarly, $\| F_m(x) + F_n(x) \| = \sqrt{2} \| F_n(x) \|$ a.e., so that (11) then gives $E_{m,n}(x) = E_m(x)/\sqrt{2} + E_n(x)/\sqrt{2}$.

Thus from (10) we conclude that $(T[\chi_A(e_m + e_n)/\sqrt{2}])(x) = h_{m,n}(x)\chi_{\Phi_{m,n}(A)}(x)E_m(x)/\sqrt{2} + h_{m,n}(x)\chi_{\Phi_{m,n}(A)}(x)E_n(x)/\sqrt{2}$. But the linearity of T, together with (8) and (9), implies that $(T[\chi_A(e_m+e_n)/\sqrt{2}])(x) = h_m(x)\chi_{\Phi_m(A)}(x)E_m(x)/\sqrt{2} + h_n(x)\chi_{\Phi_n(A)}(x)E_n(x)/\sqrt{2}$. Hence, once again employing the a.e. linear independence of $E_m(x)$ and $E_n(x)$, we find that $h_m(x)\chi_{\Phi_m(A)}(x) = h_{m,n}(x)\chi_{\Phi_{m,n}(A)}(x) = h_n(x)\chi_{\Phi_n(A)}(x)$ a.e. Since this equality holds for every measurable set A, we can conclude that $h_n = h_m$ and $\Phi_n = \Phi_m$, modulo sets of measure zero.

Thus, if we let $\Phi = \Phi_1$ and $h = h_1$, then for all $f \in L^p(X, S)$ and all n, we have $(T(fe_n))(x) = h(x)(\Phi(f))(x)E_n(x)$ a.e., and $h = h_1$ satisfies (5) by Lemma 6. This concludes the proof of lemma.

A function U defined on X and taking values in the space of bounded operators on K is called weakly measurable if $\langle U(x)e, \varphi\rangle$ is measurable for all $e, \varphi \in K$.

THEOREM. *Let T be an isometry of $L^p(X, K)$ onto itself, and let $\{e_n: n = 1, 2, \cdots\}$ be some fixed orthonormal basis for K. Then there exists a regular set isomorphism Φ of the σ-algebra Σ of measurable sets onto itself (defined modulo null sets), a scalar function h defined on X satisfying* (5), *and a weakly measurable operator-valued function U defined on X, where $U(x)$ is an isometry of K onto itself for almost all $x \in X$, such that for $F \in L^p(X, K)$,*

$$(T(F))(x) = U(x)h(x)(\Phi(F))(x) ,$$

where $\Phi(F)$ is defined by (2). *Conversely, every map T of this form is an isometry of $L^p(X, K)$ onto itself.*

Proof. If T is of this form, then it follows from (3) and the fact that $U(x)$ is almost everywhere an isometry, that

$$\| U(x)h(x)(\Phi(F))(x) \| = | h(x) | \, | \Phi(\| F \|) | (x) , \quad \text{for} \quad F \in L^p(X, K) ,$$

so that T is norm-preserving by Lamperti's result for the scalar case. The fact that T maps $L^p(X, K)$ onto itself can readily be established, for example, by noting that since Φ is onto, and $U(x)$ is a.e. an isometry of K onto K, no nonzero element of $L^q(X, K)$ can annihilate the range of T.

Now suppose that T is any isometry of $L^p(X, K)$ onto itself. We define $U(x)$ on the basis vectors e_n of K by $U(x)e_n = E_n(x)$, where the E_n are determined as in Lemma 7, and then extend $U(x)$ linearly to K. Since by Lemma 5, $\{E_n(x): n = 1, 2, \cdots\}$ is almost everywhere an orthonormal set in K, $U(x)$ is an isometry of K into itself a.e., and if K is of finite dimension, the remaining assertions of the theorem then follow immediately from Lemma 7.

Thus we may as well assume that K is infinite dimensional. Let $F(x) = \sum_n f_n(x)e_n$ belong to $L^p(X, K)$. Then the sequence $\{F_N\}$, where $F_N(x) = \sum_{n=1}^N f_n(x)e_n$, converges a.e. to F and is dominated by $\| F \|$. Hence by the dominated convergence theorem, $\| F_N - F \|_p \to 0$. We thus have $T(F) = \lim_N T(F_N)$ in $L^p(X, K)$, and so at least a subsequence of the $T(F_N)$ converges a.e. to $T(F)$. But we know from (3) and the fact that $U(x)$ is almost everywhere norm-preserving that $U(x)h(x)(\Phi(F))(x) = \lim_N U(x)h(x)(\Phi(F_N))(x) = \lim_N (T(F_N))(x)$ exists in K for almost all $x \in X$, and thus it follows that $(T(F))(x) = U(x)h(x)(\Phi(F))(x)$, as claimed. Finally, since the elements of $T(L^p(X, K))$ take their values a.e. in the range of $U(x)$, and since T is onto, $U(x)$ must map K onto K for almost all $x \in X$.

3. Remarks and problems. (i) Throughout we have assumed that the measure space is finite, but the theorem is also valid for σ-finite measure spaces, and the generalization to this latter case is largely straightforward. We say "largely" only because there are a few modifications (other than the obvious ones) of statements and proofs necessary for the σ-finite case, whose necessity might easily be overlooked. For example, if the space is σ-finite, a suitable reformulation of Lemma 4 is the following:

Let A be a measurable subset of X with finite positive measure and let e be an element of K with $\|e\| = 1$. If $T(\chi_A e) = F$, and if E is that vector function defined by $E(x) = F(x)/\|F(x)\|$ if x belongs to the support of F, and $E(x) = 0$ otherwise, then $T^{*-1}(\chi_A e)$ is determined by $(T^{*-1}(\chi_A e))(x) = \|F(x)\|^{p-1}E(x)$, for almost all $x \in X$.

The proof of this fact is analogous to that given for Lemma 4, provided $p > 1$. However, in the case $p = 1$, additional arguments, unnecessary if $\mu(X)$ is finite, have to be introduced.

(ii) For a certain class of measure spaces, the set isomorphism Φ may, of course, be repleaced by a measurable point mapping [5, Chap. 15].

(iii) In [4], Lamperti provides a description of all isometries of $L^p(X, S)$ *into* itself, not just the surjective ones. One may ask if such a description is attainable in the vector case. The type of argument needed would presumably differ substantially from that used here, since we often rely on the existence of the mapping T^{*-1} from $L^q(X, K)$ to itself.

(iv) Can a reasonable description of the isometries be obtained if the Hilbert space K is replaced by a suitable class of Banach spaces? In particular, it might be of interest to see if K can be replaced by an arbitrary finite dimensional Banach space.

REFERENCES

1. S. Banach, *Théorie des opération linéaires*, Monografje Matematyczne, Warsaw, 1932.
2. N. Dinculeanu, *Vector Measures*, Pergaman Press, New York, 1967.
3. J. L. Doob, *Stochastic Processes*, Wiley and Sons, New York, 1953.
4. J. Lamperti, *On the isometries of certain function spaces*, Pacific J. Math., **8** (1958), 459-466.
5. H. L. Royden, *Real Analysis*, 2nd ed., Macmillan Co., New York, 1968.

Received July 3, 1973.

UNIVERSITY OF CALIFORNIA, SANTA BARBARA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

TWO RELATED INTEGRALS OVER SPACES OF CONTINUOUS FUNCTIONS

R. H. CAMERON AND D. A. STORVICK

In this paper the authors evaluate Yeh-Wiener integrals (which apply to functionals of a variable continuous function of two arguments) in terms of multiple Wiener integrals (which apply to functionals of several variable continuous functions of one argument). First somewhat specialized cases are given where the multiplicity of the Wiener integral is finite, and then quite general Yeh-Wiener integrals are evaluated in terms of limits of n-fold Wiener integrals as $n \to \infty$.

Introduction. James Yeh [5][1] defined Wiener measure in the space $C_2[S]$ of continuous real valued functions of two variables defined on the square S: $0 \leqq s \leqq 1$, $0 \leqq t \leqq 1$ and vanishing whenever s or t equals zero. More recently James Kuelbs [3, 4] extended Yeh's integral to integration over $C_2[X]$, the space of continuous real valued functions on any compact subset X of the plane. Kuelbs also defined a similar integral over spaces of functions of several variables and even infinitely many variables [4].

In the present paper we shall consider integration over $C_2[X]$ in the case where X is the rectangle $R = \{(s, t) \mid a \leqq s \leqq b, \alpha \leqq t \leqq \beta\}$. We note that this is closely connected with Yeh's integral over $C_2[S]$ and that Kuelbs has given a formula for relating integrals over $C_2[R]$ with integrals over $C_2[S]$, [3, p. 18].

Yeh's measure as applied to the space

$$C_2[R] \equiv \{x(\cdot, \cdot) \mid x(a, t) = x(s, \alpha) = 0, x(s, t) \text{ continuous for } a \leqq s \leqq b, \alpha \leqq t \leqq \beta\}$$

is defined as follows. Let $a = s_0 < s_1 < \cdots < s_m = b$, and $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ be subdivisions of $[a, b]$ and $[\alpha, \beta]$ respectively and let $-\infty \leqq P_{j,k} \leqq Q_{j,k} \leqq +\infty$ be given for $j = 1, \cdots, m$ and $k = 1, \cdots, n$. Then

$$I = \{x \in C_2[R] \mid P_{j,k} < x(s_j, t_k) \leqq Q_{j,k} \text{ for } j = 1, \cdots, m, k = 1, \cdots, n\}$$

will be called an "interval" in $C_2[R]$. He defines the measure of the interval I by

[1] See also reference to Kitagawa in [5].

$$m(I) = \pi^{-mn/2}[(s_1 - s_0) \cdots (s_m - s_{m-1})]^{-n/2}[(t_1 - t_0) \cdots (t_n - t_{n-1})]^{-m/2}$$
$$\cdot \int_{P_{m,n}}^{Q_{m,n}} \overset{(mn)}{\cdots} \int_{P_{1,1}}^{Q_{1,1}} \exp\left\{-\sum_{j=1}^{m}\sum_{k=1}^{n} \frac{[u_{j,k}-u_{j-1,k}-u_{j,k-1}+u_{j-1,k-1}]^2}{(s_j-s_{j-1})(t_k-t_{k-1})}\right\} du_{1,1} \cdots du_{m,n}$$

where $u_{0,k} \equiv u_{j,0} \equiv 0$ for $j = 1, \cdots, m$; $k = 1, \cdots, n$.

This measure is countably additive on the set of intervals in $C_2[R]$ and can be extended in the usual way to the sigma-algebra of sets generated by the intervals and can then be further extended so as to be a complete measure. Thus "Yeh-Wiener measurable set" and its "measure" are defined in $C_2[R]$.

The integrals of functionals integrable with respect to this measure will be called "Yeh-Wiener integrals".

In Theorem 1 of the present paper we establish a formula for evaluating in terms of a Wiener integral the Yeh-Wiener integral of a functional of $x(\cdot, \cdot)$ which actually depends solely on the values of x on one horizontal line.

Theorem 2 treats the case of a functional depending only on the values of x on a finite number of horizontal lines.

Theorem 4 deals with the case of a functional depending only on the values of x on the two (perpendicular) free edges of R. Examples are given to show how Theorem 4 can be used to evaluate Yeh-Wiener integrals of specific functionals.

Finally in Theorem 5 we consider a class of functionals that may depend on the values which x assumes at all points of the rectangle R and not only on the values x assumes on some restricted set.

1. The one line theorem. Let $C_1[a, b] \equiv \{y(\cdot) \mid y(a) = 0,\ y(t)$ continuous on $[a, b]\}$, let $R \equiv [a, b] \times [\alpha, \beta]$ and let

$$C_2[R] \equiv \{x(\cdot, \cdot) \mid x(a, t) = x(s, \alpha) = 0,\ x(s, t) \text{ continuous for } a \leqq s \leqq b,\ \alpha \leqq t \leqq \beta\}.$$

THEOREM 1. *Let $\alpha < \gamma \leqq \beta$, and let $f(\cdot)$ be a real or complex valued functional defined on $C_1[a, b]$ such that $f(\sqrt{(\gamma - \alpha)/2}\, y)$ is a Wiener measurable functional of y on $C_1[a, b]$. Then $f(x(\cdot, \gamma))$ is a Yeh-Wiener measurable functional of $x(\cdot, \cdot)$ on $C_2[R]$ and*

$$\int_{C_2[R]} f(x(\cdot, \gamma))dx = \int_{C_1[a,b]} f\left(\sqrt{\frac{\gamma - \alpha}{2}}\, y\right)dy, \tag{1.1}$$

where the existence of either integral implies the existence of the other and their equality.

Proof. Let $g(y) \equiv f(\sqrt{(\gamma - \alpha)/2}\, y)$. Then it suffices to prove that $g(\sqrt{2/(\gamma - \alpha)}\, x(\cdot, \gamma))$ is Yeh-Wiener measurable and that

$$\int_{C_2[R]} g\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)dx = \int_{C_1[a,b]} g(y)dy \tag{1.2}$$

where the existence of either implies the existence of the other and their equality.

Case I. Let us consider a subdivision $a = s_0 < s_1 < \cdots < s_m = b$ and let $g(y) = \chi_I(y)$ where I is the interval

$$I = \{y \in C_1[a, b] \mid -\infty \leqq z_i < y(s_i) \leqq w_i \leqq +\infty,\ i = 1, \cdots, m\}$$

so that

$$g\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big) = \chi_I\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big) = \chi_K(x(\cdot,\cdot))$$

where

$$K = \Big\{x \in C_2[R] \mid -\infty \leqq \sqrt{\frac{\gamma-\alpha}{2}}z_i < x(s_i,\gamma) \leqq \sqrt{\frac{\gamma-\alpha}{2}}w_i \leqq +\infty,\ i, \cdots, m\Big\}.$$

Thus in this case, $g(\sqrt{2/(\gamma-\alpha)}x(\cdot,\gamma))$ is Yeh-Wiener measurable on $C_2[R]$ (see Definition (2.1) of [4, p. 434]).

Because $g(\sqrt{2/(\gamma-\alpha)}x(\cdot,\gamma))$ is the characteristic functional of an interval, the left member of equation (1.2) equals the measure of the interval K, i.e.,

$$\int_{C_2[R]} g\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)dx = \int_{C_2[R]} \chi_K(x(\cdot,\cdot))dx$$

$$= [\pi^m(s_1-s_0)(s_2-s_1)\cdots(s_m-s_{m-1})(\gamma-\alpha)^m]^{-1/2}\int_{\sqrt{(\gamma-\alpha)/2}z_m}^{\sqrt{(\gamma-\alpha)/2}w_m} \overset{(m)}{\cdots} \int_{\sqrt{(\gamma-\alpha)/2}z_1}^{\sqrt{(\gamma-\alpha)/2}w_1}$$

$$\cdot \exp\Big\{-\sum_{i=1}^m \frac{(u_i-u_{i-1})^2}{(s_i-s_{i-1})(\gamma-\alpha)}\Big\}du_1\cdots du_m\,,$$

where $u_0 \equiv 0$.

The right hand member of (1.2) can be evaluated in the following manner,

$$\int_{C_1[a,b]} g(y)dy = \int_{C_1[a,b]} \chi_I(y)dy = [(2\pi)^m(s_1-s_0)\cdots(s_m-s_{m-1})]^{-1/2}\int_{z_m}^{w_m} \overset{(m)}{\cdots} \int_{z_1}^{w_1}$$

$$\cdot \exp\Big\{-\sum_{i=1}^m \frac{(v_i-v_{i-1})^2}{2(s_i-s_{i-1})}\Big\}dv_1\cdots dv_m\,,$$

where $v_0 \equiv 0$. If we set $v_i = \sqrt{2/(\gamma-\alpha)}u_i$ we obtain (1.2) and hence (1.1).

Case II. Let $g(y) = \chi_\Omega(y)$ where Ω is the union of the disjoint intervals $I_1, I_2, \cdots$. Then by Case I, we have

$$\int_{C_2[R]} \chi_{I_k}\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)dx = \int_{C_1[a,b]} \chi_{I_k}(y)dy$$

including the measurability of the left hand integrand. The functional obtained by summing over k is Yeh-Wiener measurable, i.e.,

$$\sum_{k=1}^{\infty} \chi_{I_k}\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big) \equiv \chi_\Omega\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)$$

is Yeh-Wiener measurable. Then summing the integrals we have

$$\int_{C_2[R]} \chi_\Omega\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)dx = \int_{C_1[a,b]} \chi_\Omega(y)dy .$$

Thus (1.2) holds in this case.

Case III. Let $g(y) = \chi_\Delta(y)$ where Δ is a countable intersection of sets Ω of the type considered in Case II. Since finite intersections of such sets are of the same type, we can set

$$\Delta = \bigcap_{k=1}^{\infty} \Omega_k$$

where $\Omega_1 \supset \Omega_2 \supset \Omega_3 \supset \cdots$ and each Ω_k is of the type considered in Case II. Thus

$$g(y) = \lim_{k\to\infty} \chi_{\Omega_k}(y) ,$$

and $g(y)$ is Yeh-Wiener measurable. If we now apply (1.2) to χ_{Ω_k} and take limits we obtain (1.2) for $g(y) = \chi_\Delta(y)$, including the measurability of $g(\sqrt{2/(\gamma-\alpha)}x(\cdot,\gamma))$.

Case IV. Let $g(y) = \chi_N(y)$ where N is a Wiener null set. Let N_1 be a Wiener null set of the type discussed in Case III such that $N_1 \supset N$. Then (1.2) holds for $\chi_{N_1}(y)$ and we have

$$\int_{C_2[R]} \chi_{N_1}\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)dx = \int_{C_1[a,b]} \chi_{N_1}(y)dy = 0 ,$$

including the measurability of the left hand integrand which we now know to be Yeh-Wiener almost everywhere zero. Thus

$$\chi_N\Big(\sqrt{\frac{2}{\gamma-\alpha}}x(\cdot,\gamma)\Big)$$

is also Yeh-Wiener almost everywhere zero and (1.2) holds.

Case V. Let $g(y) = \chi_E(y)$ where E is any Wiener measurable set. Then $E = \Delta - N$ where Δ and N are sets of the type considered in

Cases III and IV. By applying (1.2) to Δ and to N we obtain (1.2) for E including the measurability of the left hand integrand.

Case VI. Let $g(y)$ be a simple functional (with respect to Wiener measure). Then $g(y)$ is a linear combination with constant coefficients of a finite number of functionals of the type considered in Case V. Hence (1.2) holds.

Case VII. Let $g(y)$ be a real nonnegative Wiener measurable functional. Then $g(y)$ is the limit of a monotone increasing sequence of simple functionals and (1.2) follows from Case VI by monotone convergence.

Case VIII. General case: Because any complex valued functional can be decomposed into its real and imaginary parts and they into their positive and negative parts, the theorem is proved.

2. **The n-parallel lines theorem.** Having obtained a formula for Yeh-Wiener integrals where the functional of $x(\cdot, \cdot)$ actually depends only on the values of $x(\cdot, \gamma)$, i.e., on the values of x on one horizontal line of the fundamental rectangle R, it is natural to inquire next concerning functionals that depend solely on the values of x on a finite number of horizontal lines, i.e., functionals of the form

$$F(x) = f[x(\cdot, t_1), x(\cdot, t_2), \cdots, x(\cdot, t_n)] \,. \tag{2.0}$$

One might expect to obtain the Yeh-Wiener integral of F as an n-fold Wiener integral over the product of n Wiener spaces. Since it is not immediately apparent what the formula should be, we begin with the case where f depends on the values of the $y_k(\cdot)$ at a finite number of points. Thus we let

$$\begin{aligned} &f[y_1, \cdots, y_n] \\ &\quad = \varphi[y_1(s_1), \cdots, y_1(s_m); y_2(s_1), \cdots, y_2(s_m); \cdots; y_n(s_1), \cdots, y_n(s_m)] \end{aligned} \tag{2.1}$$

where

$$\varphi(u_{1,1}, u_{2,1}, \cdots, u_{m,1}; \cdots; u_{1,n}, \cdots, u_{m,n}) \equiv \varphi(U)$$

is defined on R^{mn} and U denotes the rectangular array $\{u_{i,j}\}_{i=1,\cdots,m; j=1,\cdots,n}$. Then from (2.0) and (2.1) we have

$$F(x) = \varphi(\{x(s_i, t_j)\}_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}}) \,. \tag{2.2}$$

Integrating over $C_2[R]$ and evaluating the Yeh-Wiener integral we have

$$\int_{C_2[R]} F(x)dx = \left[\pi^{mn} \prod_{i=1}^{m}{}' \prod_{j=1}^{n} (s_i - s_{i-1})(t_j - t_{j-1})\right]^{-1/2}$$

$$\int_{R^{mn}} \varphi(U) \exp\left\{-\sum_{i=1}^{m}\sum_{j=1}^{n} \frac{(u_{i,j} - u_{i-1,j} - u_{i,j-1} + u_{i-1,j-1})^2}{(s_i - s_{i-1})(t_j - t_{j-1})}\right\} dU \tag{2.3}$$

where $dU = du_{1,1} \cdots du_{m,n}$, where $u_{0,j} \equiv u_{i,0} \equiv 0$.

We now make the transformation

$$v_{i,j} = \sqrt{\frac{2}{t_j - t_{j-1}}}(u_{i,j} - u_{i,j-1})$$

so

$$u_{i,j} = \sqrt{\frac{t_1 - \alpha}{2}} v_{i,1} + \sqrt{\frac{t_2 - t_1}{2}} v_{i,2} + \cdots + \sqrt{\frac{t_j - t_{j-1}}{2}} v_{i,j}$$

and obtain

$$\begin{aligned} \int_{C_2[R]} F(x)dx &= \left[(2\pi)^m \prod_{i=1}^{m} (s_i - s_{i-1})\right]^{-n/2} \\ &\int_{R^{mn}} \varphi\left(\left\{\sqrt{\frac{t_1 - \alpha}{2}} v_{i,1} + \sqrt{\frac{t_2 - t_1}{2}} v_{i,2} + \cdots \right.\right. \\ &\left.\left. + \sqrt{\frac{t_j - t_{j-1}}{2}} v_{i,j}\right\}_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}}\right) \cdot \end{aligned} \tag{2.4}$$

$$\prod_{j=1}^{n} \exp\left\{-\frac{1}{2}\sum_{i=1}^{m} \frac{(v_{i,j} - v_{i-1,j})^2}{(s_i - s_{i-1})}\right\} dV, \quad \text{where} \quad dV = dv_{1,1} \cdots dv_{m,n}, \quad \text{where} \quad v_{0,j} \equiv 0 .$$

For each fixed j, the sums in the exponential are those which would occur in the evaluation of a Wiener integral, and so we see that the whole expression is the evaluation of an n-fold Wiener integral. Thus

$$\begin{aligned} \int_{C_2[R]} F(x)dx &= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} \varphi\left(\left\{\sqrt{\frac{t_1 - \alpha}{2}} y_1(s_i) \right.\right. \\ &\left.\left. + \sqrt{\frac{t_2 - t_1}{2}} y_2(s_i) + \cdots + \sqrt{\frac{t_j - t_{j-1}}{2}} y_j(s_i)\right\}_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}}\right) dy_1 \cdots dy_n . \end{aligned} \tag{2.5}$$

We shall use the following notation for the cartesian product of n Wiener spaces $\overset{n}{\times} C_1[a, b] \equiv C_1[a, b] \times \overset{(n)}{\cdots} \times C_1[a, b]$.

We have given the motivation for the following theorem:

THEOREM 2. *Let* $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ *and let* $f[y_1, \cdots, y_n]$ *be a real or complex valued functional defined on* $\overset{n}{\times} C_1[a, b]$ *such that*

$$f\left[\sqrt{\frac{t_1-t_0}{2}}y_1, \sqrt{\frac{t_1-t_0}{2}}y_1 + \sqrt{\frac{t_2-t_1}{2}}y_2, \cdots, \sqrt{\frac{t_1-t_0}{2}}y_1 + \cdots + \sqrt{\frac{t_n-t_{n-1}}{2}}\, y_n\right] \tag{2.6}$$

is a Wiener measurable functional of $(y_1, \cdots, y_n)$ *on* $\overset{n}{\times} C_1[a, b]$. *Then* $f[x(\cdot, t_1), \cdots, x(\cdot, t_n)]$ *is a Yeh-Wiener measurable functional of* $x(\cdot, \cdot)$ *on* $C_2[R]$ *and*

$$\begin{aligned} &\int_{C_2[R]} f[x(\cdot, t_1), \cdots, x(\cdot, t_n)]dx \\ &= \int_{\overset{n}{\times} C_1[a,b]} f\left[\sqrt{\frac{t_1-t_0}{2}}y_1, \sqrt{\frac{t_1-t_0}{2}}y_1 + \sqrt{\frac{t_2-t_1}{2}}y_2, \cdots, \sqrt{\frac{t_1-t_0}{2}}y_1 + \cdots \right. \\ &\qquad \left. + \sqrt{\frac{t_n-t_{n-1}}{2}}y_n\right]d(y_1 \times \cdots \times y_n) \end{aligned} \tag{2.7}$$

where the existence of either integral implies the existence of the other and their equality.

Proof.[2] Let

$$g(y_1, \cdots, y_n) = f\left(\sqrt{\frac{t_1-t_0}{2}}y_1, \cdots, \sum_{k=1}^{n} \sqrt{\frac{t_k-t_{k-1}}{2}}y_k\right).$$

Making the substitution $z_\nu = \sum_{k=1}^{\nu} \sqrt{(t_k - t_{k-1})/2}\, y_k$, we have

$$f(z_1, \cdots, z_n) = g\left(\sqrt{\frac{2}{t_1-t_0}}\, (z_1 - z_0), \cdots, \sqrt{\frac{2}{t_n-t_{n-1}}}(z_n - z_{n-1})\right).$$

Thus it suffices to prove that if $g(y_1, \cdots, y_n)$ is a Wiener measurable functional of $(y_1, \cdots, y_n)$, then

$$g\left[\sqrt{\frac{2}{t_1-t_0}}x(\cdot, t_1), \cdots, \sqrt{\frac{2}{t_n-t_{n-1}}}(x(\cdot, t_n) - x(\cdot, t_{n-1}))\right] \tag{2.8}$$

is a Yeh-Wiener measurable functional on $C_2[R]$ and

$$\begin{aligned} &\int_{C_2[R]} g\left[\sqrt{\frac{2}{t_1-t_0}}x(\cdot, t_1), \cdots, \sqrt{\frac{2}{t_n-t_{n-1}}}(x(\cdot, t_n) - x(\cdot, t_{n-1}))\right]dx \\ &= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} g(y_1, \cdots, y_n)dy_1 \cdots dy_n \,. \end{aligned} \tag{2.9}$$

Case I. Let $g(y_1, \cdots, y_n) = \chi_I(y_1, \cdots, y_n)$, where I is the interval

[2] The proof has to proceed in the opposite order from the motivation because of the measurability argument.

$I = \{(y_1, \cdots, y_n) \in \overset{n}{\times} C_1[a, b] \mid -\infty \leqq z_{j,k} < y_k(s_j) \leqq w_{j,k} \leqq +\infty$, for $j = 1, \cdots, m$, $k = 1, \cdots, n\}$. Clearly $I = I_1 \cap I_2 \cap \cdots \cap I_m$ where $I_j = \{(y_1, \cdots, y_n) \in \overset{n}{\times} C_1[a, b] \mid -\infty \leqq z_{j,k} < y_k(s_j) \leqq w_{j,k} \leqq +\infty$, for $k = 1, \cdots, n\}$. Now

$$\begin{aligned} & g\left[\sqrt{\frac{2}{t_1 - t_0}}x(\cdot, t_1), \cdots, \sqrt{\frac{2}{t_n - t_{n-1}}}(x(\cdot, t_n) - x(\cdot, t_{n-1}))\right] \\ & = \prod_{j=1}^{m} \chi_{I_j}\left[\sqrt{\frac{2}{t_1 - t_0}}x(\cdot, t_1), \cdots, \sqrt{\frac{2}{t_n - t_{n-1}}}(x(\cdot, t_n) - x(\cdot, t_{n-1}))\right] \\ & = \prod_{j=1}^{m} \chi_{K_j}(x(\cdot, \cdot)) \end{aligned} \tag{2.10}$$

where

$$K_j = \left\{x \in C_2[R] \mid -\infty \leqq \sqrt{\frac{t_k - t_{k-1}}{2}} z_{j,k} < [x(s_j, t_k) - x(s_j, t_{k-1})] \right.$$
$$\left. \leqq \sqrt{\frac{t_k - t_{k-1}}{2}} w_{j,k} \leqq +\infty \quad \text{for} \quad k = 1, \cdots, n\right\},$$

and

$$\chi_{K_j}(x(\cdot, \cdot)) = \chi_{L_j}[x(s_j, t_1), \cdots, x(s_j, t_n)], \tag{2.11}$$

where

$$L_j = \left\{[u_{j,1}, \cdots, u_{j,n}] \in R^n \mid -\infty \leqq \sqrt{\frac{t_k - t_{k-1}}{2}} z_{j,k} < u_{j,k} - u_{j,k-1} \right.$$
$$\left. \leqq \sqrt{\frac{t_k - t_{k-1}}{2}} w_{j,k} \leqq +\infty \quad \text{for} \quad k = 1, \cdots, n\right\}.$$

Thus in this case (2.8) is Yeh-Wiener measurable on $C_2[R]$ since χ_{L_j} is a Lebesgue measurable function in R^{mn}. Integrating the expression (2.8) we obtain by using (2.10) and (2.11),

$$\begin{aligned} G & \equiv \int_{C_2[R]} g\left[\sqrt{\frac{2}{t_1 - t_0}}x(\cdot, t_1), \cdots, \sqrt{\frac{2}{t_n - t_{n-1}}}(x(\cdot, t_n) - x(\cdot, t_{n-1}))\right] dx \\ & = \int_{C_2[R]} \prod_{j=1}^{m} \chi_{L_j}[x(s_j, t_1), \cdots, x(s_j, t_n)] dx \\ & = \pi^{-(mn)/2}[(s_1 - s_0) \cdots (s_m - s_{m-1})]^{-n/2}[(t_1 - t_0) \cdots (t_n - t_{n-1})]^{-m/2} \\ & \quad \cdot \int_{-\infty}^{\infty} \overset{(mn)}{\cdots} \int_{-\infty}^{\infty} \prod_{j=1}^{m} \chi_{L_j}[u_{j,1}, \cdots, u_{j,n}] \\ & \quad \cdot \exp\left\{-\sum_{j=1}^{m} \sum_{k=1}^{n} \frac{(u_{j,k} - u_{j,k-1} - u_{j-1,k} + u_{j-1,k-1})^2}{(s_j - s_{j-1})(t_k - t_{k-1})}\right\} dU, \end{aligned}$$

where $u_{j,0} \equiv u_{0,k} \equiv 0$. If we set $v_{j,k} = \sqrt{2/(t_k - t_{k-1})}(u_{j,k} - u_{j,k-1})$ so that $u_{j,k} = \sum_{i=1}^{k} \sqrt{(t_i - t_{i-1})/2}\, v_{j,i}$, we obtain

$$G \equiv (2\pi)^{-(mn)/2}[(s_1 - s_0) \cdots (s_m - s_{m-1})]^{-n/2}$$
$$\cdot \int_{-\infty}^{\infty} \overset{(mn)}{\cdots} \int_{-\infty}^{\infty} \prod_{j=1}^{m} \chi_{L_j}\Big(\sqrt{\frac{t_1 - t_0}{2}} v_{j,1}, \cdots, \sum_{i=1}^{k} \sqrt{\frac{t_i - t_{i-1}}{2}} v_{j,i}, \cdots,$$
$$\cdot \sum_{i=1}^{n} \sqrt{\frac{t_i - t_{i-1}}{2}} v_{j,i}\Big) \exp\Big\{- \sum_{j=1}^{m} \sum_{k=1}^{n} \frac{(v_{j,k} - v_{j-1,k})^2}{2(s_j - s_{j-1})}\Big\} dv_{1,1} \cdots dv_{m,n}$$
$$= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} \prod_{j=1}^{m} \chi_{I_j}(y_1(s_j), \cdots, y_i(s_j), \cdots, y_n(s_j)) dy_1 \cdots dy_n$$
$$= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} g(y_1, \cdots, y_n) dy_1 \cdots dy_n$$

and Case I is proved. The remaining cases are analogous to those of Theorem 1 and are proved in the same way.

3. **The orthogonal transformation.** Theorem 2 which we have just proved gives us an evaluation of the Yeh-Wiener integral of a functional $F(x(\cdot, \cdot))$ which depends only on the values of x on n parallel lines. It is natural to inquire next concerning functionals that depend solely on the values of x on two perpendicular lines. We shall limit our investigation in this paper to the case where the two perpendicular lines are the free edges of the fundamental rectangle. Before we can obtain such a theorem, we will need to establish a generalization of Bearman's theorem [1, 130] on rotations in the product of two Wiener spaces. (A theorem of this sort was once proved by Edwin Sheffield, but so far as the authors know, it was never published.)

THEOREM 3. *Let $F(y_1, \cdots, y_n)$ be any Wiener integrable functional of $y_1(\cdot), \cdots, y_n(\cdot)$ on $\overset{n}{\times} C_1[a, b]$ and let $(c_{i,j})_{i,j=1,\cdots,n}$ be a real orthogonal matrix (so that $\sum_{k=1}^{n} c_{i,k}c_{j,k} = \delta_{i,j}$ for $i, j = 1, \cdots, n$). Then the transformation*

$$y_i(\cdot) = \sum_{j=1}^{n} c_{i,j} z_j(\cdot) \quad \textit{for} \quad i = 1, \cdots, n \tag{3.0}$$

is a measure preserving transformation of $\overset{n}{\times} C_1[a, b]$ onto itself. Moreover,

$$\begin{aligned}\int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} & F(y_1, \cdots, y_n) dy_1 \cdots dy_n \\ = \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} & F\Big(\sum_{j=1}^{n} c_{1,j} z_j, \cdots, \sum_{j=1}^{n} c_{n,j} z_j\Big) dz_1 \cdots dz_n .\end{aligned} \tag{3.1}$$

Proof. Case I. Let F depend only on the values of $y_1, \cdots, y_n$ at a certain finite set of points, $a = s_0 < s_1 < \cdots < s_m = b$, i.e., let

$$\begin{aligned}(3.2)\qquad F(y_1, \cdots y_n) = f(y_1(s_1), \cdots y_1(s_m); y_2(s_1), \cdots ,\\ y_2(s_m); \cdots; y_n(s_1), \cdots, y_n(s_m)) ,\end{aligned}$$

where $f(u_{1,1}, \cdots u_{1,m}; \cdots; u_{n,1}, \cdots, u_{n,m})$ is a bounded measurable function of its nm arguments. It is clear that F is Wiener measurable and bounded on $\overset{n}{\times} C_1[a, b]$. Now we have

$$\begin{aligned}I &\equiv \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} F(y_1, \cdots, y_n) dy_1 \cdots dy_n\\ (3.3)\quad &= (2\pi)^{-(mn)/2}[(s_1 - a) \cdots (s_m - s_{m-1})]^{-n/2} \int_{-\infty}^{\infty} \overset{(mn)}{\cdots} \int_{-\infty}^{\infty}\\ &\quad \cdot f(u_{1,1}, \cdots, u_{n,m}) \exp\left\{- \sum_{i=1}^{n}\sum_{k=1}^{m} \frac{(u_{i,k} - u_{i,k-1})^2}{2(s_k - s_{k-1})}\right\} du_{1,1} \cdots du_{n,m} ,\end{aligned}$$

where $u_{i,0} \equiv 0$.

Let us make the transformation $u_{i,k} = \sum_{j=1}^{n} c_{i,j} v_{j,k}$ where $i = 1, \cdots, n$ and $k = 1, \cdots, m$, to obtain

$$\begin{aligned}I &= (2\pi)^{-(mn)/2}[(s_1 - a) \cdots (s_m - s_{m-1})]^{-n/2} \int_{-\infty}^{\infty} \overset{(mn)}{\cdots} \int_{-\infty}^{\infty}\\ &\quad \cdot f\left(\sum_{j=1}^{n} c_{1,j} v_{j,1}, \cdots, \sum_{j=1}^{n} c_{n,j} v_{j,m}\right)\\ &\quad \cdot \exp\left\{- \sum_{k=1}^{m}\sum_{i=1}^{n} \frac{\left(\sum_{j=1}^{n} c_{i,j}(v_{j,k} - v_{j,k-1})\right)^2}{2(s_k - s_{k-1})}\right\} dv_{1,1} \cdots dv_{n,m} .\end{aligned}$$

Since $(c_{i,j})$ is an orthogonal matrix,

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n} c_{i,j}(v_{j,k} - v_{j,k-1})\right)^2 = \sum_{j=1}^{n} (v_{j,k} - v_{j,k-1})^2$$

and we obtain

$$\begin{aligned}I &= (2\pi)^{-(mn)/2}[(s_1 - a) \cdots (s_m - s_{m-1})]^{-n/2} \int_{-\infty}^{\infty} \overset{(mn)}{\cdots} \int_{-\infty}^{\infty}\\ &\quad \cdot f\left(\sum_{j=1}^{n} c_{1,j} v_{j,1}, \cdots, \sum_{j=1}^{n} c_{n,j} v_{j,m}\right) \exp\left\{- \sum_{k=1}^{m}\sum_{j=1}^{n} \frac{(v_{j,k} - v_{j,k-1})^2}{2(s_k - s_{k-1})}\right\}\\ (3.4)\quad &\quad \cdot dv_{1,1} \cdots dv_{n,m}\\ &= \int_{C_1[a,b]} \cdots \int_{C_1[a,b]} f\left(\sum_{j=1}^{n} c_{1,j} z_j(s_1), \cdots, \sum_{j=1}^{n} c_{n,j} z_j(s_m)\right) dz_1 \cdots dz_n\\ &= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} F\left(\sum_{j=1}^{n} c_{1,j} z_j(\cdot), \cdots, \sum_{j=1}^{n} c_{n,j} z_j(\cdot)\right) dz_1 \cdots dz_n .\end{aligned}$$

In the above argument, the measurability of each successive integrand follows from the measurability of $f(u_{1,1}, \cdots, u_{n,m})$, and the boundedness of f implies the integrability of each integrand. Thus (3.1) is established for Case I. If we apply (3.1) to the case where f is a characteristic function of a measurable set we observe that (3.0) is a measure preserving transformation of $\overset{n}{\times} C_1[a, b]$ onto itself.

Case II. Let $F(y_1, \cdots, y_n) = \chi_\Omega(y_1, \cdots, y_n)$, where Ω is the union of a countable disjoint set of intervals $\Omega = \bigcup_{j=1}^\infty I_j$ and each I_j is an interval in the product space $\overset{n}{\times} C_1[a, b]$, (as in the proof of Theorem 2, Case I). Because each χ_{I_j} satisfies the hypothesis of Case I, the theorem holds when F is of the form $F(y_1, \cdots, y_n) = \chi_{I_j}(y_1, \cdots, y_n)$. Since Ω is the countable union of measurable sets, it is measurable, and by summing both sides of (3.1) applied to χ_{I_j} we obtain (3.1) applied to χ_Ω.

Case III. Let $F = \chi_E(y_1, \cdots, y_n)$ where E is a Wiener measurable set in $\times^n C_1[a, b]$. The result of Case II can be extended from $\Omega = \bigcup_{j=1}^\infty I_j$ to countable intersections of sets of this form and then to null sets and then to general measurable sets in the usual way.

Case IV. Let F be a nonnegative functional. If F is actually a simple functional the result follows from Case III by multiplication by constants and addition. If F is not a simple functional, it can be expressed as a limit of a monotone increasing sequence of simple functionals, and the theorem follows for this case.

Case V. General Case: If F is real, we write $F = F^+ - F^-$ and apply Case IV to F^+ and to F^- and thus establish the theorem for real functionals. The extension to complex functionals is immediate.

4. **The two perpendicular lines theorem.** We now proceed to establish a formula for the evaluation of the Yeh-Wiener integral of a functional that depends solely on the values of x on two perpendicular lines.

THEOREM 4. *Let $f(z, y)$ defined on $C_1[a, b] \times C_1[\alpha, \beta]$ be a functional such that*

$$f\left(\sqrt{\frac{\beta-\alpha}{2}}z,\ [(\cdot)-\alpha]\left[\sqrt{\frac{b-a}{2}}\int_{(\cdot)}^{\beta}\frac{dy(\tau)}{\tau-\alpha}+\frac{z(b)}{\sqrt{2(\beta-\alpha)}}\right]\right) \tag{4.0}$$

is Wiener measurable on $C_1[a, b] \times C_1[\alpha, \beta]$. Then it follows that $f[x(\cdot, \beta), x(b, \cdot)]$ is Yeh-Wiener measurable on $C_2[R]$, where $R \equiv [a, b] \times [\alpha, \beta]$. Moreover,

$$\begin{aligned} &\int_{C_2[R]} f[x(\cdot, \beta), x(b, \cdot)]dx \\ &= \int_{C_1[a,b]\times C_1[\alpha,\beta]} f\left\{\sqrt{\frac{\beta-\alpha}{2}}z,\ [(\cdot)-\alpha]\left[\sqrt{\frac{b-a}{2}}\int_{(\cdot)}^{\beta}\frac{dy(\tau)}{\tau-\alpha}\right.\right. \\ &\quad \left.\left. + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\right]\right\} d(z \times y) \end{aligned} \tag{4.1}$$

where the existence of either member implies the existence of the other and their equality.

Proof.[3] Case I. Let $f(z, y) = g(z; y(t_1), \cdots, y(t_n))$, where $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ and let $g(z; u_1, \cdots, u_n)$ be the characteristic functional of a half-open interval I in $C_1[a, b] \times R^n$; i.e., $I = \{(z; u_1, \cdots, u_n) \mid -\infty \leqq \gamma_j < z(s_j) \leqq \delta_j \leqq +\infty$ for $j = 1, \cdots, m$; $-\infty \leqq c_k < u_k \leqq d_k \leqq +\infty$ for $k = 1, \cdots, n\}$, $a = s_0 < s_1 < \cdots < s_m = b$.

The right member of (4.1) becomes

$$\begin{aligned}I_2 \equiv \int_{C_1[a,b]} \int_{C_1[\alpha,\beta]} g\Big\{ & \sqrt{\frac{\beta-\alpha}{2}} z; (t_1 - \alpha) \\ & \cdot \left[\sqrt{\frac{b-a}{2}} \int_{t_1}^{\beta} \frac{dy(\tau)}{\tau - \alpha} + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\right], \cdots, \\ & (t_{n-1} - \alpha)\left[\sqrt{\frac{b-a}{2}} \int_{t_{n-1}}^{\beta} \frac{dy(\tau)}{\tau - \alpha} + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\right], \\ & (t_n - \alpha)\frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big\} dy dz \,. \end{aligned} \tag{4.2}$$

We now apply the well-known result: If $\varphi_1, \cdots, \varphi_n$ are orthonormal on $[a, b]$ and of bounded variation on $[a, b]$ and if $h(u_1, \cdots, u_n)$ is measurable, then

$$\begin{aligned}&\int_{C_1[a,b]} h\left\{\int_a^b \varphi_1(t)dx(t), \cdots, \int_a^b \varphi_n(t)dx(t)\right\} dx \\ &= (2\pi)^{-n/2} \int_{-\infty}^{\infty} \overset{(n)}{\cdots} \int_{-\infty}^{\infty} h(u_1, \cdots, u_n) \exp\left\{-\sum_{j=1}^{n} \frac{u_j^2}{2}\right\} du_1 \cdots du_n \,,\end{aligned} \tag{4.3}$$

where the existence of either member implies that of the other and their equality.

To apply this result, we let

$$\theta_j(t) = \begin{cases} 0 & \text{if} \quad \alpha \leqq t \leqq t_j \,, \\ \dfrac{1}{t-\alpha} & \text{if} \quad t_j < t \leqq \beta \end{cases} \qquad \text{for} \quad j = 1, \cdots, n$$

and note that $\theta_n(t) \equiv 0$, so

$$\int_{t_j}^{\beta} \frac{dy(\tau)}{\tau - \alpha} = \int_{\alpha}^{\beta} \theta_j(t) dy(t), \quad \text{for} \quad j = 1, \cdots, n \,. \tag{4.4}$$

We define

[3] This proof is given in the logical order. For motivation read in reverse order, using the inverse of the matrix $(c_{k,j})$.

$$\begin{aligned}\varphi_j(t) &= \frac{-[\theta_{j+1}(t) - \theta_j(t)]}{\sqrt{\int_\alpha^\beta (\theta_{j+1}(\tau) - \theta_j(\tau))^2 d\tau}} \\ &= -\sqrt{\frac{(t_j - \alpha)(t_{j+1} - \alpha)}{t_{j+1} - t_j}}[\theta_{j+1}(t) - \theta_j(t)] \text{ for } j = 1, \cdots, n-1 ,\end{aligned}$$

and observe that $\{\varphi_1, \cdots, \varphi_{n-1}\}$ forms an orthonormal set on $[\alpha, \beta]$. To solve for θ_j we write

$$\theta_{j+1}(t) - \theta_j(t) = -\sqrt{\frac{t_{j+1} - t_j}{(t_j - \alpha)(t_{j+1} - \alpha)}}\varphi_j(t)$$

and sum from $j = k$ to $j = n - 1$ to obtain

$$\theta_k(t) = \sum_{j=k}^{n-1} \sqrt{\frac{t_{j+1} - t_j}{(t_j - \alpha)(t_{j+1} - \alpha)}}\varphi_j(t)$$

and consequently (4.4) becomes

$$\int_{t_k}^\beta \frac{dy(\tau)}{\tau - \alpha} = \sum_{j=k}^{n-1} \sqrt{\frac{t_{j+1} - t_j}{(t_j - \alpha)(t_{j+1} - \alpha)}} \int_\alpha^\beta \varphi_j(t)dy(t) .$$

Substituting the value of $\int_{t_k}^\beta (dy(\tau))/(\tau - \alpha)$ for $k = 1, \cdots, n - 1$ into (4.2) we obtain

$$\begin{aligned} I_2 = \int_{C_1[a,b]} \int_{C_1[\alpha,\beta]} g\Big\{\sqrt{\frac{\beta-\alpha}{2}}z; (t_1 - \alpha)\Big[\sqrt{\frac{b-a}{2}} \\ \cdot \sum_{j=1}^{n-1} \sqrt{\frac{t_{j+1} - t_j}{(t_j - \alpha)(t_{j+1} - \alpha)}} \int_\alpha^\beta \varphi_j(t)dy(t) + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big], \cdots, \\ (t_{n-1} - \alpha)\Big[\sqrt{\frac{b-a}{2}}\sqrt{\frac{t_n - t_{n-1}}{(t_{n-1} - \alpha)(t_n - \alpha)}} \int_\alpha^\beta \varphi_{n-1}(t)dy(t) \\ + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big], (t_n - \alpha)\frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big\}dydz . \end{aligned} \tag{4.4}$$

We now use (4.3) to evaluate the inner Wiener integral above and obtain

$$\begin{aligned} I_2 = \int_{C_1[a,b]} (2\pi)^{-(n-1)/2} \int_{-\infty}^\infty \overset{(n-1)}{\cdots} \int_{-\infty}^\infty g\Big\{\sqrt{\frac{\beta-\alpha}{2}}z; \\ (t_1 - \alpha)\Big[\frac{z(b)}{\sqrt{2(\beta-\alpha)}} + \sqrt{\frac{b-a}{2}} \sum_{j=1}^{n-1} \sqrt{\frac{t_{j+1} - t_j}{(t_j - \alpha)(t_{j+1} - \alpha)}}u_j\Big], \cdots, \\ (t_{n-1} - \alpha)\Big[\frac{z(b)}{\sqrt{2(\beta-\alpha)}} + \sqrt{\frac{b-a}{2}}\sqrt{\frac{t_n - t_{n-1}}{(t_{n-1} - \alpha)(t_n - \alpha)}}u_{n-1}\Big], \\ (t_n - \alpha)\frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big\} \exp\Big\{-\frac{u_1^2}{2} - \cdots - \frac{u_{n-1}^2}{2}\Big\}du_1 \cdots du_{n-1}dz . \end{aligned}$$

If we set $v_j = u_j\sqrt{b-a}$, then

$$I_2 = \int_{C_1[a,b]} (2\pi(b-a))^{-((n-1)/2)} \int_{-\infty}^{\infty} \overset{(n-1)}{\cdots} \int_{-\infty}^{\infty} g\Big\{\sqrt{\frac{\beta-\alpha}{2}}z;$$

$$(t_1-\alpha)\Big[\frac{z(b)}{\sqrt{2(\beta-\alpha)}} + \frac{1}{\sqrt{2}}\sum_{j=1}^{n-1}\sqrt{\frac{t_{j+1}-t_j}{(t_j-\alpha)(t_{j+1}-\alpha)}}v_j\Big], \cdots,$$

$$(t_{n-1}-\alpha)\Big[\frac{z(b)}{\sqrt{2(\beta-\alpha)}} + \frac{1}{\sqrt{2}}\sqrt{\frac{t_n-t_{n-1}}{(t_{n-1}-\alpha)(t_n-\alpha)}}v_{n-1}\Big],$$

$$(t_n-\alpha)\frac{z(b)}{\sqrt{2(\beta-\alpha)}}\Big\} \exp\Big\{-\frac{v_1^2}{2(b-a)} - \cdots - \frac{v_{n-1}^2}{2(b-a)}\Big\}$$

$$\cdot\, dv_1 \cdots dv_{n-1} dz\,.$$

Using the formula

$$\frac{1}{\sqrt{2\pi(b-a)}}\int_{-\infty}^{\infty} F(v)e^{-v^2/(2(b-a))}dv = \int_{C_1[a,b]} F(x(b))dx$$

$(n-1)$ times, we see that (replacing z by z_n)

$$I_2 = \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} g\Big\{\sqrt{\frac{\beta-\alpha}{2}}z_n; \frac{t_1-\alpha}{\sqrt{2}}\Big[\frac{z_n(b)}{\sqrt{\beta-\alpha}}$$

$$+ \sum_{j=1}^{n-1}\sqrt{\frac{t_{j+1}-t_j}{(t_j-\alpha)(t_{j+1}-\alpha)}}z_j(b)\Big], \cdots, \Big(\frac{t_{n-1}-\alpha}{\sqrt{2}}\Big)\Big[\frac{z_n(b)}{\sqrt{\beta-\alpha}}$$

$$+ \sqrt{\frac{t_n-t_{n-1}}{(t_{n-1}-\alpha)(t_n-\alpha)}}z_{n-1}(b)\Big], \frac{(t_n-\alpha)}{\sqrt{2}}\frac{z_n(b)}{\sqrt{\beta-\alpha}}\Big\}dz_1\cdots dz_n\,.$$

We next apply Theorem 3, using the transformation

$$z_k = \sum_{j=1}^{n} c_{k,j}y_j$$

where for $k \leqq n-1$

$$c_{k,j} = \begin{cases} \sqrt{\dfrac{(t_{k+1}-t_k)(t_j-t_{j-1})}{(t_{k+1}-\alpha)(t_k-\alpha)}} & \text{if} \quad j \leqq k \\ -\sqrt{\dfrac{t_k-\alpha}{t_{k+1}-\alpha}} & \text{if} \quad j = k+1 \\ 0 & \text{if} \quad j > k+1 \end{cases}$$

and

$$c_{n,j} = \sqrt{\frac{t_j-t_{j-1}}{t_n-\alpha}} \qquad \text{for} \quad j = 1, \cdots, n\,.$$

We note that $(c_{k,j})$ forms a real orthogonal matrix and that

$$(t_k-\alpha)\Big[\frac{z_n}{\sqrt{\beta-\alpha}} + \sum_{j=k}^{n-1}\sqrt{\frac{t_{j+1}-t_j}{(t_j-\alpha)(t_{j+1}-\alpha)}}z_j\Big]$$

$$= \sum_{1}^{k}\sqrt{t_j-t_{j-1}}\,y_j\,, \quad \text{for} \quad k = 1, \cdots, n\,.$$

Thus by Theorem 3

$$I_2 = \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} g\Big\{\sum_{j=1}^{n} \sqrt{\frac{t_j - t_{j-1}}{2}} y_j;$$
$$\sqrt{\frac{t_1 - \alpha}{2}} y_1(b), \cdots, \sum_{j=1}^{n} \sqrt{\frac{t_j - t_{j-1}}{2}} y_j(b)\Big\} dy_1 \cdots dy_n$$

and by Theorem 2,

$$I_2 = \int_{C_2[R]} g(x(\cdot, t_n); x(b, t_1), \cdots, x(b, t_n))dx = \int_{C_2[R]} f[x(\cdot, \beta), x(b, \cdot)]dx ,$$

and Case I is established.

We then proceed as in Theorem 1 to establish Theorem 4.

5. Applications of Theorem 4.

EXAMPLE 1. Let us apply Theorem 4 to the functional:

$$f(z, y) = \int_a^b p(s)[z(s)]^2 ds \int_\alpha^\beta q(t)[y(t)]^2 dt \tag{5.0}$$

where $p \in L_1[a, b]$ and $q \in L_1[\alpha, \beta]$. Then

$$\begin{aligned} I &\equiv \int_{C_2[R]} \Big\{\int_a^b p(s)[x(s, \beta)]^2 ds \int_\alpha^\beta q(t)[x(b, t)]^2 dt\Big\} dx \\ &= \int_{C_1[a,b]} \int_{C_1[\alpha,\beta]} \int_a^b p(s)\Big(\frac{\beta - \alpha}{2}\Big) z^2(s) ds \int_\alpha^\beta q(t)[(t - \alpha)^2] \\ &\quad \cdot \Big[\sqrt{\frac{b - a}{2}} \int_t^\beta \frac{dy(\tau)}{\tau - \alpha} + \frac{z(b)}{\sqrt{2(\beta - \alpha)}}\Big]^2 dt dy dz \end{aligned} \tag{5.1}$$

and each expression can be evaluated by known techniques to yield

$$\begin{aligned} I = \frac{1}{4} \int_a^b \int_\alpha^\beta & p(s)q(t)(s - a)(t - \alpha)[(b - a)(\beta - \alpha) \\ & + 2(s - a)(t - \alpha)]dxdt . \end{aligned} \tag{5.2}$$

EXAMPLE 2. We next show how to calculate the following integral using Theorem 4: (the authors know of no way of evaluating the integral without applying Theorem 4)

$$I \equiv \int_{C_2[R]} \exp\Big\{A \int_a^b [x(s, \beta)]^2 ds + B \int_\alpha^\beta x(b, t)dt\Big\} dx . \tag{5.3}$$

Let us set

$$f(z, y) = \exp\Big\{A \int_a^b [z(s)]^2 ds + B \int_\alpha^\beta y(t)dt\Big\} . \tag{5.4}$$

By Theorem 4,

$$(5.5)\qquad \begin{aligned} I &= \int_{C_1[a,b]}\int_{C_1[\alpha,\beta]} \exp\left\{\frac{A(\beta-\alpha)}{2}\int_a^b [z(s)]^2 ds \right.\\ &\left.\quad + B\int_\alpha^\beta [t-\alpha]\left[\sqrt{\frac{b-a}{2}}\int_t^\beta \frac{dy(\tau)}{\tau-\alpha} + \frac{z(b)}{\sqrt{2(\beta-\alpha)}}\right]dt\right\} dy dz \\ &= I_1 \cdot I_2 \text{ where} \end{aligned}$$

$$(5.6)\qquad I_1 = \int_{C_1[a,b]} \exp\left\{\frac{A(\beta-\alpha)}{2}\int_a^b [z(s)]^2 ds\right\} \exp\left\{B\int_\alpha^\beta \frac{(t-\alpha)z(b)}{\sqrt{2(\beta-\alpha)}}dt\right\} dz$$

and

$$(5.7)\qquad I_2 = \int_{C_1[\alpha,\beta]} \exp\left\{B\sqrt{\frac{b-a}{2}}\int_\alpha^\beta (t-\alpha)\int_t^\beta \frac{dy(\tau)}{\tau-\alpha}dt\right\} dy\ .$$

To evaluate I_1, we shall use the following theorem of Cameron and Martin [2, 75] where we have changed the scale and the variance:

THEOREM 1a. *Let $q(t)$ be continuous and positive on $[a, b]$ and let μ_0 be the least characteristic value of the differential equation*

$$(5.8)\qquad h''(s) + \mu q(s) h(s) = 0$$

subject to the boundary conditions

$$(5.9)\qquad h(a) = h'(b) = 0\ .$$

Then if $F(x)$ is any Wiener measurable functional, if $\mu < \mu_0$, and if $h_\mu(t)$ is any nontrivial solution of (5.8) *satisfying $h'_\mu(b) = 0$, we have*

$$(5.10)\qquad \begin{aligned} &\int_{C_1[a,b]} F(x) \exp\left\{\frac{\mu}{2}\int_a^b q(s)x^2(s)ds\right\}dx \\ &= \sqrt{\frac{h_\mu(b)}{h_\mu(a)}}\int_{C_1[a,b]} F\left[h_\mu(\cdot)\int_a^{(\cdot)} \frac{dy(\sigma)}{h_\mu(\sigma)}\right]dy \end{aligned}$$

where the existence of either member implies that of the other and their equality.

We now identify in the expression for I_1 in (5.6)

$$(5.11)\qquad F(z) = \exp\left\{\frac{Bz(b)}{\sqrt{2(\beta-\alpha)}}\int_\alpha^\beta (t-\alpha)dt\right\} = \exp\left\{Bz(b)\left(\frac{\beta-\alpha}{2}\right)^{3/2}\right\}\ .$$

Let $q(s) \equiv 1$, $\mu = A(\beta - \alpha)$. An examination of the differential system shows that the least characteristic value is $\mu_0 = \pi^2/(4(b-a)^2)$. We must therefore have $A < \pi^2[4(b-a)^2(\beta-\alpha)]^{-1}$. Now

$$h_\mu(s) = \cos((s-b)\mu^{1/2}) \\ = \cos((s-b)\sqrt{A(\beta-\alpha)}) ,$$

and our integral I_1 may be evaluated:

$$I_1 = \sqrt{\frac{1}{\cos((b-a)\sqrt{A(\beta-\alpha)})}} \int_{C_1[a,b]} \cdot \exp\left\{B\left(\frac{\beta-\alpha}{2}\right)^{3/2}\int_a^b \frac{dy(\sigma)}{\cos((\sigma-b)\sqrt{A(\beta-\alpha)})}\right\} dy .$$

In order to employ (4.3), we normalize the secant function appearing in the Stieltjes integral, i.e., since

$$\int_a^b \sec^2[(\sigma-b)\sqrt{A(\beta-\alpha)}]d\sigma = \frac{\tan[(b-a)\sqrt{A(\beta-\alpha)}]}{\sqrt{A(\beta-\alpha)}}$$

we let $p(\sigma) = \sec[(\sigma-b)\sqrt{A(\beta-\alpha)}](\tan\gamma)^{-1/2}[A(\beta-\alpha)]^{1/4}$, where $\gamma = (b-a)\sqrt{A(\beta-\alpha)}$. Our integral I_1 becomes

$$I_1 = \sqrt{\sec\gamma}\int_{C_1[a,b]} \exp\left\{c\cdot\int_a^b p(\sigma)dy(\sigma)\right\}dy$$

where $c = B((\beta-\alpha)/2)^{3/2}\sqrt{\tan\gamma}[A(\beta-\alpha)]^{-(1/4)}$.

We apply (4.3) to obtain

$$\begin{aligned} I_1 &= \sqrt{\frac{\sec\gamma}{2\pi}}\int_{-\infty}^{\infty} e^{cu}e^{-u^2/2}du \\ &= \sqrt{\sec\gamma}\exp\left\{\frac{B^2\left(\frac{\beta-\alpha}{2}\right)^3\tan\gamma(A(\beta-\alpha))^{-1/2}}{2}\right\} \\ &= [\sec[(b-a)(A(\beta-\alpha))^{1/2}]]^{1/2} \\ &\quad\cdot\exp\left\{\frac{B^2(\beta-\alpha)^{5/2}\tan[(b-a)[A(\beta-\alpha)]^{1/2}]}{A^{1/2}2^4}\right\} . \end{aligned}$$

In I_2, we set

$$\begin{aligned} J &= \int_\alpha^\beta (t-\alpha)\int_t^\beta \frac{dy(\tau)}{\tau-\alpha}dt \\ &= \int_\alpha^\beta\left(\int_\alpha^\tau (t-\alpha)dt\right)\frac{dy(\tau)}{\tau-\alpha} \\ &= \frac{1}{2}\int_\alpha^\beta (\tau-\alpha)dy(\tau) . \end{aligned}$$

We normalize the integrand of this Stieltjes integral and set $p(\tau) = (\tau-\alpha)(\beta-\alpha)^{-3/2}\sqrt{3}$, so the integral becomes

$$J = \frac{(\beta-\alpha)^{3/2}}{2\sqrt{3}} \int_\alpha^\beta p(\tau)dy(\tau) ,$$

and (5.7) becomes

$$\begin{aligned} I_2 &= \int_{C_1[a,b]} \exp\left\{\frac{B\sqrt{b-a}(\beta-\alpha)^{3/2}}{2\sqrt{2}\sqrt{3}} \int_\alpha^\beta p(\tau)dy(\tau)\right\}dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left\{\frac{B\sqrt{b-a}(\beta-\alpha)^{3/2}u}{2\sqrt{6}} - \frac{u^2}{2}\right\}du \\ &= \exp\left\{\frac{B^2(b-a)(\beta-\alpha)^3}{3\cdot 2^4}\right\} . \end{aligned}$$

Thus our original integral has the value $I = I_1 \cdot I_2$, so that

$$\begin{aligned} I = {} & [\sec[(b-a)(A(\beta-\alpha))^{1/2}]]^{1/2} \\ & \cdot \exp\left\{\frac{B^2(\beta-\alpha)^{5/2}\tan[(b-a)(A(\beta-\alpha))^{1/2}]}{A^{1/2}2^4}\right\} \\ & \cdot \exp\left\{\frac{B^2(b-a)(\beta-\alpha)^3}{3\cdot 2^4}\right\} , \end{aligned}$$

where $A < \pi^2/(4(b-a)^2(\beta-\alpha))$ and $A \neq 0$.

6. **General functionals.** Finally we consider a class of functionals which are not required to depend only on the values of x on a restricted set. We do this by approximating $F(x)$ by a sequence of functionals $F(x_n)$ where x_n is determined by the values of x on n horizontal lines and is defined in between the lines by linear interpolation. We then apply Theorem 2 to $F(x_n)$ and take limits.

THEOREM 5. *Let $F(x)$ be a functional which is bounded and continuous in the uniform topology on $C_2[R]$. Let*

$$g_\sigma[y_1, \cdots, y_n; s, t] = \left(\frac{t_k - t}{t_k - t_{k-1}}\right)y_{k-1}(s) + \left(\frac{t - t_{k-1}}{t_k - t_{k-1}}\right)y_k(s) \tag{6.0}$$

for $a \leqq s \leqq b$, $t_{k-1} \leqq t \leqq t_k$, $y_k \in C_1[a, b]$ for $k = 1, \cdots, n$; where σ is a subdivision, $\alpha = t_0 < t_1 < \cdots < t_n = \beta$, and norm $\sigma = \max_{k=1,\cdots,n} |t_k - t_{k-1}|$, *and $y_0 \equiv 0$.*

Then

$$\begin{aligned} \int_{C_2[R]} F(x)dx = \lim_{\text{norm }\sigma\to 0} \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} F\Big\{g_\sigma\Big[\sqrt{\frac{t_1-t_0}{2}}y_1, \cdots, \\ \sqrt{\frac{t_1-t_0}{2}}y_1 + \cdots + \sqrt{\frac{t_n-t_{n-1}}{2}}y_n\Big]\Big\}dy_1 \cdots dy_n . \end{aligned} \tag{6.1}$$

Proof. If we set

$$f(y_1, \cdots, y_n) = F\{g_\sigma[y_1, \cdots, y_n; \cdot, \cdot]\}$$

so that

$$f[x(\cdot, t_1), \cdots, x(\cdot, t_n)] = F\{g_\sigma[x(\cdot, t_1), \cdots, x(\cdot, t_n); \cdot, \cdot]\},$$

our functional f satisfies the hypotheses of Theorem 2 and we have

$$\int_{C_2[R]} f(x(\cdot, t_1), \cdots, x(\cdot, t_n))dx$$
$$= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} f\Big[\sqrt{\frac{t_1 - t_0}{2}}y_1, \sqrt{\frac{t_1 - t_0}{2}}y_1 + \sqrt{\frac{t_2 - t_1}{2}}y_2, \cdots,$$
$$\sqrt{\frac{t_1 - t_0}{2}}y_1 + \cdots + \sqrt{\frac{t_n - t_{n-1}}{2}}y_n\Big]dy_1 \cdots dy_n .$$

If we let

$$F_\sigma(x) \equiv F\{g_\sigma[x(\cdot, t_1), \cdots, x(\cdot, t_n); \cdot, \cdot]\}$$

we obtain

$$\int_{C_2[R]} F_\sigma(x)dx$$
$$= \int_{C_1[a,b]} \overset{(n)}{\cdots} \int_{C_1[a,b]} F\Big\{g_\sigma\Big[\sqrt{\frac{t_1 - t_0}{2}}y_1, \cdots, \sum_{k=1}^{n}\sqrt{\frac{t_k - t_{k-1}}{2}}y_k\Big]\Big\}dy_1 \cdots dy_n .$$

It is clear that $\lim_{\text{norm}\,\sigma\to 0} F_\sigma(x) = F(x)$ for all $x \in C_2[R]$ and since F is bounded we may apply Lebesgue's convergence theorem to obtain (6.1).

REFERENCES

1. J. E. Bearman, *Rotations in the product of two Wiener spaces*, Proc. Amer. Math. Soc., **3** (1952), 129-137.
2. R. H. Cameron and W. T. Martin, *Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations*, Bull. Amer. Math. Soc., **51** (1945), 73-89.
3. J. E. Kuelbs, *Integration on Spaces of Continuous Functions*, Thesis, University of Minnesota, 1965.
4. ———, *Measures on $C(Y)$ when Y is a compact metric space*, Proc. Amer. Math. Soc., **18** (1967), 248-254.
5. J. Yeh, *Wiener measure in a space of functions of two variables*, Trans. Amer. Math. Soc., **95** (1960), 433-450.

Received March 16, 1973. Research sponsored by N.S.F. Grant GP 28732.

UNIVERSITY OF MINNESOTA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

RAMSEY THEORY AND CHROMATIC NUMBERS

GARY CHARTRAND AND ALBERT D. POLIMENI

Let $\chi(G)$ denote the chromatic number of a graph G. For positive integers $n_1, n_2, \cdots, n_k$ $(k \geqq 1)$ the chromatic Ramsey number $\chi(n_1, n_2, \cdots, n_k)$ is defined as the least positive integer p such that for any factorization $K_p = \bigcup_{i=1}^{k} G_i$, $\chi(G_i) \geqq n_i$ for at least one $i, 1 \leqq i \leqq k$. It is shown that $\chi(n_1, n_2, \cdots, n_k) = 1 + \prod_{i=1}^{k} (n_i - 1)$. The vertex-arboricity $a(G)$ of a graph G is the fewest number of subsets into which the vertex set of G can be partitioned so that each subset induces an acyclic graph. For positive integers $n_1, n_2, \cdots, n_k$ $(k \geqq 1)$ the vertex-arboricity Ramsey number $a(n_1, n_2, \cdots, n_k)$ is defined as the least positive integer p such that for any factorization $K_p = \bigcup_{i=1}^{k} G_i$, $a(G_i) \geqq n_i$ for at least one $i, 1 \leqq i \leqq k$. It is shown that $a(n_1, n_2, \cdots, n_k) = 1 + 2k \prod_{i=1}^{k} (n_i - 1)$.

Introduction. The classical *Ramsey number* $r(m, n)$, for positive integers m and n, is the least positive integer p such that for any graph G of order p, either G contains the complete graph K_m of order m as a subgraph or the complement $\bar{G}$ of G contains K_n as a subgraph. More generally, for $k(\geqq 1)$ positive integers $n_1, n_2, \cdots, n_k$, the *Ramsey number* $r(n_1, n_2, \cdots, n_k)$ is defined as the least positive integer p such that for any *factorization* $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$ (i.e., the G_i are spanning, pairwise edge-disjoint, possibly empty subgraphs of K_p such that the union of the edge sets of the G_i equals the edge set of K_p), G_i contains K_{n_i} as a subgraph for at least one $i, 1 \leqq i \leqq k$. It is known (see [5]) that all such Ramsey numbers exist; however, the actual values of $r(n_1, n_2, \cdots, n_k)$, $k \geqq 1$, are known in only seven cases (see [2, 3]) for which $\min \{n_1, n_2, \cdots, n_k\} \geqq 3$.

A *clique* in a graph G is a maximal complete subgraph of G. The *clique number* $\omega(G)$ is the maximum order among the cliques of G. The Ramsey number $r(n_1, n_2, \cdots, n_k)$ may be alternatively defined as the least positive integer p such that for any factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$, $\omega(G_i) \geqq n_i$ for at least one $i, 1 \leqq i \leqq k$.

The foregoing observation suggests the following definition. Let f be a graphical parameter, and let $n_1, n_2, \cdots, n_k$, $k \geqq 1$ be positive integers. The *f-Ramsey number* $f(n_1, n_2, \cdots, n_k)$ is the least positive integer p such that for any factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$, $f(G_i) \geqq n_i$ for at least one $i, 1 \leqq i \leqq k$. Hence, $\omega(n_1, n_2, \cdots, n_k) = r(n_1, n_2, \cdots, n_k)$, i.e., the ω-Ramsey number is the Ramsey number.

The object of this paper is to investigate f-Ramsey numbers for two graphical parameters f, namely chromatic number and vertex-arboricity.

Chromatic Ramsey numbers. The *chromatic number* $\chi(G)$ of a graph G is the fewest number of colors which may be assigned to the vertices of G so that adjacent vertices are assigned different colors. For positive integers $n_1, n_2, \cdots, n_k$, the *chromatic Ramsey number* $\chi(n_1, n_2, \cdots, n_k)$ is the least positive integer p such that for any factorization $K_p = G_1 \cup G_2 \cup \cdots G_k$, $\chi(G_i) \geqq n_i$ for some i, $1 \leqq i \leqq k$. The existence of the numbers $\chi(n_1, n_2, \cdots, n_k)$ is guaranteed by the fact that $\chi(n_1, n_2, \cdots, n_k) \leqq r(n_1, n_2, \cdots, n_k)$. We are now prepared to present a formula for $\chi(n_1, n_2, \cdots, n_k)$. We begin with a lemma.

LEMMA. *If $G = G_1 \cup G_2 \cup \cdots \cup G_k$, then*

$$\chi(G) \leqq \sum_{i=1}^{k} \chi(G_i) .$$

Proof. For $i = 1, 2, \cdots, k$, let a $\chi(G_i)$-coloring be given for G_i. We assign to a vertex v of G the color $(c_1, c_2, \cdots, c_k)$, where c_i is the color assigned to v in G_i. This produces a coloring of G using at most $\prod_{i=1}^{k} \chi(G_i)$ colors; hence, $\chi(G) \leqq \prod_{i=1}^{k} \chi(G_i)$.

THEOREM 1. *For positive integers $n_1, n_2, \cdots, n_k$,*

$$\chi(n_1, n_2, \cdots, n_k) = 1 + \prod_{i=1}^{k} (n_i - 1) .$$

Proof. The result is immediate if $n_i = 1$ for some i; hence, we assume that $n_i \geqq 2$ for all i, $1 \leqq i \leqq k$. First, we verify that

$$\chi(n_1, n_2, \cdots, n_k) \leqq 1 + \prod_{i=1}^{k} (n_i - 1) .$$

Let $p = 1 + \prod_{i=1}^{k} (n_i - 1)$, and assume there exists a factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$ such that $\chi(G_i) \leqq n_i - 1$ for each $i = 1, 2, \cdots, k$. Then by the Lemma, it follows that

$$1 + \prod_{i=1}^{k} (n_i - 1) = \chi(K_p) \leqq \prod_{i=1}^{k} \chi(G_i) \leqq \prod_{i=1}^{k} (n_i - 1) ,$$

which produces a contradiction. Thus, in any factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$ for $p = 1 + \prod_{i=1}^{k} (n_i - 1)$, we have $\chi(G_i) \geqq n_i$ for at least one i, $1 \leqq i \leqq k$.

In order to show that

$$\chi(n_1, n_2, \cdots, n_k) \geqq 1 + \prod_{i=1}^{k} (n_i - 1) ,$$

we exhibit a factorization $K_{N_k} = G_1 \cup G_2 \cup \cdots \cup G_k$, where $N_k =$

$\prod_{i=1}^{k}(n_i - 1)$ and $\chi(G_i) \leqq n_i - 1$ for $i = 1, 2, \cdots, k$. The factorization is accomplished by employing induction on k. For $k = 1$, we simply observe that $\chi(K_{N_1}) = \chi(K_{n_1-1}) = n_1 - 1$. Assume there exists a factorization $K_{N_{k-1}} = H_1 \cup H_2 \cup \cdots \cup H_{k-1}$ such that $\chi(H_i) \leqq n_i - 1$ for $i = 1, 2, \cdots, k - 1$. Let F denote $n_k - 1$ (pairwise disjoint) copies of $K_{N_{k-1}}$ and define G_k by $G_k = \bar{F}$. Thus, $\bar{G}_k$ contains $n_k - 1$ pairwise disjoint copies of H_i for $i = 1, 2, \cdots, k - 1$, which we denote by G_i. Hence, $K_{N_k} = G_1 \cup G_2 \cup \cdots \cup G_k$, where $\chi(G_i) \leqq n_i - 1$ for each i, $1 \leqq i \leqq k$, which produces the desired result.

Vertex-arboricity Ramsey numbers. The *vertex-arboricity* $a(G)$ of a graph G is the minimum number of subsets into which the vertex set of G may be partitioned so that each subset induces an acyclic subgraph. As with the chromatic number, the vertex-arboricity may be considered a coloring number since $a(G)$ is the least number of colors which may be assigned to the vertices of G so that no cycle of G has all of its vertices assigned the same color.

Our next result will establish a formula for the *vertex-arboricity Ramsey number* $a(n_1, n_2, \cdots, n_k)$, defined as the least positive integer p such that for every factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$, $a(G_i) \geqq n_i$ for some i, $1 \leqq i \leqq k$. Since $a(K_n) = \{n/2\}$, it follows that $a(n_1, n_2, \cdots, n_k) \leqq r(2n_1 - 1, 2n_2 - 1, \cdots, 2n_k - 1)$. In the proof of the following result, we shall make use of the *(edge) arboricity* $a_1(G)$ of a graph, which is the minimum number of subsets into which the edge set of G may be partitioned so that the subgraph induced by each subset is acyclic. It is known (see [1, 4]) that $a_1(K_n) = \{n/2\}$.

THEOREM 2. *For positive integers* $n_1, n_2, \cdots, n_k$,

$$a(n_1, n_2, \cdots, n_k) = 1 + 2k \prod_{i=1}^{k} (n_i - 1) .$$

Proof. In order to show that

$$a(n_1, n_2, \cdots, n_k) \leqq 1 + 2k \prod_{i=1}^{k} (n_i - 1) ,$$

we let $p = 1 + 2k \prod_{i=1}^{k}(n_i - 1)$ and assume there exists a factorization $K_p = G_1 \cup G_2 \cup \cdots \cup G_k$ such that $a(G_i) \leqq n_i - 1$ for each $i = 1, 2, \cdots, k$. For each $i = 1, 2, \cdots, k$, there is a partition $\{U_{i,1}, U_{i,2}, \cdots, U_{i,n_i-1}\}$ of the vertex set $V(G_i)$ of G_i such that the subgraph $\langle U_{i,j} \rangle$ of G_i induced by $U_{i,j}$ is acyclic, $j = 1, 2, \cdots, n_i - 1$. At least one of the sets $U_{1,1}, U_{1,2}, \cdots, U_{1,n_1-1}$, say U_{1,m_1}, contains at least $1 + 2k \prod_{i=2}^{k}(n_i - 1)$ vertices. Thus, at least one of the sets $U_{2,1}, U_{2,2}, \cdots,$

U_{2,n_2-1}, say U_{2,m_2}, contains at least $1 + 2k \prod_{i=3}^{k} (n_i - 1)$ vertices of U_{1,m_1}. Proceeding inductively, we arrive at subsets $U_{1,m_1}, U_{2,m_2}, \cdots, U_{k,m_k}$ such that $\bigcap_{i=1}^{t} U_{i,m_i}$ contains at least $1 + 2k \prod_{i=t+1}^{k} (n_i - 1)$ vertices, $1 \leqq t \leqq k - 1$. In particular, $\bigcap_{i=1}^{k} U_{i,m_i}$, contains a set U having $1 + 2k$ vertices. For each $i = 1, 2, \cdots, k$, $\langle U \rangle$ is an acyclic subgraph of the graph $\langle U_{i,m_i} \rangle$. This implies that $a_1(K_{1+2k}) \leqq k$, which is contradictory. Therefore, $a(G_i) \geqq n_i$ for at least one i, $1 \leqq i \leqq k$.

The proof will be complete once we have verified that

$$a(n_1, n_2, \cdots, n_k) \geqq 1 + 2k \prod_{i=1}^{k} (n_i - 1) .$$

Let $r = \prod_{i=1}^{k} (n_i - 1)$. We shall exhibit a factorization $K_{2kr} = G_1 \cup G_2 \cup \cdots \cup G_k$ such that $a(G_i) \leqq n_i - 1$ for $i = 1, 2, \cdots, k$. We begin with r pairwise disjoint copies of K_{2k}, labeled $K_{2k}^1, K_{2k}^2, \cdots, K_{2k}^r$. Since $a_1(K_{2k}) = k$, it follows that $K_{2k} = \bigcup_{i=1}^{k} F_i$, where each F_i is an acyclic graph. We introduce the notation F_{il} to denote the F_i contained in K_{2k}^l, $l = 1, 2, \cdots, r$ and $i = 1, 2, \cdots, k$. With each of the r k-tuples $(c_1, c_2, \cdots, c_k)$, $c_j = 1, 2, \cdots, n_j - 1$ and $j = 1, 2, \cdots, k$, we identify a complete graph K_{2k}^l, $l = 1, 2, \cdots, r$, in such a way that the identification is one-to-one. Then, for each $i = 1, 2, \cdots, k$ and $l = 1, 2, \cdots, r$, we associate with F_{il} the k-tuple identified with K_{2k}^l. Define the graph G_i, $i = 1, 2, \cdots, k$, to consist of the graphs $F_{i1}, F_{i2}, \cdots, F_{ir}$; in addition, each vertex of F_{is} is adjacent to each vertex of F_{it}, $s, t = 1, 2, \cdots, r$, provided the ith coordinate is the first coordinate in which their associated k-tuples differ (otherwise, there are no edges between F_{is} and F_{it}). It is then seen that $K_{2kr} = \bigcup_{i=1}^{k} G_i$. For each $i = 1, 2, \cdots, k$, define $V_{i,j}$ to be the set of all vertices v such that v is a vertex of an F_{il} whose associated k-tuple $(c_1, c_2, \cdots, c_k)$ has $c_i = j$; $j = 1, 2, \cdots, n_i - 1$. Then $\{V_{i,1}, V_{i,2}, \cdots, V_{i,n_i-1}\}$ is a partition of $V(G_i)$ for which the subgraph $\langle V_{i,j} \rangle$ consists of $r/(n_i - 1)$ pairwise disjoint copies of F_i, $j = 1, 2, \cdots, n_i - 1$. Thus, $\langle V_{i,j} \rangle$ is an acyclic graph for each such j. Hence, $a(G_i) \leqq n_i - 1$, $i = 1, 2, \cdots, k$.

References

1. L. W. Beineke, *Decompositions of complete graphs into forests*, Magyar Tud. Akad. Mat. Kutato Int. Kozl., **9** (1964), 589-594.
2. J. E. Graver and J. Yackel, *Some graph theoretic results associated with Ramsey's theorem*, J. Combinatorial Theory, **4** (1968), 125-175.
3. R. E. Greenwood and A. M. Gleason, *Combinatorial relations and chromatic graphs*, Canad. J. Math., **7** (1955), 1-7.
4. C. St. J. A. Nash-Williams, *Edge-disjoint spanning trees of finite graphs*, J. London Math. Soc., **36** (1961), 445-450.
5. F. P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc., **30** (1930), 264-286.

Received October 4, 1974. The second author's research was supported by the Research Foundation of the State of New York.

WESTERN MICHIGAN UNIVERSITY
AND
SUNY, COLLEGE AT FREDONIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

CHARACTERIZATION OF COLLECTIVELY COMPACT SETS OF LINEAR OPERATORS

J. D. DEPREE AND H. S. KLEIN

The basic results in this paper show that each collectively compact set of linear operators can be viewed as an equicontinuous collection followed by a single compact operator. This observation not only gives insight into the character of collectively compact sets of linear operators, but also yields easier proofs of many of the results obtained by earlier workers in the field.

1. Factorizations of collectively compact operators. A fairly complete treatment, with applications, of collectively compact sets of linear operators is given in the recent book [1] by Anselone. Collectively compact sets of linear operators on normed linear spaces were originally studied by Anselone and Moore [2] in connection with approximate solutions of integral and operator equations.

The general properties of such sets of operators, again in normed linear spaces, were studied by Anselone and Palmer in [3] and [4]. Collectively compact sets of linear operators were studied in the more general setting of linear topological spaces by DePree and Higgins [5]. In the current work new characterizations are given for collectively compact sets of operators on a linear topological space.

We assume that X and Y are separated topological vector spaces and that $[X, Y]$ is the space of all continuous linear operators from X to Y. For a collection $\mathscr{T} \subseteqq [X, Y]$ and U a subset of X, let $\mathscr{T}(U) = \{T(x): x \in U, T \in \mathscr{T}\}$. For a set Ω with topology τ, we adopt the notation $\langle \Omega, \tau \rangle$. For a set $\mathscr{N}$ of operators, we will be making statements of the following nature: Viewed as mappings between the unit ball of Y^* endowed with its relative weak-star topology and X^* equipped with norm topology, $\mathscr{N}$ is equicontinuous; we shall simply say that $\mathscr{N}: \langle Y^*, \text{weak}^* \rangle \to \langle X^*, \text{norm} \rangle$ is equicontinuous.

Following the work of DePree and Higgins [5], we make the following definition.

DEFINITION 1.1. Let X and Y be separated topological vector spaces. Then $\mathscr{T} \subseteqq [X, Y]$ is *collectively compact* if there exists a neighborhood U of the origin in X such that $\mathscr{T}(U)$ has compact closure in Y.

The easy proof of the following lemma belies its importance for the situation described in it is typical of collectively compact sets of operators; i.e., they can always be factored as in Theorem 1.3.

LEMMA 1.2. *Let X, Y, and Z be separated topological vector spaces, $\mathscr{N} \subseteqq [X, Z]$ an equicontinuous collection, and $K \in [Z, Y]$ a compact operator. Then the collection $K\mathscr{N} = \{KN: N \in \mathscr{N}\}$ is collectively compact.*

Proof. Let V be a 0-neighborhood in Z such that $\overline{K(V)}$ is compact. Since the family $\mathscr{N}$ is equicontinuous, there is a 0-neighborhood U in X such that $\mathscr{N}(U) \subseteqq V$. Thus $\overline{K\mathscr{N}(U)} \subseteqq \overline{K(V)}$. It follows that $K\mathscr{N}$ is collectively compact.

THEOREM 1.3. *Let $\mathscr{F} \subseteqq [X, Y]$ be such that there exists a 0-neighborhood U in X with the closure of the balanced convex hull of $\mathscr{F}(U)$ compact in Y. Then there exist*

(a) *a Banach space Z,*

(b) *an equicontinuous collection $\mathscr{N} \subseteqq [X, Z]$, and*

(c) *a compact operator $K \in [Z, Y]$*

such that $\mathscr{F} = KN$.

Proof. The following proof is based upon the construction of an auxiliary normed space.

Let the set C be the closure of the balanced convex hull of $\mathscr{F}(U)$ and Z the span of C in Y. Since C is balanced and convex, $Z = \bigcup_{n=1}^{\infty} nC$ and C is absorbing in Z. Hence p, the Minkowski functional of C, is defined on Z.

If $\langle Z, p\rangle$ denotes the set Z endowed with the topology generated by p, then let $K: \langle Z, p\rangle \rightarrow Y$ be the natural injection which maps a point $z \in Z$ to the same point z considered as an element of Y. K is a compact operator since C, the unit ball of Z, is compact in Y. In particular, K is continuous and the p-topology on Z is stronger than the Hausdorff relative topology on Z. So $\langle Z, p\rangle$ is Hausdorff and Z is a normed linear space.

Let $\{z_n\}$ be a Cauchy sequence in $\langle Z, p\rangle$. Since $\{z_n\}$ is a bounded subset of $\langle Z, p\rangle$ and K is a compact operator, $\{K(z_n)\}$ is a Cauchy sequence with $\overline{\{K(z_n)\}}$ compact in Y. So there exists a $y \in Y$ such that $\lim_n K(z_n) = y$. For $\alpha > 0$, choose l such that $n, m \geqq l$ implies that $p(z_n - z_m) \leqq \alpha$. For these n and m, $z_n - z_m$ is an element of αC. In Y, C is a closed set and $y = \lim_n z_n$. So $y - z_m$ is an element of αC for $m \geqq l$. We see that $y \in Z$ and that $p(y - z_m) \leqq \alpha$ for $m \geqq l$. It follows that $\langle Z, p\rangle$ is a Banach space.

Let $x \in X$ and $T \in \mathscr{F}$. Since 0-neighborhoods are absorbing and

U is such a neighborhood in X, there exists an $\alpha > 0$ such that $x \in \alpha U$. Hence, $T(x) \in \alpha C$ and $T(x) \in Z$. So let $\mathscr{N} \subseteqq [X, Z]$ be defined as the collection $\mathscr{F}$ mapping X to $\langle Z, p\rangle$. The collection $\mathscr{N}$ is equicontinuous since $\mathscr{N}(U)$ is a subset of the unit ball of Z. Obviously, $\mathscr{F} = K\mathscr{N}$.

Suppose $\mathscr{F} \subseteqq [X, Y]$ satisfies the hypothesis of Theorem 2.3 and can be factored, $\mathscr{F} = K\mathscr{N}$, as above. The single compact operator K has been the object of study for years. The next section shows that our knowledge about K gives insight into the collection $\mathscr{F}$.

2. Characterizations of collectively compact operators defined on Banach spaces. Throughout this section, X and Y will be Banach spaces with closed unit balls X_1 and Y_1, respectively, X^* and X^{**} will denote the first and second duals of X with their usual norm topologies, and $[X, Y]$ will be the space of continuous linear operators from X to Y endowed with the uniform operator topology.

Note that $\mathscr{F} \subseteqq [X, Y]$ is collectively compact if and only if $\mathscr{F}(F_1)$ has compact closure in Y.

LEMMA 2.1. *Let $\mathscr{F} \subseteqq [X, Y]$. $\mathscr{F}$ is collectively compact if and only if there exist*

(a) *a Banach space Z,*

(b) *an equicontinuous collection $\mathscr{N} \subseteqq [X, Z]$, and*

(c) *a compact operator $K \in [Z, Y]$*

such that $\mathscr{F} = K\mathscr{N}$.

Proof. Mazur's theorem [6, p. 416] states that if $\mathscr{F}(X_1)$ is relatively compact, then so is the balanced convex hull of $\mathscr{F}(X_1)$. Apply Lemma 1.2 and Theorem 1.3.

For $T \in [X, Y]$, let $T^* \in [Y^*, X^*]$ denote the adjoint of T. While Schauder's theorem implies that the adjoint of a compact operator is compact, the following example shows that this phenomenon has no generalization for collectively compact sets of operators. This example will also serve as an illustration for the results of the remainder of this paper.

EXAMPLE 2.2. Let $X = Y = l_2$ with the usual orthonormal basis $\{e_k: k = 1, 2, \cdots\}$. For each positive integer n, define T_n by letting $T_n(x) = (x, e_n)e_1$. The set $\{T_n: n \geqq 1\}$ is collectively compact since $\bigcup_n T_n(X_1)$ is a bounded subset of the finite dimensional subspace generated by e_1.

However, the collection of adjoints $\{T_n^*\}$ is not collectively compact since $T_n^*(x) = (x, e_1)e_n$ and $\bigcup_n T_n^*(X_1^*)$ contains the orthonormal basis $\{e_k\}$.

LEMMA 2.3. *Let* $\mathscr{F} \subseteqq [X, Y]$. *Then* $\mathscr{F}^* = \{T^*: T \in \mathscr{F}\} \subseteqq [Y^*, X^*]$ *is collectively compact if and only if there exists a Banach space* Z, *an equicontinuous collection* $\mathscr{S} \subseteqq [Z, Y]$, *and a compact operator* $K \in [X, Z]$ *such that* $\mathscr{F} = \mathscr{S}K = \{SK: S \in \mathscr{S}\}$.

Proof. Assume that $\mathscr{F} = \mathscr{S}K$, with K a compact operator and $\mathscr{S}$ equicontinuous. The process of taking adjoints is an anti-homomorphism which preserves operator norms. So if $\mathscr{F} = \mathscr{S}K$, then $\mathscr{S}^* = K^*\mathscr{S}^*$. If $\mathscr{S}$ is an equicontinuous (i.e., bounded) subset of $[Z, Y]$, then $\mathscr{S}^*$ is an equicontinuous subset of $[Y^*, Z^*]$. Lemma 2.1 implies that $\mathscr{F}^*$ is collectively compact.

Conversely, if $\mathscr{F}^*$ is collectively compact, there exists a Banach space W, an equicontinuous collection $\mathscr{P} \subseteqq [Y^*, W]$, and a compact operator $L \in [W, X^*]$ such that $\mathscr{F}^* = L\mathscr{P}$.

Let J_x and J_y denote the natural injections of X into X^{**} and Y into Y^{**}, respectively. Note that $\mathscr{F} = J_y^{-1}\mathscr{F}^{**}J_x = (J_y^{-1}\mathscr{P}^*)(L^*J_x)$. Let $K = L^*J_x$. Then K is a compact operator mapping X into W^*.

Since J_y is an isometry and $\mathscr{P}^*$ is equicontinuous, it follows that $\mathscr{S} = J_y^{-1}\mathscr{P}^*$ is an equicontinuous subset of $[W^*, Y]$ such that $\mathscr{S}K = (J_y^{-1}\mathscr{P}^*)K = \mathscr{F}$.

A description of the bounded weak-star topology of a Banach space Y is given in [6, pages 427–430]. The feature of the bounded weak-star topology that will be of interest to us is the equivalence of parts (i) and (ii) of the following theorem.

THEOREM 2.4. *Let* Y *be a Banach space. If* Y_1^* *denotes the closed unit ball of* Y^*, *then a set* $U \subseteqq Y^*$ *is a bounded weak-star neighborhood of* 0 *if and only if either one of the following are satisfied*:

(i) *For each* $\alpha > 0$, $U \cap \alpha Y_1^*$ *is a relative weak-star neighborhood of* 0 *in* αY_1^*.

(ii) *There exists a sequence* $\{y_n\} \subseteqq Y$ *such that* $\lim_n \|y_n\| = 0$ *and* $\{y^* \in Y^*: |\langle y^*, y_n\rangle| \leqq 1$ *for each* $n\}$ *is a subset of* U.

Of course, statement (ii) *may be rephrased in the form*: *There is a sequence* $\{y_n\} \subseteqq Y$ *converging to* 0 *in norm such that the polar of* $\{y_n\}$ *is a subset of* U.

THEOREM 2.5. *For* $\mathscr{F} \subseteqq [X, Y]$, *the following are equivalent*:

(a) $\mathscr{F}$ *is collectively compact.*

(b) $\mathscr{F}^*: \langle Y_1^*, \text{weak}^*\rangle \to \langle X^*, \text{norm}\rangle$ *is equicontinuous.*

(c) $\mathscr{F}^*: \langle Y_1^*, \text{weak}^*\rangle \to \langle X^*, \text{norm}\rangle$ *is equicontinuous at the origin.*

(d) $\{y^*: \|T^*(y^*)\| \leqq 1$ *for each* $T^* \in \mathscr{F}^*\}$ *is a bounded weak-star neighborhood of* 0.

(e) *There exists a sequence* $\{y_n\} \subseteqq Y$ *such that* $\|y_n\| \to 0$ *and* $\mathscr{F}(X_1)$ *is a subset of the closure of the balanced convex hull of* $\{y_n\}$.

Proof. (a) implies (b). If $\mathscr{F}$ is collectively compact, Lemma 2.1 implies that there exists a Banach space Z and a factorization of $\mathscr{F}$, which, after taking adjoints, is of the form:

(1) $\mathscr{F}^* = \mathscr{N}^* K^*$.

(2) $K^*: \langle Y^*, \text{norm}\rangle \to \langle Z^*, \text{norm}\rangle$ is a compact operator.

(3) $\mathscr{N}^*: \langle Z^*, \text{norm}\rangle \to \langle X^*, \text{norm}\rangle$ is equicontinuous.

Now since K^* is a compact operator, $K^*: \langle Y_1^*, \text{weak}^*\rangle \to \langle Z^*, \text{norm}\rangle$ is continuous: It maps bounded nets which converge in the weak-star topology to weak-star convergent nets which are also totally bounded in the norm topology of Z^*. By (3), $\mathscr{N}^* K^*$ is an equicontinuous collection of mappings of $\langle Y_1^*, \text{weak}^*\rangle$ into $\langle X^*, \text{norm}\rangle$. We see that (b) follows immediately from (1).

(b) implies (c). This implication is obvious.

(c) implies (d). If the situation $\|T^*(y^*)\| \leqq r$ for each $T^* \in \mathscr{F}^*$ is abbreviated $\|\mathscr{F}^*(y^*)\| \leqq r$, then for any $\alpha > 0$, (c) implies that $\{y^*: \|\mathscr{F}^*(y^*)\| \leqq 1/\alpha\} \cap Y_1^*$ is a relative weak-star neighborhood of 0 in Y_1^*. Multiplication by α yields that $\{y^*: \|\mathscr{F}^*(y^*)\| \leqq 1\} \cap \alpha Y_1^*$ is a relative weak-star neighborhood of 0 in αY_1^*. Theorem 2.4, part (i), yields (d).

(d) implies (e). Statement (d) together with Theorem 2.4, part (ii), guarantee the existence of a sequence $\{y_n\} \subseteqq Y$ such that $\|y_n\| \to 0$ and

(4) $\{y^*: |\langle y^*, y_n\rangle| \leqq 1 \text{ for each } n\} \subseteqq \{y^*: \|\mathscr{F}^*(y^*)\| \leqq 1\}$. Now take polars in Y of both of the above sets. By the Bipolar Theorem [7, p. 141], the polar of the left-hand side of (4) is the closure of the balanced convex hull of $\{y_n\}$. Since $\|\mathscr{F}^*(y^*)\| \leqq 1$, implies that $\|y^*(\mathscr{F}(X_1))\| \leqq 1$, the polar of the right-hand of (4) contains $\mathscr{F}(X_1)$.

(e) implies (a). The set $\{y_n\}$ is compact. Therefore, the closure of the balanced convex hull of $\{y_n\}$ is also compact.

The following corollary was first proved by Palmer [8]. A new and simpler proof is given below.

COROLLARY 2.6. *Let* $\mathscr{F} \subseteqq [X, Y]$. *If*

(a) *F is collectively compact, and if*

(b) *for each* $y^* \in Y^*$, $\mathscr{F}^*(y^*)$ *is totally bounded in the norm topology of* X^*,
then $\mathscr{F}$ *is totally bounded in* $[X, Y]$.

Proof. By Theorem 2.5, (a) implies that if we consider $\mathscr{F}^*$ as a set of mappings between $\langle Y_1^*, \text{weak}^*\rangle$ and $\langle X^*, \text{norm}\rangle$, then $\mathscr{F}^*$ is equicontinuous with respect to these topologies. Since $\langle Y_1^*, \text{weak}^*\rangle$ is a compact topological space, (b) together with the Ascoli theorem [7, p. 81] imply that the collection $\mathscr{F}^*$ is totally bounded in the topology of uniform convergence on Y_1^*, that is, in the uniform operator topology. Since the adjoint is an isometry between $[X, Y]$ and $[Y^*, X^*]$, $\mathscr{F}$ is totally bounded in $[X, Y]$.

COROLLARY 2.7. *If* $\mathscr{N} \subseteqq [X, Y]$ *is totally bounded and each member of* $\mathscr{N}$ *is a compact operator, then* $\mathscr{N}$ *is collectively compact.*

Proof. If $\mathscr{N} \subseteqq [X, Y]$ is totally bounded, so is $\mathscr{N}^* \subseteqq [Y^*, X^*]$, i.e., $\mathscr{N}^*$ is totally bounded in the topology of uniform convergence on Y_1^*. Since each $T^* \in \mathscr{N}^*$ is a compact operator, each

$$T^*: \langle Y_1^*, \text{weak}^*\rangle \longrightarrow \langle X^*, \text{norm}\rangle \tag{5}$$

is continuous. Considered as a collection of mappings between the topological spaces of (5), $\mathscr{N}^*$ must be equicontinuous. By Theorem 2.5, $\mathscr{N}$ is collectively compact.

In order to extend the range of application of Corollary 2.6, the following theorem is stated.

THEOREM 2.8. *Let* $\mathscr{S} = \{S_n: n \geqq 1\}$ *be a sequence of bounded linear maps from* X *to* Y. *Suppose there exists a collectively compact set* $\{V_n: n \geqq 1\} \subseteqq [X, Y]$ *such that* $\lim_{n\to\infty} \| S_n - V_n \| = 0$. *If* $\mathscr{S}^*(y^*)$ *is a totally bounded subset of* X^* *for each* $y^* \in Y^*$, *then* $\mathscr{S}$ *is a totally bounded subset of* $[X, Y]$.

Proof. Since $\lim_n \| S_n - V_n \| = 0$, $\lim_n \| S_n^* - V_n^* \| = 0$. Let $y^* \in Y^*$ *and* $\varepsilon > 0$ be given. Choose an integer N such that

$$\| S_n^*(y^*) - V_n^*(y^*) \| \leqq \varepsilon/3 \quad \text{for} \quad n \geqq N\,. \tag{6}$$

$\mathscr{S}^*(y^*)$ is totally bounded and consequently has a finite $\varepsilon/3$-net. The inequality (6) then implies that $\{V_n^*(y^*): n \geqq N\}$ has a finite ε-net. Since the excluded points are finite in number and $\varepsilon > 0$ is arbitrary, the set $\{V_n^*(y^*): n \geqq 1\}$ is totally bounded. By Corollary

2.6, $\{V_n: n \geqq 1\}$ is a totally bounded subset of $[X, Y]$. However, $\lim_n \| S_n - V_n \| = 0$. By using an argument similar to the one above, it follows that $\mathscr{S}$ is a totally bounded subset of $[X, Y]$.

LEMMA 2.9. *Let $\mathscr{S} \subseteqq [X, Y]$. If for each $x \in X$, $\mathscr{S}(x)$ is totally bounded in Y, then $\mathscr{S}^*: \langle Y_1^*, \text{weak}^* \rangle \to \langle X^*, \text{weak}^* \rangle$ is equicontinuous at the origin.*

Proof. Let x be any fixed element of X. Then

$$W = \{x^*: | \langle x^*, x \rangle | \leqq 1\}$$

is a neighborhood of 0 in the weak-star topology of X^*. In fact, the family of all such W form a sub-basis of the neighborhood system of 0 for the weak-star topology. Therefore, it suffices to show that $Y_1^* \cap \{y^*: \mathscr{S}^*(y^*) \subseteqq W\} = Y_1^* \cap \{y^*: | \langle \mathscr{S}^*(y^*), x \rangle | \leqq 1\} = Y_1^* \cap \{y^*: | < y^*, \mathscr{S}(x) \rangle | \leqq 1\}$ is a neighborhood of 0 in the relative weak-star topology on Y_1^*.

Let $\{y_1, y_2, \cdots, y_n\}$ be a 1/2-net for $\mathscr{S}(x)$. Consider $V = \{y^*: | \langle y^*, y_i \rangle | \leqq 1/2, \ 1 \leqq i \leqq n\}$. If $y^* \in V \cap Y_1^*$, and $y \in \mathscr{S}(x)$, choose j such that $\| y - y_j \| \leqq 1/2$. Then $| \langle y^*, y \rangle \leqq | \langle y^*, y_j \rangle | + | \langle y^*, y - y_j \rangle | \leqq 1/2 + \| y^* \| \, \| y - y_j \| \leqq 1$ since $y^* \in Y_1^*$. So $V \cap Y_1^* \subseteqq Y_1^* \cap \{y^*: \mathscr{S}^*(y^*) \subseteqq W\}$. It follows that

$$\mathscr{S}^*: \langle Y_1^*, \text{weak}^* \rangle \longrightarrow \langle X^*, \text{weak}^* \rangle$$

is equicontinuous at the origin.

THEOREM 2.10. *Let X, Y, and Z be Banach spaces and let $\mathscr{T} \subseteqq [X, Z]$ be collectively compact. For $\mathscr{S} \subseteqq [Z, Y]$, suppose $\mathscr{S}(z)$ is totally bounded in Y for each $Z \in Z$. Then $\mathscr{S}\mathscr{T} = \{ST: S \in \mathscr{S}, T \in \mathscr{T}\}$ is collectively compact.*

Proof. Since $\mathscr{S}(z)$ is bounded for each $z \in Z$ and Z is complete, there exists a constant m such that $\| S \| \leqq m$ for each $S \in \mathscr{S}$. If U is any 0-neighborhood in the norm topology of X^*, choose, by Theorem 2.5, a weak-star neighborhood W of 0 in Z^* such that $\mathscr{T}^*(W \cap Z_1^*) \subseteqq (1/m)U$. Lemma 2.9 guarantees that there exists a weak-star neighborhood V of 0 in Y^* such that $\mathscr{S}^*(V \cap Y_1^*) \subseteqq mW$. So $(1/m)\mathscr{S}^*(V \cap Y_1^*) \subseteqq W \cap (1/m)\mathscr{S}^*(Y_1^*) \subseteqq W \cap Z_1^*$. It follows that $\mathscr{T}^*((1/m)\mathscr{S}^*(V \cap Y_1^*)) \subseteqq (1/m)U$ and that $(\mathscr{S}\mathscr{T})^*: \langle Y_1^*, \text{weak}^* \rangle \to \langle X^*, \text{norm} \rangle$ is equicontinuous at the origin. Theorem 2.5 implies that the set $\mathscr{S}\mathscr{T}$ is collectively compact.

EXAMPLE 2.2 *continued.* Let $C = \{x: \| T_n(x) \| \leqq 1 \text{ for each } n\}$.

Then $C = \{x: |(x, e_n)| \leqq 1$ for each $n\}$ and C is the polar of $\{e_n\}$. C is not a bounded weak* neighborhood of 0, since if it were, the Bipolar Theorem would then imply that the orthonormal basis $\{e_n\}$ is a compact subset of l_2. Since l_2 is a Hilbert space, one can view $\{T_n\}$ as the adjoint of the collection $\{T_n^*\}$. Theorem 2.5, part (d), implies that $\{T_n^*\} \subseteqq [l_2, l_2]$ is not collectively compact. In particular, an explicit calculation of the adjoints is unnecessary in determining whether or not $\{T_n^*\}$ is collectively compact. The next theorem shows that it is unnecessary to calculate the adjoints even when the operators involved are acting on arbitrary Banach spaces.

THEOREM 2.11. *For* $\mathscr{F} \subseteqq [X, Y]$, *the following are equivalent:*

(a) $\mathscr{F}^*[Y^*, X^*]$ *is collectively compact.*

(b) $\mathscr{F}: \langle X_1, \text{weak topology}\rangle \rightarrow \langle Y, \text{norm}\rangle$ *is equicontinuous.*

(c) $\mathscr{F}: \langle X_1, \text{weak topology}\rangle \rightarrow \langle Y, \text{norm}\rangle$ *is equicontinuous at the origin.*

(d) *There exists a sequence* $\{x_n^*\} \subseteqq X^*$ *such that* $\|x_n^*\| \rightarrow 0$ *and* $\{x \in X: |\langle x_n^*, x\rangle| \leqq 1$ *for each* $n\} \subseteqq \{x: \|\mathscr{F}(x)\| \leqq 1\}$.

(e) *There exists a sequence* $\{x_n^*\} \subseteqq X^*$ *such that* $\|x_n^*\| \rightarrow 0$ *and* $\mathscr{F}^*(Y_1^*)$ *is a subset of the closure of the balanced convex hull of* $\{x_n^*\}$.

Proof. The equivalence of (a) and (e) follows from Theorem 2.5.

The polar of the closure of the balanced convex hull of $\{x_n^*\}$ is $\{x: |\langle x_n^*, x\rangle| \leqq 1, n \geqq 1\}$. Also, the polar of $\mathscr{F}^*(Y_1^*)$ is $\{x: \|\mathscr{F}(x)\| \leqq 1\}$ since $|\langle \mathscr{F}^*(Y_1^*), x\rangle| \leqq 1$ if and only if $|\langle Y_1^*, \mathscr{F}(x)\rangle| \leqq 1$. The equivalence of (d) and (e) follows from these two observations.

(a) implies (b). By Lemma 2.3, there exists a Banach space Z, a compact operator $K \in [X, Z]$, and an equicontinuous collection $\mathscr{S} \subseteqq [Z, Y]$ such that $\mathscr{F} = \mathscr{S}K$. Since K is a compact operator

$$K: \langle X_1, \text{weak}\rangle \longrightarrow \langle Z, \text{norm}\rangle \text{ is continuous .}$$

Moreover, $\mathscr{S}: \langle Z, \text{norm}\rangle \rightarrow \langle Y, \text{norm}\rangle$ is equicontinuous. Hence, $\mathscr{F}: \langle X_1, \text{weak}\rangle \rightarrow \langle Y, \text{norm}\rangle$ is equicontinuous.

(b) implies (c). This implication is obvious.

(c) implies (a). By Theorem 2.5, it suffices to show that $\mathscr{F}^{**}: \langle X_1^{**}, \text{weak*}\rangle \rightarrow \langle Y^{**}, \text{norm}\rangle$ is equicontinuous at the origin. If J denotes the natural injection of X into X^{**}, (c) implies that

$$\mathscr{F}^{**}: \langle J(X_1), \text{weak*}\rangle \longrightarrow \langle Y^{**}, \text{norm}\rangle \tag{7}$$

is equicontinuous at the origin.

Let V be a 0-neighborhood in the norm topology of Y^{**}. Choose a 0-neighborhood U such that $\bar{U} \subseteqq V$, where the bar denotes closure in the norm topology of Y^{**}. By (8), choose W, a 0-neighborhood in the weak-star topology of X^{**}, such that

$$\mathscr{F}^{**}(J(X_1) \cap W) \subseteqq U . \tag{8}$$

Let T^{**} be an element of $\mathscr{F}^{**}$ and $x^{**} \in X_1^{**} \cap W$. Since $J(X_1)$ is weak-star dense in X_1^{**}, it is possible to choose a net $\{x_\alpha\} \subseteqq X_1$ such that the weak-star limit of $\{J(X_\alpha)\}$ is x^{**}. $\{J(x_\alpha)\}$ is eventually in W since $x^{**} \in W$. Therefore, $\{T^{**}(J(x_\alpha))\}$ is eventually in U, by (8). Since T^{**} is a compact operator, $\| T^{**}(x^{**})\text{-}T^{**}(J(x_\alpha)) \| \to 0$. Hence $T^{**}(x^{**}) \in \bar{U} \subseteqq V$. So in addition to (8),

$$\mathscr{F}^{**}(X_1^{**} \cap W) \subseteqq V .$$

In order to indicate how some previous results in the theory of collectively compact operators follow from our results, we prove the following lemma.

LEMMA 2.12. *Let* $L \subseteqq [X, Y]$ *be bounded in the uniform operator topology. The following are equivalent:*

(a) L: $\langle X_1, \text{weak}\rangle \to \langle Y, \text{norm}\rangle$ *is equicontinuous at the origin.*

(b) *For each* $\varepsilon > 0$, *there exists a subspace* $X(\varepsilon)$ *of finite codimension in* X *such that the restrictions of operators in* L *to* $X(\varepsilon)$ *have operator norms no greater than* ε.

Proof. Let the bound on $\mathscr{L} \subseteqq [X, Y]$ be M, i.e., $\| T \| \leqq M$ for each $T \in \mathscr{L}$.

(a) implies (b). Let $\varepsilon > 0$. By (a), there exists a finite set $\{x_1^*, \cdots, x_p^*\} \subseteqq X^*$ such that $\{x\colon |\langle x_i^*, x\rangle| \leqq 1,\ 1 \leqq i \leqq p\} \cap X_1 \subseteqq \{x\colon \| \mathscr{L}(x) \| \leqq \varepsilon\}$. Let $X(\varepsilon) = \{x\colon \langle x_i^*, x\rangle = 0,\ 1 \leqq i \leqq p\}$. Then (b) follows since $X(\varepsilon) \cap X_1 \subseteqq \{x\colon \| \mathscr{L}(x) \| \leqq \varepsilon\}$.

(b) implies (a). Let ε and $X(\varepsilon)$ be given. If $T \in \mathscr{L}$, then the operator norms of the restrictions of T to $X(\varepsilon)$ and $\overline{X(\varepsilon)}$ are the same. Consequently, we may assume that $X(\varepsilon)$ is a closed subspace of X.

Choose linearly independent $\{x_1, \cdots, x_p\} \subseteqq X$ such that $\| x_i \| = 1$ for each i and $X = X(\varepsilon) \oplus \operatorname{Span}\{x_i\}$. Since for each j, $X(\varepsilon) \oplus \operatorname{Span}\{x_1, \cdots, x_{j-1}, x_{j+1}, \cdots, x_p\}$ is a closed subspace which does not contain x_j, there exists $\{x_i^*\colon 1 \leqq i \leqq p\} \subseteqq X^*$ such that $x_i^*(X(\varepsilon)) = 0$, $1 \leqq i \leqq p$, and $x_i^*(x_j) = \delta_{i,j}$, $1 \leqq i, j \leqq p$. Consider the weak open set

$$W = \left\{x\colon \sum_{i=1}^{p} |\langle x_i^*, x\rangle| \leqq \min\{\varepsilon, 1\}\right\} .$$

If $x \in W \cap X_1$, then x has the representation

$$x = x_\varepsilon \oplus \sum_{i=1}^{p} \langle x_i^*, x\rangle x_i$$

with $x_\varepsilon \in X(\varepsilon)$. Since $\| x_i \| = 1$,

$$\| x_\varepsilon \| \leqq \| x \| + \sum_{i=1}^{p} | \langle x_i^*, x\rangle | \leqq 2 .$$

Then for $T \in L$, $\| T(X_\varepsilon) \| \leqq 2\varepsilon$ and

$$\| T(x) \| \leqq 2\varepsilon + \sum_{i=1}^{p} | \langle x_i^*, x\rangle | \, \| T(x_i) \| \leqq 2\varepsilon + M\varepsilon .$$

We have shown that for any $\varepsilon > 0$, $\{x: \| \mathscr{L}(x) \| \leqq 2\varepsilon + M\varepsilon\} \cap X_1$ is a relative weak neighborhood of the origin. Hence, statement (a) follows.

Finally, in view of Theorem 2.11, one obtains the result of Palmer [8] that for the collection $\mathscr{L}$ above, $\mathscr{L}^*$ is collectively compact if and only if condition (b) of the above lemma is satisfied.

References

1. P. M. Anselone, *Collectively Compact Operator Approximation Theory and Applications to Integral Operators*, Prentice-Hall, 1971.
2. P. M. Anselone and R. H. Moore, *Approximate solutions of integral and operator equations*, J. Math. Anal. Appl., **9** (1964), 268–277.
3. P. M. Anselone and T. W. Palmer, *Collectively compact sets of linear operators*, Pacific J. Math., **25** (1968), 417–422.
4. ———, *Spectral analysis of collectively compact, strongly convergent operator sequences*, Pacific J. Math., **25** (1968), 423–431.
5. J. D. DePree and J. A. Higgins, *Collectively compact sets of linear operators*, Math. Z., **115** (1970), 366–370.
6. N. Dunford and J. T. Schwartz, *Linear Operators*, Part I. New York: Interscience, 1958.
7. J. L. Kelly and I. Namioka, *Linear Topological Spaces*, Princeton: Van Nostrand, 1963.
8. T. W. Palmer, *Totally bounded sets of precompact linear operators*, Amer. Math. Soc., **29** No. 1, (1969), 101–106.

Received April 4, 1973.

New Mexico State University

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

SEMI-GROUPS AND COLLECTIVELY COMPACT SETS OF LINEAR OPERATORS

J. D. DEPREE AND H. S. KLEIN

A set of linear operators from one Banach space to another is collectively compact if and only if the union of the images of the unit ball has compact closure. Semi-groups $S = \{T(t): t \geqq 0\}$ of bounded linear operators on a complex Banach space into itself and in which every operator $T(t)$, $t > 0$ is compact are considered. Since $T(t_1 + t_2) = T(t_1)T(t_2)$ for each operator in the semi-group, it would be expected that the theory of collectively compact sets of linear operators could be profitably applied to semi-groups.

1. Introduction. Let X be a complex Banach space with unit ball X_1 and let $[X, X]$ denote the space of all bounded linear operators on X equipped with the uniform operator topology. The semi-group definitions and terminology used are those of Hille and Phillips [6]. Let S be a semi-group of vector-valued functions $T: [0, \infty) \rightarrow [X, X]$. It is assumed that $T(t)$ is strongly continuous for $t \geqq 0$. If $\lim_{t \to t_0} \| T(t)x - T(t_0)x \| = 0$ for each $t_0 \geqq 0$, $x \in X$ and if there is a constant M such that the $\| T(t) \| \leqq M$ for each $t \geqq 0$, then $S = \{T(t): t \geqq 0\}$ is called an equicontinuous semi-group of class C_0. The infinitesimal generator A of the semi-group S is defined by

$$Ax = \lim_{S \to 0} \frac{1}{S}[T(s)x - x]$$

whenever the limit exists. The domain $D(A)$ of A is a dense subset of X consisting of just those elements x for which this limit exists. A is a closed linear operator having resolvents $R(\lambda)$ which, for each complex number λ with the real part of λ greater than zero, are given by the absolutely summable Riemann-Stieltjes integral

$$R(\lambda)x = \int_0^\infty e^{-\lambda t} T(t)x dt, \; x \in X . \tag{1}$$

It follows from (1) that

$$\| R(\lambda) \| \leqq \frac{M}{re(\lambda)}, \; re(\lambda) > 0 . \tag{2}$$

In particular, sets of the type $\{R(\lambda): re(\lambda) \geqq \alpha > 0\}$ are equicontinuous subsets of $[X, X]$.

Results yielding the collective compactness of the resolvents of

A have recently been obtained independently by N. E. Joshi and M. V. Deshpande.

2. **Semi-groups of compact operators.** First, note that (1) states that the resolvents of A are Laplace transforms of the semi-group S. Consequently, there are many other important integral expressions involving the elements of the semi-group and the resolvents. In order to take advantage of these, we prove the following lemma, in which $|v|$ denotes the total variation of a complex measure v.

LEMMA 2.1. *Let Ω be a topological space and $\mathscr{M}$ a collection of complex-valued Borel measures on Ω. Suppose there exists a constant α for which $|v|\,\Omega \leqq \alpha$ for each $v \in \mathscr{M}$. Let $\mathscr{K}: \Omega \to [X, X]$ be an operator-valued function defined on Ω which is strongly measurable with respect to each $v \in M$* [6, page 74] *and suppose $\mathscr{K} = \{K(w): w \in \Omega\}$ is a bounded subset of $[X, X]$. For each $v \in \mathscr{M}$ and $x \in X$, let $F_v(x) = \int_\Omega K(w)x\,dv$, where the integral exists in the Bochner sense since $\int_\Omega \|K(w)x\|\,d|v| < \infty$* [6, page 80]. *Let $\mathscr{F} = \{F_v: v \in \mathscr{M}\}$. Whenever $\mathscr{K}(\mathscr{K}^*)$ is collectively compact, $\mathscr{F}(\mathscr{F}^*)$ is also collectively compact.*

Proof. Assume that $\mathscr{K}$ is collectively compact. Let $B = \{K(w)x: w \in \Omega,\ \|x\| \leqq 1\}$ and let C denote the balanced convex hull of B. Both B and C are totally bounded subsets of X. It suffices to show that $F_v(x) \in \alpha\bar{C}$ for any $F_v \in \mathscr{F}$ and x with $\|x\| \leqq 1$. Let $\varepsilon > 0$ and choose $\{K(w_1)x_1, \cdots, K(w_n)x_n\}$, an ε/α-net for B. For $i = 1, \cdots, n$, let $\Omega_i = \{w: \|K(w)x - K(w_i)x_i\| \leqq \varepsilon/\alpha\}$ and let $\Omega_i' = \Omega_j \backslash \bigcup_{j=1}^{i-1} \Omega_j$ be a decomposition of the Ω_i into pairwise disjoint sets. Then

$$\left\|F_v(x) - \sum_{i=1}^n K(w_i)x_i v(\Omega_i')\right\| \leqq \sum_{i=1}^n \int_{\Omega_i'} \|K(w)x - K(w_i)x_i\|\,d|v|(w) \leqq (\varepsilon/\alpha)\,|v|\,(\Omega) \leqq \varepsilon\,.$$

Since $\sum_{i=1}^n |v(\Omega_i')| \leqq \alpha$, $\sum_{i=1}^n K(w_i)x_i v(\Omega_i')$ is an element of αC. It follows that $F_v(x) \in \alpha\bar{C}$ and so $\mathscr{F}$ is also collectively compact.

Now assume that $\mathscr{K}^*$ is collectively compact. Let V be any neighborhood of 0 in the norm topology of X. There exists an $\varepsilon > 0$ such that $U = \{x: \|x\| \leqq \varepsilon\} \subseteqq V$. Since $\mathscr{K}^*$ is collectively compact, [2, Theorem 2.11, part (c)] implies that there exists a weak neighborhood W of the origin with $\mathscr{K}(W \cap X_1) \subseteqq (1/\alpha)U$. For $F_v \in \mathscr{F}$ and $x \in W \cap X_1$, $\|F_v(x)\| \leqq \int_\Omega \|K(w)x\|\,d|v| \leqq (\varepsilon/\alpha)\,|v|\,(\Omega) \leqq$

ε. So $\mathscr{F}(W \cap X_1) \subseteq V$. Again using [2, Theorem 2.1, part (c)], we see that $\mathscr{F}^*$ is also collectively compact.

The following is essentially a result of P. Lax [6, page 304]. Rephrased in the terminology of collectively compact sets of operators, it becomes quite transparent.

THEOREM 2.2. *Suppose that some* $T(t_0)$, $t_0 > 0$, *is a compact operator. Then* $\mathscr{K} = \{T(t): t \geqq t_0\}$ *is a totally bounded, collectively compact subset of* $[X, X]$. *Consequently,* $T(t)$ *is continuous in the uniform operator topology for* $t \geqq t_0$.

Proof. Since $T(t) = T(t - t_0)T(t_0) = T(t_0)T(t - t_0)$ for $t \geqq t_0$, it follows that $\mathscr{K} = T(t_0)\mathscr{S} = \mathscr{S}T(t_0)$. $T(t_0)$ is a compact operator and the collection $\mathscr{S}$ is equicontinuous. By Lemmas 2.1 and 2.3 of [2], both $\mathscr{K}$ and $\mathscr{K}^*$ are collectively compact. [2, Corollary 2.6] implies that $\mathscr{K}$ is a totally bounded subset of $[X, X]$. Since $T(t)$ is continuous in the strong operator topology, $T(t)$ is continuous in the uniform operator topology for $t \geqq t_0$.

COROLLARY 2.3. *Suppose every* $T(t)$, $t > 0$, *is a compact operator. Let* $\mathscr{F} = \{R(\lambda): re(\lambda) \geqq 1\}$ *be the collection of the resolvents of the infinitesimal generator* A *corresponding to the half-plane* $\{\lambda \in C: re(\lambda) \geqq 1\}$. *Then* $\mathscr{F}$ *is a totally bounded, collectively compact set of operators.*

It should be noted that for any $\alpha > 0$, the following arguments can be applied to $\{R(\lambda): re(\lambda) \geqq \alpha\}$. One particular half-plane is chosen simply to keep the notation as uncomplicated as possible.

Proof. It will suffice to show that for each $\varepsilon > 0$, there exists a totally bounded, collectively compact set of operators $\mathscr{K}$ such that for any $R(\lambda) \in \mathscr{F}$, there exists a $K \in \mathscr{K}$ with $\|R(\lambda) - K\| \leqq \varepsilon$. For this ε, choose $\delta > 0$ with $\int_0^\delta e^{-t}dt < \varepsilon/M$, where M is such that $\|T(\lambda)\| \leqq M$ for $t > 0$. Let λ be any complex number with $re(\lambda) \geqq 1$ and $x \in X$. Since $R(\lambda)x = \int_0^\infty e^{-\lambda t}T(t)xdt$, $\left\|R(\lambda)x - \int_\delta^\infty e^{-\lambda t}T(t)xdt\right\| \leqq \int_0^\delta e^{-\lambda t}\|T(t)x\|\,dt \leqq \int_0^\delta e^{-t}dtM\|x\| \leqq \varepsilon\|x\|$. Consequently, $\left\|R(\lambda) - \int_\delta^\infty e^{-\lambda t}T(t)dt\right\| \leqq \varepsilon$. Now $\mathscr{K} = \left\{\int_\delta^\infty e^{-\lambda t}T(t)dt: re(\lambda) \geqq 1\right\}$ is a totally bounded, collectively compact set of operators. To see this, note that $\sup\left\{\int_\delta^\infty |e^{-\lambda t}|\,dt: re(\lambda) \geqq 1\right\} \leqq 1$ and that both $\{T(t): t \geqq \delta\}$ and $\{T^*(t): t \geqq \delta\}$ are collectively compact. Lemma 2.1 implies that both

$\mathscr{K}$ and $\mathscr{K}^*$ are collectively compact. As before, [2, Corollary 2.6] implies that $\mathscr{K}$ is a totally bounded subset of $[X, X]$.

The following lemma will be useful in the next section. Since a quotable reference cannot be found, a brief proof is included.

LEMMA 2.4. *Let $\mathscr{S}$ be an equicontinuous semi-group of class C_0. Then $R(\lambda)$ converges to zero in the strong operator topology as $|\lambda| \to \infty$, $re(\lambda) \geqq 1$. Whenever $\{R(\lambda)\colon re(\lambda) \geqq 1\}$ is a totally bounded subset of $[X, X]$, the $R(\lambda)$ converge to zero in the uniform operator topology as $|\lambda| \to \infty$, $re(\lambda) \geqq 1$.*

Proof. The second assertion follows immediately from the first.

Let $x \in D(A)$, the domain of the infinitesimal generator A. Since $R(\lambda)(\lambda - A)x = x$, we have the identity

$$R(\lambda)x = \frac{1}{\lambda}[x + R(\lambda)Ax] .$$

By (2) of § 1, $\{R(\lambda)Ax\colon re(\lambda) \geqq 1\}$ is a bounded subset of X. It follows that $\| R(\lambda)x \| \to 0$ as $|\lambda| \to \infty$, $re(\lambda) \geqq 1$, for each $x \in D(A)$. Since $D(A)$ is dense in X, the Banach-Steinhaus theorem implies that this type of convergence holds for each $x \in X$. We see that the first assertion of this lemma holds also.

3. Semi-groups with compact resolvents. Suppose that the domain of the infinitesimal generator of a semi-group can be given a topology τ such that the topological space $\langle D(A), \tau \rangle$ is a Banach space and the natural injection $i\colon \langle D(A), \tau \rangle \to X$ is a compact operator. In such cases, it might be possible to prove that certain sets of the resolvents of A are equicontinuous subsets of $[X, \langle D(A), \tau \rangle]$, i.e., collectively compact subsets of $[X, X]$. A specific example is the case in which X is some L^p space and A is the negative of a uniformly strongly elliptic differential operator defined on a Sobolev space $H = \langle D(A), \tau \rangle$. The so-called "a priori inequalities" [4, Theorems 18.2 and 19.2, pages 69 and 77] imply that, after a suitable translation, $\{R(\lambda)\colon re(\lambda) \geqq 1\}$ is an equicontinuous subset of $[L^p, H]$. Since the injection $i\colon H \to L^p$ is a compact operator [4, Theorem 11.2, page 31], $\{R(\lambda)\colon re(\lambda) \geqq 1\}$ is a collectively compact subset of $[L^p, L^p]$. The obvious question is what are the implications of such assumptions for a general semi-group $\mathscr{S}$.

We first consider the case in which A has one compact resolvent. Of course, the first resolvent equation,

$$R(\lambda_1) - R(\lambda_2) = (\lambda_2 - \lambda_1)R(\lambda_1)R(\lambda_2) ,$$

then implies that all resolvents of A are compact operators.

LEMMA 3.1. *Suppose A has one compact resolvent. Let Ω be a compact subset of $\{\lambda: re(\lambda) > 0\}$. Then $\{R(\lambda): \lambda \in \Omega\}$ is collectively compact.*

Proof. Since $R(\lambda)$ is a holomorphic function in the right half-plane, $\{R(\lambda): \lambda \in \Omega\}$ is a totally bounded subset of $[X, X]$. Each element in this collection is a compact operator. So [2, Corollary 2.7] implies that $\{R(\lambda): \lambda \in \Omega\}$ is collectively compact.

The following is a partial converse of Theorem 2.2.

PROPOSITION 3.2. *Suppose A has compact resolvents. Let $t_0 > 0$. If $T(t)$ is continuous in the uniform operator topology for $t \in [t_0, \infty)$, then $T(t_0)$ is a compact operator.*

Proof. Since the resolvents are Laplace transforms of $\{T(t): t \geqq 0\}$, we may use the formula based upon fractional integration of order two [6, page 220] which states that

$$\int_0^s (s - t)T(t)dt = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} \frac{e^{\lambda s}}{\lambda^2} R(\lambda)d\lambda,\ s > 0 .$$

For $\varepsilon > 0$, choose N such that

$$\int_{1-i\infty}^{1-iN} + \int_{1+iN}^{1+i\infty} \frac{1}{|\lambda^2|} \| e^{\lambda s}R(\lambda) \| \, d\,|\lambda| < \varepsilon .$$

Then

$$\left\| \int_0^s (s - t)T(t)dt - \frac{1}{2\pi i}\int_{1-iN}^{1+iN} \frac{e^{\lambda s}}{\lambda^2} R(\lambda)d\lambda \right\| < \varepsilon .$$

By Lemmas 3.1 and 2.1, the integral of $(e^{\lambda s}/\lambda^2)R(\lambda)$ over the finite segment of the vertical line is a compact operator. It follows that for each $s \geqq 0$, $\int_0^s (s - t)T(t)dt$ is a compact operator.

Consider the function

$$F(s) = \int_0^s (s - t)T(t)dt,\ s \geqq 0 .$$

Each value of F is a compact operator. Elementary calculations show that F is differentiable in the uniform operator topology. Consequently, each

$$F'(s) = \int_0^s T(t)dt,\ s \geqq 0\ ,$$

is the limit in the uniform operator topology of a sequence of compact operators. Hence, each $F'(s)$, $s \geqq 0$, is a compact operator. In taking derivatives again, we see that for $h > 0$,

$$\left\| \frac{1}{h}\int_{t_0}^{t_0+h} T(t)dt - T(t_0) \right\| \leqq \sup\{\| T(t_0 + \alpha) - T(t_0)\|: 0 \leqq \alpha \leqq h\}\ .$$

If $T(t_0 + \alpha)$ is continuous in the uniform operator topology for $\alpha \geqq 0$, then

$$T(t_0) = \text{uniform} - \lim_{h\to 0+} \frac{1}{h}\int_{t_0}^{t_0+h} T(t)dt\ .$$

It follows that $T(t_0)$ is a compact operator.

See [6, page 537] for a discussion of the following example.

EXAMPLE 3.3. Consider the semi-group $\mathscr{S}$ of left translations on the space $C_0[0, 1]$ consisting of continuous functions $x(u)$ vanishing at 1, where the norm $\| x \| = \sup\{| x(u)|: 0 \leqq u \leqq 1\}$. Let $[T(t)x](u) = x(u + t)$, for $0 \leqq u \leqq \max\{0, 1 - t\}$, and 0 for $\max\{0, 1 - t\} \leqq u \leqq 1$. The infinitesimal generator of $\mathscr{S}$ is the operator of differentiation $d/(du)$ with domain

$$D\left(\frac{d}{du}\right) = \{x: x' \in C_0[0, 1]\}\ .$$

The compact resolvents are given by

$$[R(\lambda)x](u) = \int_0^{1-u} e^{-\lambda t}x(u + t)dt,\ \lambda \in \boldsymbol{C}\ .$$

For $t \geqq 1$, $T(t)$ is the compact operator 0 while for $t, s < 1$, $\| T(t) - T(s)\| = 2$. This can easily be seen by evaluating $T(t) - T(s)$ at a function $x \in C_0[0, 1]$ with $\| x \| \leqq 1$ and $x(t) = 1$, $x(s) = -1$. So $T(t)$ is continuous in the uniform operator topology only for $t \geqq 1$.

Choose a monotonically increasing sequence of positive functions $\{y_n\} \subseteqq C_0[0, 1]$ such that $\lim_n y_n(u) = 1$ for each $u < 1$. For $t < 1$, $\{T(t)y_n\}$ is a sequence of functions having no subsequence which can converge uniformly. So $T(t)$, $t < 1$, is not a compact operator.

For $\lambda = \sigma + i\tau$, let $x_n(u) = e^{i\tau u}y_n(u)$ in the definition of $R(\lambda)$. We see that

$$[R(\lambda)x_n](0) = \int_0^1 e^{-\sigma t}y_n(t)dt\ .$$

Since $\|x_n\| = 1$ for each n,

$$\|R(\lambda)\| \geqq \sup_n |[R(\lambda)x_n](0)| = \int_0^1 e^{-\sigma t}dt .$$

It follows immediately from the definition of $R(\lambda)$ that the reverse inequality holds also. Consequently, $\|R(\lambda)\| = \int_0^1 e^{-\sigma t}dt$. In particular, $\lim_{|\tau|\to\infty} \|R(\sigma + i\tau)\| \neq 0$. This serves to distinguish this differential operator from the class of infinitesimal generators which we consider next.

LEMMA 3.4. *Suppose $\mathscr{S}$ is a semi-group such that the set of resolvents $\{R(\lambda): re(\lambda) = 1\}$ corresponding to the vertical line $re(\lambda) = 1$ is collectively compact. Then $\{R(\lambda): re(\lambda) \geqq 1\}$ is also collectively compact.*

Proof. For each $x \in X$, $R(\lambda)x$ is a holomorphic and bounded function of λ, $re(\lambda) > 1/2$. So $R(\lambda)x$ admits Poisson's integral representation [6, page 229]

$$R(\sigma + i\tau)x = \frac{\sigma - 1}{\pi}\int_{-\infty}^{\infty} \frac{R(1 + i\beta)x}{(\sigma - 1)^2 + (\tau - \beta)^2}d\beta$$

for $\sigma > 1$, $x \in X$. Since $\{R(1 + i\beta): -\infty < \beta < \infty\}$ is collectively compact and the integral of the Poisson kernel over $-\infty < \beta < \infty$ is identically one, Lemma 2.1 implies that $\{R(\lambda): re(\lambda) > 1\}$ is collectively compact. Taking the union of this set and $\{R(\lambda): re(\lambda) = 1\}$, one obtains the desired result.

For $x \in X$ and $x^* \in X^*$,

$$\langle x^*, R(\sigma + i\tau)x\rangle = \int_0^\infty e^{-i\tau t}(e^{-\sigma t}\langle x^*, T(t)x\rangle)dt .$$

This is this Fourier transform of the absolutely summable function $e^{-\sigma t}\langle x^*, T(t)x\rangle$, $t \geqq 0$. The convergence of

$$\|R(\sigma + i\tau)\| = \sup\{|\langle x^*, R(\sigma + i\tau)x\rangle| : \|x\|, \|x^*\| \leqq 1\}$$

to 0 as $|\sigma|$ and $|\tau|$ approach infinity can be viewed as a "uniform" Riemann-Lebesgue lemma.

THEOREM 3.5. *If $\mathscr{T} = \{R(\lambda): re(\lambda) \geqq 1\}$ is collectively compact, then $\|R(\lambda)\|$ converges to 0 as $|\lambda|$ approaches ∞, $re(\lambda) \geqq 1$.*

Proof. Throughout the following proof, we assume that $re(\lambda) \geqq 1$.

Let $\varepsilon > 0$ be given and choose real β so large that $1 + \beta \geqq M/\varepsilon$, where M is the constant in § 1 which bounds the operator norms of elements of $\mathscr{S}$. By (2),

$$\| R(\lambda + \beta) \| \leqq \frac{M}{re(\lambda) + \beta} \leqq \frac{M}{1 + \beta} \leqq \varepsilon .$$

In view of Lemma 2.4, $\mathscr{F}$ is an equicontinuous collection with $R(\lambda)$ converging to zero as $|\lambda| \to \infty$ pointwise on the relatively compact set $\mathscr{F}(X_1)$. Therefore, $\| R(\lambda)F \| \to 0$ as $|\lambda| \to \infty$ uniformly for $F \in \mathscr{F}$. Choose N such that $|\lambda| \geqq N$ implies that

$$\| R(\lambda)R(\lambda + \beta) \| \leqq \varepsilon/\beta .$$

The first resolvent equation states that

$$R(\lambda) - R(\lambda + \beta) = (\lambda + \beta - \lambda)R(\lambda)R(\lambda + \beta) .$$

So, for $|\lambda| \geqq N$,

$$\| R(\lambda) \| \leqq \| \beta R(\lambda)R(\lambda + \beta) \| + \| R(\lambda + \beta) \| \leqq 2\varepsilon .$$

Note that we have used the fact that $\mathscr{F}$ contains those resolvents $R(\lambda)$ with $re(\lambda)$ arbitrarily large in an essential way.

COROLLARY 3.6. *Let $\mathscr{S}$ be any semi-group whose infinitesimal generator A has compact resolvents, i.e., each $R(\lambda)$, $re(\lambda) > 0$, is a compact operator. Then $\mathscr{F} = \{R(\lambda): re(\lambda) \geqq 1\}$ is collectively compact if and only if $\| R(\lambda) \| \to 0$ as $|\lambda| \to \infty$, $re(\lambda) \geqq 1$.*

Proof. The assumption that $\| R(\lambda) \| \to 0$ as $|\lambda| \to \infty$, $re(\lambda) \geqq 1$, simply implies that $R(\lambda)$ can be extended to a continuous function on the compactification of the half-plane $\{\lambda: re(\lambda) \geqq 1\}$. Consequently, if A has compact resolvents, $\mathscr{F}$ is a totally bounded set of compact operators. [2, Corollary 2.7] implies that $\mathscr{F}$ is collectively compact.

The converse is simply Theorem 3.5.

The behavior of the holomorphic function $R(\lambda)$ on the vertical line $re(\lambda) = 1$ is of fundamental importance. For example, if $d(\lambda)$ denotes the distance of the complex number λ from the spectrum of A, then [3, page 566]

$$d(1 + i\tau) \geqq \frac{1}{\| R(1 + i\tau) \|} .$$

We see that the spectrum of A must be bounded on the right by the curve

$$\gamma(\tau) = 1 - \frac{1}{\| R(1 + i\tau) \|} + i\tau, \ -\infty < \tau < \infty \ .$$

In particular, it follows from Theorem 3.5 and Lemma 3.4 that when $\{R(\lambda): re(\lambda) = 1\}$ is collectively compact, the spectrum of A is severely restricted.

The usual methods of inverting Fourier transforms can be typified by the use of $(C, 1)$ means. In [5, page 350], it is shown that for each $t > 0$

$$T(t) = \lim_{w \to \infty} \frac{1}{2\pi} \int_{-w}^{w} \left(1 - \frac{|\tau|}{w}\right) e^{(1+i\tau)t} R(1 + i\tau) d\tau \ .$$

However, the measures involved no longer satisfy the requirements of Lemma 2.1. As this situation is typical, we are not able to prove that if $\{R(\lambda): re(\lambda) = 1\}$ is collectively compact, then each $T(t) \in \mathscr{S}$, $t > 0$, is a compact operator.

REFERENCES

1. P. M. Anselone and R. H. Moore, *Approximate solutions of integral and operator equations*, J. Math. Anal. Appl., **9** (1964), 268-277.
2. J. D. DePree and H. S. Klein, *Characterizations of collectively compact set, of linear operators*, Pacific J. Math., **55** (1974), 45-54.
3. N. Dunford and J. T. Schwartz, *Linear Operators*, Part I, New York: Interscience, 1958.
4. A. Friedman, *Partial Differential Equations*, New York: Holt, Rinehart and Winston, 1969.
5. A. Grothendieck, *Sur les Applications Lineaires Faiblement Compactes D'Epaces Du Type C(K)*, Canad. J. Math., **5** (1953), 129-173.
6. E. Hille and R. Phillips, *Functional Analysis and Semi-Groups*, Providence, R. I.: American Mathematical Society, 1957.

Received April 4, 1973.

NEW MEXICO STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

CHAIN BASED LATTICES

G. EPSTEIN AND A. HORN

In recent years several weakenings of Post algebras have been studied. Among these have been P_0-lattices by T. Traczyk, Stone lattice of order n by T. Katrinak and A. Mitschke, and P-algebras by the present authors. Each of these system is an abstraction from certain aspects of Post algebras, and no two of them are comparable. In the present paper, the theory of P_0-lattices will be developed further and two new systems, called P_1-lattices and P_2-lattices are introduced. These systems are referred to as chain based lattices. P_2-lattices form the intersection of all three weakenings mentioned above. While P-algebras and weaker systems such as L-algebras, Heyting algebras, and B-algebras, do not require any distinguished chain of elements other than 0, 1, chain based lattices require such a chain.

Definitions are given in § 1. A P_0-lattice is a bounded distributive lattice A which is generated by its center and a finite subchain containing 0 and 1. Such a subchain is called a chain base for A. The order of a P_0-lattice A is the smallest number of elements in a chain base of A. In § 2, properties of P_0-lattices are given which are used in later sections. If a P_0-lattice A is a Heyting algebra, then it is shown in § 3, that there exists a unique chain base $0 = e_0 < e_1 < \cdots < e_{n-1} = 1$ such that $e_{i+1} \to e_i = e_i$ for all $i > 0$. A P_0-lattice with such a chain base is called a P_1-lattice. Every P_1-lattice of order n is a Stone lattice of order n. If a P_1-lattice is pseudo-supplemented then it is called a P_2-lattice. It turns out that P_2-lattices of order n are direct products of finitely many Post algebras whose maximum order is n. In § 4, properties of P_2-lattices are studied. In § 5, equational axioms are given for P_2-lattices. P_2-lattices share many of the properties of Post algebras and have application to computer science. Among examples of P_2-lattices are direct products of finitely many p-rings. These further remarks on P_2-lattices are in § 6. In § 7, prime ideals in P_0-lattices are studied. It is shown that the order of a P_0-lattice is one more than the number of elements in a chain of prime ideals of maximum length. A characterization of P_1-lattices by properties of their prime ideals is given. Such a characterization of P_2-lattices is also indicated.

1. DEFINITIONS. We use ϕ for the empty set. Let A be a distributive lattice which is *bounded*, that is, has a largest element 1 and a smallest element 0. The dual of A is denoted by A^d. The

complement of x is denoted by $\bar{x}$ or $-x$. The *center* of A is the set B of all complemented elements of A. We use $x \vee y$ for the join, and $x \wedge y$ or xy for the meet of two elements x, y in A. $x \rightarrow y$ denotes the largest $z \in A$ (if it exists) such that $xz \leqq y$. A is called a *Heyting algebra* if $x \rightarrow y$ exists for all $x, y \in A$. $\neg x = x \rightarrow 0$ is called the *pseudo-complement* of x (when it exists). An element x is called *dense* if $\neg x = 0$. A is called *pseudo-complemented* if $\neg x$ exists for all $x \in A$. $x \Rightarrow y$ denotes the largest $z \in B$ such that $xz \leqq y$. A is called a *B-algebra* if $x \Rightarrow y$ exists for all $x, y \in A$. $!x = 1 \Rightarrow x$ is called the *pseudo-supplement* of x. A is called pseudo-supplemented if $!x$ exists for all $x \in A$.

A *Stone lattice* is a pseudo-complemented lattice satisfying the identity $\neg x \vee \neg\neg x = 1$. An *L-algebra* is a Heyting algebra satisfying $(x \rightarrow y) \vee (y \rightarrow x) = 1$. A *P-algebra* is a B-algebra satisfying $(x \Rightarrow y) \vee (y \Rightarrow x) = 1$. We denote the interval $\{z: x \leqq z \leqq y\}$ by $[x, y]$. A is an L-algebra if and only if every interval in A is a Stone lattice [1, 3.11]. The identity $x \rightarrow (y \vee z) = (x \rightarrow y) \vee (x \rightarrow z)$ is satisfied in an L-algebra. The identity $x \Rightarrow (y \vee z) = (x \Rightarrow y) \vee (x \Rightarrow z)$ is satisfied in a P-algebra.

2. *P_0-lattices.* Let A be a bounded distributive lattice and let B be a Boolean subalgebra of the center of A. A *chain base* of A is a finite sequence $0 = e_0 \leqq e_1 \leqq \cdots \leqq e_{n-1} = 1$ such that A is generated by $B \cup \{e_0, \cdots, e_{n-1}\}$. If A has a chain base then A is called a P_0-lattice [13], in which case every element $x \in A$ can be written in the form

$$x = \bigvee_{i=1}^{n-1} b_i e_i \,, \tag{1}$$

where $b_i \in B$. If $b_i \geqq b_{i+1}$ for all i, then (1) is called a *monotone representation* (abbreviated mon. rep.) of x. If $b_i b_j = 0$ for $i \neq j$, then (1) is called a *disjoint representation* (disj. rep.) of x. Every element in a P_0-lattice has both a mon. rep. and a disj. rep.

LEMMA 2.1. *If* (1) *is a mon. rep. of x and $y = \bigvee_i c_i e_i$ is a mon. rep., then $x \vee y = \bigvee_i (b_i \vee c_i) e_i$ and $xy = \bigvee_i b_i c_i e_i$ are mon. reps.*

Proof. This follows from the distributivity of A.

The following theorem shows that B must coincide with the center of the P_0-lattice A, and gives a method for constructing P_0-lattices.

THEOREM 2.2. *Let A be a bounded distributive lattice. Let*

B be a subalgebra of the center of A and let $0 = e_0 \leqq \cdots \leqq e_{n-1} = 1$. If A_0 is the sublattice generated by $B \cup \{e_0, \cdots, e_{n-1}\}$, and B_0 is the center of A_0, then $B_0 = B$.

Proof. Let $x = \bigvee_j b_j e_j$ be a disj. rep. of an element $x \in B_0$. For each i, $xb_i = b_i e_i$ is in B_0. Let $\bigvee_j c_j e_j = 0$ be a mon. rep. of the complement of $b_i e_i$. Then $b_i e_i \bigvee_j c_j e_j = 0$ implies $b_i c_i e_i = 0$, hence $b_i e_i \leqq b_i \bar{c}_i$. Also $1 = b_i e_i \vee \bigvee_j c_j e_j$ implies $1 \leqq e_i \vee c_i$, hence $b_i \bar{c}_i \leqq b_i e_i$. Thus $b_i e_i = b_i \bar{c}_i \in B$ for all i, and so $x \in B$.

DEFINITION 2.3. A P_0-lattice A is said to be of *order* n if n is the smallest integer such that A has a chain base with n terms.

LEMMA 2.4. *If $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_0-lattice, then $\langle A^d; e_{n-1}, \cdots, e_0\rangle$ is a P_0-lattice. A^d has the same order as A.*

Proof. This is obvious by inspection.

THEOREM 2.5. *If $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_0-lattice with center B and $A' = [e_i, e_j]$, where $i \leqq j$, then $\langle A'; e_i, \cdots, e_j\rangle$ is a P_0-lattice with center $B' = \{e_i \vee be_j : b \in B\}$. If $e_i = f_0 \leqq \cdots \leqq f_{r-1} = e_j$ is a chain base of A', then $e_0, \cdots, e_{i-1}, f_0, \cdots, f_{r-1}, e_{j+1}, \cdots, e_{n-1}$ is a chain base of A. If A has order n, then A' has order $j - i + 1$.*

Proof. Let $x = \bigvee_{k=1}^{n-1} b_k e_k$ be a mon. rep. of an element $x \in A'$. Then

$$x = (e_i \vee x)e_j = e_i \vee \bigvee_{k=i+1}^{j} b_k e_k = \bigvee_{k=i+1}^{j} (e_i \vee b_k e_j)e_k .$$

B' is clearly a subalgebra of the center of A'. Therefore by 2.2, B' is the center of the P_0-lattice $\langle A'; e_i, \cdots, e_j\rangle$. The remaining parts of the theorem hold because if $i \leqq k \leqq j$, then e_k is in the sublattice generated by $B' \cup \{f_0, \cdots, f_{r-1}\}$.

LEMMA 2.6. *Let A be a bounded distributive lattice with center B, and $x, y, z \in A$.*

(i) *If $x \to z$ and $y \to z$ exist, then $(x \vee y) \to z = (x \to z)(y \to z)$.*

(ii) *If $z \to x$ and $z \to y$ exist, then $z \to xy = (z \to x)(z \to y)$.*

(iii) *If $x \to y$ exists, $b \in B$ and $c \in B$, then $bx \to (c \vee y) = \bar{b} \vee c \vee (x \to y)$.*

(iv) *If $x \Rightarrow z$ and $y \Rightarrow z$ exist, then $(x \vee y) \Rightarrow z = (x \Rightarrow z)(y \Rightarrow z)$.*

(v) *If $z \Rightarrow x$ and $z \Rightarrow y$ exist, then $z \Rightarrow xy = (z \Rightarrow x)(z \Rightarrow y)$.*

(vi) *If $x \Rightarrow y$ exists, $b \in B$ and $c \in B$, then $bx \Rightarrow (c \vee y) = \bar{b} \vee c \vee (x \Rightarrow y)$.*

Proof. The proof is straightforward.

LEMMA 2.7. *If* $a_2 \leqq \cdots \leqq a_m$ *and* $b_1 \geqq \cdots \geqq b_{m-1}$ *are elements of a distributive lattice, then* $\bigvee_{j=1}^{m-1} a_{j+1}b_j = a_m b_1 \bigwedge_{j=2}^{m-1} (a_j \vee b_j)$.

Proof. This is easily proved by induction.

THEOREM 2.8. *Let* $\langle A; e_0, \cdots, e_{n-1} \rangle$ *be a* P_0*-lattice with center* B. *Then the following are equivalent:*

(i) $e_i \Rightarrow 0$ *exists for all* i.

(ii) $\neg e_i$ *exists for all* i.

(iii) A *is pseudo-complemented.*

(iv) A *is a Stone lattice.*

(v) *Each* $x \in A$ *has a mon. rep.* $x = \bigvee_i b_i e_i$ *such that* $b_1 \leqq c_1$ *for every mon. rep.* $x = \bigvee_i c_i e_i$.

Proof. (i) implies (ii): Let $xe_i = 0$ and suppose $x = \bigvee_j b_j e_j$ is a mon. rep. of x. Then $b_j e_j = 0$ for $j \leqq i$, while if $j > i$, then $b_j e_i = 0$, so $b_j \leqq e_i \Rightarrow 0$. Hence $x \leqq e_i \Rightarrow 0$. Therefore, $\neg e_i$ exists and equals $e_i \Rightarrow 0$.

(ii) implies (iii): If $x = \bigvee_i b_i e_i$ is a mon. rep., then by 2.6(i) and 2.6(iii), $\neg x$ exists and equals $\bigwedge_i (\bar{b}_i \vee \neg e_i)$. If follows from 2.7 that

$$\neg x = \bigvee_{i=1}^{n-1} \bar{b}_i \neg e_{i-1} \,. \tag{2}$$

(iii) implies (iv): If $x, y \in A$, then by 2.1 and (2), $\neg(xy) = \neg x \vee \neg y$. This implies that A is a Stone lattice [8].

(iv) implies (v): If $x = \bigvee_i c_i e_i$ is any mon. rep., then $\bar{c}_1 x = 0$, so $\bar{c}_1 \leqq \neg x$, hence $x \leqq \neg\neg x \leqq c_1$. Therefore $x = \bigvee_i (c_i \neg\neg x) e_i$. If we set $b_i = c_i \neg\neg x$, we get a mon. rep. in which $b_1 = \neg\neg x$.

(v) implies (i): Let $e_i = \bigvee_j b_j e_j$ be a mon. rep. of e_i having the property stated in (v). Then $\bar{b}_1 e_i = 0$. If $b \in B$ and $be_i = 0$, then $e_i \leqq \bar{b}$, so $e_i = \bigvee_j \bar{b} b_j e_j$. By hypothesis, $\bar{b} b_1 \geqq b_1$. Therefore $b \leqq \bar{b}_1$, and so $e_i \Rightarrow 0 = \bar{b}_1$.

LEMMA 2.9. *If* A *is a bounded distributive lattice, then* A^d *is a Stone lattice if and only if* A *is pseudo-supplemented and* $!(x \vee y) = !x \vee !y$ *for all* $x, y \in A$.

Proof. It is easily verified that the pseudo-complement of x in A^d is $\overline{!x}$ in this case.

THEOREM 2.10. *Let* $\langle A; e_0, \cdots, e_{n-1} \rangle$ *be a* P_0*-lattice. Then the* *following are equivalent*:

(i) $!e_i$ *exists for all* i.

(ii) A *is pseudo-supplemented and* $!(x \vee y) = !x \vee !y$ *for all* $x, y \in A$.

(iii) *Each* $x \in A$ *has a mon. rep.* $\bigvee_i b_i e_i$ *such that* $b_{n-1} \geqq c_{n-1}$ *for every mon. rep.* $x = \bigvee_i c_i e_i$.

Proof. This is derived from 2.8 by using 2.4 and 2.9.

THEOREM 2.11. *Let* $\langle A; e_0, \cdots, e_{n-1} \rangle$ *be a pseudo-complemented* P_0*-lattice. Then* A *has a chain base* $0 = f_0 \leqq f_1 \leqq \cdots \leqq f_{n-1} = 1$ *such that* f_1 *is the smallest dense element of* A. *If* $0 = g_0 \leqq \cdots \leqq g_{r-1} = 1$ *is any chain base of* A *such that* g_1 *is dense, then* $g_1 = f_1$ *and for any mon. rep.* $x = \bigvee_{j=1}^{r-1} b_j g_j$, *we have* $\neg x = \bar{b}_1$.

Proof. Let $f_0 = 0$, $f_1 = \bigvee_{i=1}^{n-1} (\neg e_{i-1}) e_i$, and $f_i = e_i \vee f_1$ for $i \geqq 2$. By (2), $\neg f_1 = \bigvee_i \neg\neg\, e_{i-1} \neg\, e_i = 0$. Also $f_i = e_i \vee \bigvee_{j>i} e_j \neg\, e_{j-1}$. Therefore $f_i \neg\neg\, e_i = e_i$, since $\neg\neg e_i \neg\, e_{j-1} \leqq \neg\neg\, e_i \neg\, e_i = 0$ for $j > i$. If $x = \bigvee_i b_i e_i$ is any element of A, then $x = \bigvee_i (b_i \neg\neg\, e_i) f_i$. Thus $f_0, \cdots, f_{n-1}$ is a chain base of A. Let $g_0, \cdots, g_{r-1}$ be a chain base of A such that g_1 is dense. If $x = \bigvee_j b_j g_j$ is a mon. rep., then $\neg x = \bar{b}_1$ by (2). So if x is dense, then $x \geqq g_1$. Thus $g_1 = f_1$ is the smallest dense element of A.

3. P_1-*lattices*.

THEOREM 3.1. *Let* $\langle A; e_0, \cdots, e_{n-1} \rangle$ *be a* P_0*-lattice with center* B. *Then the following are equivalent*:

(i) $e_i \rightarrow e_j$ *exists for all* i, j.

(ii) A *is a Heyting algebra.*

(iii) A *is an* L *algebra.*

Proof. (i) implies (ii): If $x = \bigvee_i b_i e_i$ and $y = \bigvee_i c_i e_i$ are mon. reps., then by 2.7, $y = \bigwedge_{i=1}^{n-1} (c_i \vee e_{i-1})$. Therefore by 2.6, $x \rightarrow y$ exists and equals $\bigwedge_{i,j} (\bar{b}_i \vee c_j \vee (e_i \rightarrow e_j))$.

(ii) implies (iii): Let $x = \bigvee_i b_i e_i$, $y = \bigvee_i c_i e_i$ be mon. reps. Then $x \rightarrow y = \bigwedge_i (b_i e_i \rightarrow y) \geqq \bigwedge_i (\bar{b}_i \vee c_i)$. Therefore, $(x \rightarrow y) \vee (y \rightarrow x) \geqq \bigwedge_i (\bar{b}_i \vee c_i) \vee \bigwedge_i (b_i \vee \bar{c}_i) = \bigwedge_{i,j} (\bar{b}_i \vee b_j \vee c_i \vee \bar{c}_j) = 1$ since $\bar{b}_i \vee b_j = 1$ for $i \geqq j$, and $c_i \vee \bar{c}_j = 1$ for $i < j$.

(iii) implies (i): This is obvious.

DEFINITION 3.2. A P_1-lattice $\langle A; e_0, \cdots, e_{n-1} \rangle$ is a P_0-lattice together with a chain base such that $e_{i+1} \to e_i = e_i$. It follows that $e_i \to e_j = e_j$ for $i > j$ and $e_i \to e_j = 1$ for $i \leqq j$, so that (i) of 3.1 holds.

THEOREM 3.3. *If* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_0*-lattice and* A *is a Heyting algebra, then there exists a chain base* $0 = f_0 \leqq \cdots \leqq f_{n-1} = 1$ *such that* $\langle A; f_0, \cdots, f_{n-1}\rangle$ *is a* P_1*-lattice.*

Proof. This is obvious for $n = 1, 2$. Suppose $n > 2$ and the statement holds for $n - 1$. By 2.11, we may assume e_1 is dense. Let $A' = [e_1, 1]$. By 2.5, $\langle A'; e_1, \cdots, e_{n-1}\rangle$ is a P_0-lattice. If $x, y \in A'$, then $x \to y \in A'$. Therefore by the induction hypothesis, there exists a sequence $e_1 = f_1 \leqq \cdots \leqq f_{n-1} = 1$ such that $\langle A'; f_1, \cdots, f_{n-1}\rangle$ is a P_1-lattice. If we set $f_0 = 0$, then by 2.5, $\langle A; f_0, \cdots, f_{n-1}\rangle$ is a P_1-lattice.

THEOREM 3.4. *Let* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *be a* P_1*-lattice. Then for some* $m \geqq 1$, $0 = e_0 < e_1 < \cdots < e_{m-1} = e_m = \cdots = 1$. *A has order* m. *For each* i, e_{i+1} *is the smallest dense element of* $[e_i, 1]$. *Thus* $e_0, \cdots, e_{m-1}$ *is the unique strictly increasing chain such that* $\langle A; e_0, \cdots, e_{m-1}\rangle$ *is a* P_1*-lattice. If* $x = \bigvee_{i=1}^{n-1} b_i e_i$ *is a mon. rep., then* $e_i \vee b_{i+1} = (x \to e_i) \to e_i$, $0 \leqq i < n - 1$. *If* $x = \bigvee_{i=1}^{n-1} b_i e_i$ *is a disj. rep., and* $y = \bigvee_{i=1}^{n-1} c_i e_i$ *is a mon. rep., then* $x \to y = y \vee \bigvee_{i=0}^{n-1} b_i c_i$, *where* $b_0 = \bigwedge_{i=1}^{n-1} \bar{b}_i$, $c_0 = 1$.

Proof. If m is the first integer such that $e_m = e_{m-1}$, then $e_{m-1} = e_m \to e_{m-1} = 1$. Since e_{i+1} is dense in $[e_i, 1]$ it follows from 2.5 and 2.11 that e_{i+1} is the smallest dense element of $[e_i, 1]$. Using 3.3, it follows that A has order m.

If $x = \bigvee_{i=1}^{n-1} b_i e_i$ is a mon. rep., then $x \vee e_i = \bigvee_{k=i+1}^{n-1} (e_i \vee b_k)e_k$. Applying 2.11 to $[e_i, 1]$, we find $(x \vee e_i) \to e_i = e_i \vee \bar{b}_{i+1}$. Since $(x \vee e_i) \to e_i = x \to e_i$, it follows by 2.6 that $(x \to e_i) \to e_i = e_i \vee b_{i+1}$.

To prove the last statement, we observe that

$$\begin{aligned} e_i \to y &= \bigvee_{j=1}^{n-1} (e_i \to c_j e_j) = \bigvee_{j=0}^{n-1} (e_i \to c_j)(e_i \to e_j) \\ &= \bigvee_{j=1}^{i-1} c_j e_j \vee \bigvee_{j=i}^{n-1} c_j = y \vee c_i \quad \text{for} \quad 1 \leqq i \leqq n-1 \,. \end{aligned}$$

Therefore,

$$\begin{aligned} x \to y &= \bigwedge_{i=1}^{n-1} (b_i e_i \to y) = \bigwedge_{i=1}^{n-1} (\bar{b}_i \vee c_i \vee y) = y \vee \bigwedge_{i-1}^{n-1} (\bar{b}_i \vee c_i) \\ &= y \vee \bigvee_{i=0}^{n-1} b_i c_i \,, \end{aligned}$$

where the last equality is easily proved by induction.

DEFINITION 3.5. A *Stone lattice* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *of order* n is an L-algebra A in which there exists a chain $0 = e_0 < e_1 < \cdots < e_{n-1} = 1$ such that e_{i+1} is the smallest dense element of $[e_i, 1]$. If B_i is the center of $[e_i, 1]$, let $h_i: B_i \to B_{i+1}$ be the Boolean homomorphism defined by $h_i(x) = x \vee e_{i+1}$, with $B_0 = B$. These definitions are in [11].

THEOREM 3.6. *$\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice of order n, if and only if $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a Stone lattice of order n such that h_i is onto B_{i+1} for each $i \geqq 0$.*

Proof. If $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice of order n, then it is a Stone lattice of order n by 3.4, and h_i is onto by 2.5. Conversely, suppose $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a Stone lattice of order n and h_i is onto B_{i+1} for each i. Then $B_i = \{b \vee e_i : b \in B\}$ by 2.5. It was proved in [11, 3.4], that if $x \in A$, then $x = \bigwedge_{i=0}^{n-2} x_i$, where $x_i \in B_i$. Therefore $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice.

THEOREM 3.7. *If A is a Heyting algebra with center B, $0 = e_0 \leqq \cdots \leqq e_{n-1} = 1$, e_{i+1} is the smallest dense element of $[e_i, 1]$, and if whenever $i < j$, the center of $[e_i, e_j]$ is $\{e_i \vee be_j : b \in B\}$, then $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice.*

Proof. The point of this theorem is that the condition that A is an L-algebra is replaced by the condition that A is a Heyting algebra such that the center of $[e_i, e_j]$ is $\{e_i \vee be_j : b \in B\}$, for all $i < j$. We omit details of proof since this theorem is not used in what follows.

4. *P_2-lattices.*

DEFINITION 4.1. $x = \bigvee_{i=1}^{n-1} b_i e_i$ is called the *highest* monotone representation (hi. mono. rep.) of x if for every mon. rep. $\bigvee_{i=1}^{n-1} c_i e_i$ of x, the relation $b_i \geqq c_i$ holds for all i. The *lowest* monotonic representation (lo. mon. rep.) is defined in a similar manner.

THEOREM 4.2. *Let $\langle A; e_0, \cdots, e_{n-1}\rangle$ be a P_0-lattice. Then the following are equivalent:*

(i) *$e_i \Rightarrow e_j$ exists for all i, j.*
(ii) *$e_i \to e_j$ and $!e_i$ exist for all i, j.*
(iii) *every $x \in A$ has a hi. mon. rep.*
(iv) *every $x \in A$ has a lo. mon. rep.*

(v) *A is a B-algebra.*

(vi) *A is a P-algebra.*

The hi. mon. rep. of x *is* $\bigvee_i (e_i \Rightarrow x)e_i$, *and the lo. mon. rep. of* x *is* $\bigvee_i \overline{(x \Rightarrow e_{i-1})}e_i$.

Proof. The equivalence of (i), (v), and (vi) is proved exactly as in the proof of 3.1. By [7], A is a P-algebra if and only if A is a pseudo-supplemented L-algebra in which $!(x \vee y) = !x \vee !y$ for all x, y. Therefore, by 3.1 and 2.10, (ii) is equivalent to (vi).

To prove (iii) implies (i), let $\bigvee_j b_j e_j$ be the hi. mon. rep. of e_i. Then $b_{i+1}e_{i+1} \leqq e_i$. Let $b \in B$, $be_{i+1} \leqq e_i$. Thus $e_1 \vee \cdots \vee e_i \vee be_{i+1}$ is a mon. rep. of e_i. Therefore $b_{i+1} \geqq b$, which proves $b_{i+1} = e_{i+1} \Rightarrow e_i$. Hence if $i > j$, $e_i \Rightarrow e_j = \bigwedge_{k=i}^{j-1} (e_k \Rightarrow e_{k+1})$, and for $i \leqq j$, $e_i \Rightarrow e_j = 1$.

To prove (vi) implies (iii), let $x = \bigvee_i b_i e_i$ be any mon. rep. Then $e_i \Rightarrow x \geqq e_i \Rightarrow b_i e_i \geqq b_i$. Also $e_i(e_i \Rightarrow x) \leqq x$. Therefore,

$$x \geqq \bigvee_i e_i(e_i \Rightarrow x) \geqq \bigvee_i b_i e_i = x .$$

Thus $\bigvee_i e_i(e_i \Rightarrow x)$ is the hi. mon. rep. of x.

The equivalence of (iv) and (vi) is a consequence of the equivalence of (iii) and (vi), since the dual of a P-algebra is a P-algebra. The formula for the lo. mon. rep. is obtained by duality, for if $x = \bigvee_i b_i e_i$ is a mon. rep., then $x = \bigwedge_i (b_i \vee e_{i-1})$.

DEFINITION 4.3. A P_2-lattice is a P_1-lattice $\langle A; e_0, \cdots, e_{n-1}\rangle$ such that $!e_i$ exists for all i.

Using 2.2, it is easy to construct a P_1-lattice which is not a P_2-lattice.

THEOREM 4.4. *If* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_0*-lattice of order* n *and* A *is a B-algebra, then there exists a unique chain* $f_0, \cdots, f_{n-1}$ *such that* $\langle A; f_0, \cdots, f_{n-1}\rangle$ *is a* P_2*-lattice.*

Proof. This follows from 3.3, 3.4, and 4.2.

THEOREM 4.5. *Let* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *be a* P_2*-lattice. Then*

(i) *Every* $x \in A$ *has a unique mon. rep.* $\bigvee_i D_i(x)e_i$ *such that* $D_{n-1}(x) = !x$. *This representation is also the hi. mon. rep. of* x.

(ii) *Every* $x \in A$ *has a unique disj. rep.* $\bigvee_i C_i(x)e_i$ *such that* $C_{n-1}(x) = !x$.

(iii) $D_i(x) = e_i \Rightarrow x$, $C_i(x) = D_i(x) - D_{i+1}(x)$ *and for* $i < n-1$, $C_i(x) = (x \Rightarrow e_i)(e_i \Rightarrow x) - !(xe_i)$.

(iv) $D_i(x \vee y) = D_i(x) \vee D_i(y)$, $D_i(xy) = D_i(x)D_i(y)$.

(v) $x \Rightarrow y = \bigvee_{i=0}^{n-1} C_i(x)D_i(y)$, *where* $D_0(y) = 1$ *and* $C_0(x) = 1 - D_1(x)$.

Proof. (i) Let $x = \bigvee_j b_j e_j$ be a mon. rep. such that $b_{n-1} = !x$. If $i > j$, then $e_i \Rightarrow e_j = !(e_i \rightarrow e_j) = !e_j$. Therefore,

$$\begin{aligned} e_i \Rightarrow x &= \bigvee_j (e_i \Rightarrow b_j e_j) = \bigvee_j (e_i \Rightarrow b_j)(e_i \Rightarrow e_j) \\ &= \bigvee_{j<i} b_j !e_j \vee \bigvee_{j \geqq i} b_j = b_i \,, \end{aligned}$$

since $\bigvee_{j=1}^{n-1} b_j !e_j = !x \leqq b_i$. We set $D_i(x) = e_i \Rightarrow x$ for $0 \leqq i \leqq n-1$. By 4.2, the hi. mon. rep. of x is $\bigvee_i D_i(x) e_i$, and $D_{n-1}(x) = 1 \Rightarrow x = !x$.

(ii) Follows from (i), with $C_i(x) = D_i(x) - D_{i+1}(x)$.

(iii) For $0 \leqq i < n-1$,

$$\begin{aligned} x \Rightarrow e_i &= \bigwedge_j (D_j(x) e_j \Rightarrow e_i) = \bigwedge_j (\overline{D_j(x)} \vee (e_j \Rightarrow e_i)) \\ &= \bigwedge_{j>i} (\overline{D_j(x)} \vee !e_i) = \overline{D_{i+1}(x)} \vee !e_i \,. \end{aligned}$$

Therefore $(x \Rightarrow e_i)(e_i \Rightarrow x) = C_i(x) \vee D_i(x)!e_i$. Since $D_i(x)!e_i \leqq e_i(e_i \Rightarrow x) \leqq x$, we have $D_i(x)!e_i \leqq !x = D_{n-1}(x)$. Hence $D_i(x)!e_i = !x!e_i = !(xe_i)$. Also $C_i(x)!x = C_i(x)C_{n-1}(x) = 0$. Therefore,

$$C_i(x) = (x \Rightarrow e_i)(e_i \Rightarrow x) - !(xe_i) \,.$$

(iv) follows immediately from $D_i(x) = e_i \Rightarrow x$.

(v) By 3.4, $x \rightarrow y = y \vee \bigvee_{i=0}^{n-1} C_i(x) D_i(y)$. Therefore,

$$x \Rightarrow y = !y \vee \bigvee_{i=0}^{n-1} C_i(x) D_i(y) = \bigvee_{i=0}^{n-1} C_i(x) D_i(y) \,,$$

since $\bigvee_{i=0}^{n-1} C_i(x) D_i(y) \geqq D_{n-1}(y) \bigvee_{i=0}^{n-1} C_i(x) = !y$.

THEOREM 4.6. *The following are equivalent*:

(i) $\langle A; e_0, \cdots, e_{n-1} \rangle$ *is a P_2-lattice of order n.*

(ii) $\langle A; e_0, \cdots, e_{n-1} \rangle$ *is a Stone lattice of order n, the homomorphisms h_i of* 3.5 *are onto, and the kernel of h_i is a principal ideal for each i.*

(iii) $\langle A; e_0, \cdots, e_{n-1} \rangle$ *is a Stone lattice of order n and A^d is a Stone lattice.*

Proof. The equivalence of (i) and (ii) follows from 3.6 and 2.9, using the fact that the kernel of h_i is a principal ideal if and only if $!e_{i+1}$ exists. The equivalence of (ii) and (iii) was proved in [11].

The following is the dual of the definition given in [5].

DEFINITION 4.7. A Post algebra is a P_2-algebra $\langle A; e_0, \cdots, e_{n-1} \rangle$ such that $!e_{n-2} = 0$; that is, e_{n-2} is dense in A^d. Note that A has order n, unless $A = \{0\}$.

THEOREM 4.8. *If* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_0*-lattice then the following are equivalent:*

(i) A *is a Post algebra.*

(ii) *every element* $x \in A$ *has a unique mon. rep.*

(iii) $e_i \Rightarrow e_{i-1} = 0$ *for all* $i > 0$.

Proof. This was proved in [13].

LEMMA 4.9. *If* $\langle A_j; e_{j0}, \cdots, e_{j(n_j-1)}\rangle$ *is a* P_r*-lattice for* $j \in J$, *where* $r = 0, 1,$ *or* 2, $A = \prod_{j\in J} A_j$, $n = \max\{n_j\colon j \in J\} < \infty$, *and* e_{jk} *is defined to be* $e_{j(n_j-1)}$ *for* $k > n_j$, *then* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_r*-lattice, where* $e_i = \langle e_{ji}\colon j \in J\rangle$.

Proof. This is obvious.

LEMMA 4.10. *If* $\langle A; e_0, \cdots, e_{n-1})$ *is a* P_0*-lattice,* B *is a distributive lattice and* $f\colon A \to B$ *is a lattice homomorphism onto, then* $\langle B; f(e_0), \cdots, f(e_{n-1})\rangle$ *is a* P_0*-lattice. If* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_1*-lattice and* $f\colon A \to B$ *is a Heyting homomorphism onto, then* $\langle B; f(e_0), \cdots, f(e_{n-1})\rangle$ *is a* P_1*-lattice.*

Proof. This is easy to verify.

THEOREM 4.11. *Let* A *be a finite distributive lattice then the following are equivalent:*

(i) A *is a* P_0*-lattice.*

(ii) A *is a* P*-algebra.*

(iii) A *is a direct product of chains.*

(iv) A *has a chain base* $e_0, \cdots, e_{n-1}$ *such that* $\langle A; e_0, \cdots, e_{n-1}\rangle$ *is a* P_2*-lattice.*

Proof. (i) implies (ii): Since A is finite, A is a pseudo-supplemented Heyting algebra. By 4.2, A is a P-algebra.

(ii) implies (iii) was proved in [7].

(iii) implies (iv): If A is a finite chain $0 = a_0 < \cdots < a_{n-1} = 1$, then $\langle A; a_0, \cdots, a_{n-1}\rangle$ is a P_2-lattice. Therefore (iv) follows by 4.9.

(iv) implies (i) is obvious.

A finite chain with n elements has exactly one chain base with n terms. If $\langle A; e_0, \cdots, e_{n-1}\rangle$ and $\langle B; f_0, \cdots, f_{m-1}\rangle$ are P_0-lattices of orders n and m respectively, where $n < m$, then $A \times B$ has more than one chain base. In addition to the chain base described in 4.9, there is also the chain base $(e_0, f_0), \cdots, (e_0, f_{m-n}), (e_1, f_{m-n+1}), (e_2, f_{m-n+2}), \cdots, (e_{n-1}, f_{m-1})$. These remarks lead to the next theorem.

THEOREM 4.12. *A distributive lattice A is a Post algebra of order n if and only if A has a unique n-term chain base.*

Proof. Let A be a Post algebra of order n, and let $e_0, \cdots, e_{n-1}$ be an n-term chain base. A is a subdirect power of an n element chain C. If $f_j = A \to C$ is the jth projection, then by 4.10, $f_j(e_0), \cdots, f_j(e_{n-1})$ is a chain base of C. This determines $f_j(e_i)$ uniquely for all i, j. Therefore $e_0, \cdots, e_{n-1}$ is unique.

Conversely, suppose A has a unique n-term chain base $e_0, \cdots, e_{n-1}$. We prove A is a Post algebra of order n by induction. This is obvious for $n = 1, 2$. Suppose $n > 2$ and the statement holds for $n - 1$. By 2.5, $[e_1, 1]$ has a unique chain base with $n - 1$ terms. Therefore, $[e_1, 1]$ is a Post algebra of order $n - 1$. This implies $e_{i+1} \Rightarrow e_i = 0$ in $[e_1, 1]$ for $i \geqq 1$. This implies $e_{i+1} \Rightarrow e_i = 0$ in A since the center of $[e_1, 1]$ is $\{b \vee e_1 : b \in B\}$, where B is the center of A. By 4.8, we need only show $e_1 \Rightarrow 0 = 0$. If not, there exists $b \in B$ with $be_1 = 0$, $b \neq 0$. Let $B_1 = \{0, b, \bar{b}, 1\}$, and let A_1 be the sublattice of A generated by $B_1 \cup \{e_0, \cdots, e_{n-1}\}$. By 2.2, A_1 has center B_1 and so every chain base of A_1 is a chain base of A. Thus A_1 is a finite lattice with a unique n-term chain base. By 4.11, A_1 is a direct product of finite chains. If all the chains have the same cardinal, then A_1 is a Post algebra with unique n-term chain base $e_0, \cdots, e_{n-1}$, and by 4.8, we have $e_1 \Rightarrow 0 = 0$, which contradicts $be_1 = 0$, $b \neq 0$. If two of the chains have different cardinal, then A_1 has more than one n-term chain base. This contradiction proves $e_1 \Rightarrow 0 = 0$.

THEOREM 4.13. *If $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice with center B, then there exists a P_2-lattice $\langle A'; e_0, \cdots, e_{n-1}\rangle$ with center B' such that B is a Boolean subalgebra of B' and A is the sublattice of A' generated by $B \cup \{e_0, \cdots, e_{n-1}\}$.*

Proof. By 3.1, A is an L-algebra. By [9], we may assume A is a Heyting subalgebra of a direct product of chains C_j, $j \in J$ and the projections $f_j: A \to C_j$ are onto. Then by 4.10, $\langle C_j; f_j(e_0), \cdots, f_j(e_{n-1})\rangle$ is a P_1-lattice. Therefore, C_j has at most n elements and $\langle C_j; f_j(e_0), \cdots, f_j(e_{n-1})\rangle$ is a P_2-lattice. Let $A' = \prod_{j \in J} C_j$. By 4.9, $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_2-lattice. Since A is a sublattice of A' containing 0, 1, the center of A is a subalgebra of the center of A'.

THEOREM 4.14. *Let $\langle A; e_0, \cdots, e_{n-1}\rangle$ be a P_2-lattice of order n with center B. Then A is order isomorphic with a direct product of Post algebras of maximum order n.*

Proof. Let $u_k = !e_k - !e_{k-1}$, $1 \leqq k \leqq n - 1$. Then $u_j u_k = 0$ for

$j \neq k$, and $u_1 \vee \cdots \vee u_{n-1} = 1$. Let $P_k = [0, u_k]$. Clearly the center of P_k is $B \cap P_k$. Let $e_{ki} = e_i u_k$, $0 \leqq i \leqq k$. If $x = \bigvee_{i=1}^{n-1} b_i e_i$ is a mon. rep. of any $x \in P_k$, then

$$x - xu_k = \bigvee_{i=1}^{n-1} b_i e_i u_k = \bigvee_{i=1}^{k-1} b_i e_i u_k \vee \bigvee_{i=k}^{n-1} b_i u_k = \bigvee_{i=1}^{k} (b_i u_k) e_{ki} .$$

Therefore $\langle P_k; e_{k0}, \cdots, e_{kk} \rangle$ is a P_0-lattice. If $b \in P_k \cap B$, $be_{ki} \leqq e_{k(i-1)}$, $0 < i \leqq k$, then $be_i \leqq e_{i-1}$. Therefore $b \leqq e_i \Rightarrow e_{i-1} = !e_{i-1}$. This implies $b = 0$, since $b \leqq u_k$. Thus by 4.8, P_k is a Post algebra of order $k + 1$, or else $P_k = \{0\}$. Define $f: A \rightarrow \prod_{k=1}^{n-1} P_k$ by $f(x) = (xu_1, \cdots, xu_{n-1})$. f is onto since if $z_i \in P_i$, then $f(z_1 \vee \cdots \vee z_{n-1}) = (z_1, \cdots, z_{n-1})$. If $x \leqq y$ then $f(x) \leqq f(y)$, and $f(x) \leqq f(y)$ implies

$$x = \bigvee_{i=1}^{n-1} xu_i \leqq \bigvee_{i=1}^{n-1} yu_i = y .$$

Therefore f is an order isomorphism. Finally, P_{n-1} has order n since $u_{n-1} \neq 0$.

Theorem 4.14 may be used to apply known results on Post algebras to P_2-lattices. For example, since every Post algebra is isomorphic with the set of all continuous functions on a Boolean space to a finite discrete chain, it follows that every P_2-lattice is isomorphic with the set of all such functions which are $\leqq$ some fixed continuous function. In other words, a P_2-lattice is a principal ideal in a Post algebra. It also follows from 4.14 that a P_2-lattice is complete if and only if its center is complete, and that the normal completion of a P_2-lattice A is a P_2-lattice whose center is the normal completion of the center of A. Also every P_2-algebra is isomorphic with its dual. This isomorphism is given explicitly in the following theorem.

THEOREM 4.15. *Let* $\langle A; e_0, \cdots, e_{n-1} \rangle$ *be a* P_2*-lattice. Let* $f_i = \bigvee_{k=1}^{n-1-i} e_k \overline{!e_{k+i-1}}$, $0 \leqq i < n - 1$ *and* $f_{n-1} = 0$. *Then* A *is isomorphic with* A^d *under the normal involution*

$$\beta(x) = \bigvee_{j=1}^{n-1} \overline{D_j(x)} f_{j-1} = \bigwedge_{j=1}^{n-1} (\overline{D_j(x)} \vee f_j) .$$

Proof. We have $f_0 \geqq \bigvee_{k=1}^{n-1} (!e_k - !e_{k-1}) = 1$. For $0 < i < n - 1$, $!f_i = 0$, so that by 4.5(i), $D_k(f_i) = \overline{!e_{k+i-1}}$ for $1 \leqq k \leqq n - 1 - i$, and $D_k(f_i) = 0$ for $k \geqq n - i$.

If $1 \leqq i \leqq n - 2$,

$$\begin{aligned}
\beta(f_i) &= \bigvee_{j=1}^{n-1} f_{j-1} \overline{D_j(f_i)} = \bigvee_{j=1}^{n-1-i} f_{j-1} !e_{j+i-1} \vee \bigvee_{j=n-i}^{n-1} f_{j-1} \\
&= \bigvee_{j=1}^{n-1} f_{j-1} !e_{j+i-1} = \bigvee_{j=1}^{n-i} !e_{j+i-1} \bigvee_{k=1}^{n-j} e_k \overline{!e_{j+k-2}} \\
&= \bigvee_{k=1}^{i} e_k \bigvee_{j=1}^{n-i} (!e_{j+i-1} - !e_{j+k-2}) \vee \bigvee_{k=i+1}^{n-1} e_k \bigvee_{j=1}^{n-k} (!e_{j+i-1} - !e_{j+k-2}) .
\end{aligned}$$

But $\bigvee_{j=1}^{n-k}(!e_{j+i-1} - !e_{j+k-2}) = 0$ if $k > i$, and by 2.7, if $k \leqq i$,

$$\bigvee_{j=1}^{n-i}(!e_{j+i-1} - !e_{j+k-2}) = \overline{!e_{k-1}} \bigwedge_{j=2}^{n-i}(!e_{j+i-2} \vee \overline{!e_{j+k-2}}) = \overline{!e_{k-1}} .$$

Therefore, $\beta(f_i) = \bigvee_{k=1}^{i}(e_k - !e_{k-1}) = e_i$.

Now $x \leqq y$ implies $\beta(x) \leqq \beta(y)$ and

$$\beta(\beta(x)) = \bigvee_{j=1}^{n-1} f_{j-1} \bigvee_{i=1}^{n-1} D_i(x)\overline{D_j(f_i)} = \bigvee_{i=1}^{n-1} \beta(f_i)D_i(x) = \bigvee_{i=1}^{n=1} D_i(x)e_i = x .$$

This implies that $\beta\colon A \to A^d$ is an isomorphism. The proof that β is a normal involution—that is, that $x\beta(x) \leqq y \vee \beta(y)$ for all $x, y \in A$—is omitted since this fact will not be used here [10].

5. **Axioms and P_2-functions.** P_2-algebras $\langle A; e_0, \cdots, e_{n-1}\rangle$ of order $\leqq n$ may be regarded as algebras $\langle A; \vee, \wedge, C_0, \cdots, C_{n-1}, e_0, \cdots, e_{n-1}\rangle$ with two binary operations, n binary operations, and n distinguished constants. This class of algebras can be characterized by the following set of equational axioms, in which $x \leqq y$ is used as an abbreviation for $x \wedge y = x$.

H1. Identities characterizing $\langle A; \vee, \wedge\rangle$ as a distributive lattice [8, pp. 5, 35].

H2. (a) $e_0 \leqq x$

(b) $e_i \leqq e_j$ for $0 \leqq i \leqq j \leqq n-1$

(c) $x \leqq e_{n-1}$

H3. (a) $C_i(x) \wedge C_j(x) = e_0$ for $i \neq j$

(b) $C_0(x) \vee C_1(x) \vee \cdots \vee C_{n-1}(x) = e_{n-1}$

H4. (a) $C_i(x \wedge y) = (C_i(x) \wedge \bigvee_{k=i}^{n-1} C_k(y)) \vee (C_i(y) \wedge \bigvee_{k=i}^{n-1} C_k(x))$

(b) $C_{n-1}(x \vee y) = C_{n-1}(x) \vee C_{n-1}(y)$

H5. (a) $C_i(e_j) = e_0$ for $j \neq i$ and $i < n-1$

(b) $C_{n-1}(e_0) = e_0$

H6. $x = (C_1(x) \wedge e_1) \vee \cdots \vee (C_{n-1}(x) \wedge e_{n-1})$.

Note that in every P_2-lattice H4 holds by 4.5(iv) and H5 holds by 4.5(ii). Conversely, if A satisfies the axioms then one proves $C_{n-1}(1) = 1$, $C_0(0) = 0$, $C_0(x) = \neg x$ and $C_{n-1}(x) = !x$. Then using H4 and H5, it can be proved that $xe_i = e_{i-1}$ implies $x = e_{i-1}$. This shows that $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_2-lattice.

The class of Post algebras of order n (together with the trivial lattice $\{0\}$) can be characterized by adding the axiom $C_{n-1}(e_{n-2}) = 0$ (see also [6]).

We may also characterize P_2-lattices equationally as the class of all algebras $\langle A; \vee, \wedge, \Rightarrow, e_0, \cdots, e_{n-1}\rangle$ with 3 binary operations and n constants which satisfy the following axioms.

K1. Identities characterizing $\langle A; \vee, \wedge, \Rightarrow, e_0, e_{n-1}\rangle$ as a P-algebra (see [7]).

K2. $e_i \leqq e_j$ for $i \leqq j$

K3. $e_{i+1}(e_i \Rightarrow e_j) \leqq e_j$ for $j < i < n-1$

K4. $x = \bigvee_{j=1}^{n-1} (e_i \wedge (e_i \Rightarrow x))$.

Indeed, if we set $D_i(x) = e_i \Rightarrow x$ for $0 \leqq i \leqq n-1$ and let $C_i(x) = D_i(x) - D_{i+1}(x)$ for $i < n$, then H1–3, H5(b), and H6 are obvious. By properties of P-algebras, $D_i(x \vee y) = D_i(x) \vee D_i(y)$ and $D_i(x \wedge y) = D_i(x) \wedge D_i(y)$. This proves H4. H5(a) is equivalent to $e_i \Rightarrow e_j = e_{i+1} \Rightarrow e_j$ for $j \neq i$, $i < n-1$. This is obvious for $i < j$, and follows from K3 for $i > j$.

P_2-lattices may also be characterized equationally as algebras $\langle A; \vee, \wedge, \rightarrow, !, e_0, \cdots, e_{n-1}\rangle$, since $x \Rightarrow y = !(x \rightarrow y)$, $x \rightarrow y = y \vee (x \Rightarrow y)$ and $!x = 1 \Rightarrow x$.

A P_2-function of order n in m variables is a function built from the identity functions $I_j(x_1, \cdots, x_m) = x_j$, $1 \leqq j \leqq m$, and the operations in any of the fundamental sets of operations described above. A normal form for such functions is given in the next theorem.

THEOREM 5.1. *If h is a P_2-function of order n in m variables, then*

$$h(x_1, \cdots, x_m) = \bigvee_{0 \leqq i_k \leqq n-1} h(e_{i_1}, \cdots, e_{i_m}) C_{i_1}(x_1) \cdots C_{i_m}(x_m) .$$

Proof. The n^m terms $C_{i_1}(x_1) \cdots C_{i_m}(x_m)$ are pairwise disjoint and have join 1, by axiom H3. By H6, the statement holds when h is one of the identity functions. If the statement holds for h_1 and h_2, then it holds for $h_1 \vee h_2$ and $h_1 \wedge h_2$. If it holds for h, then it holds for $D_i(h)$ by 4.5(iv). From this it follows that the statement holds for $C_j(h)$.

The normal form in 5.1 was proved for Post algebras in [5], and gives a truth table approach to Post functions. However, in a P_2-lattice, $h(e_{i_1}, \cdots, e_{i_m})$ is not necessarily in $\{e_0, \cdots, e_{n-1}\}$, as is the case for Post algebras.

6. Applications. P_2-lattices are of interest in computer science. They can be applied to the theory of machines with m_i-stable devices, $2 \leqq m_i \leqq n$, and to the analysis of machines with 2-stable devices Q_i (flip-flops) whose outputs are discretized as signals in transition $0 = e_0 < e_1 < \cdots < e_{n-1} = 1$. The case $n = 3$ is of special interest and is studied in [2] and [3]. P_2-lattices provide the complete multiple-valued logics for these applications.

P_2-lattices which admit operations of ring addition and multiplication are of interest in information processing. It is known that if R is a ring with unit element which satisfies the identities $x^p = x$ and $px = 0$, where p is a prime (so-called p-rings [12]), then lattice operations can be defined as polynomials in such a way that R becomes a Post algebra of order p. Conversely in any Post algebra of order p, ring operations can be defined in terms of the Post operations so that we obtain a p-ring. Therefore, direct products of finitely many p-rings are P_2-lattices. Such direct products can be characterized equationally. Indeed one can show that a ring R with unit element is a direct product of rings $R_1, \cdots, R_t$, where R_i is a p_i-ring and $p_i \neq p_j$ for $i \neq j$, if and only if R satisfies the following set of identities:

(1) $x^m = x$, where $m = 1 + \text{l.c.m.}\ (p_1 - 1, \cdots, p_t - 1)$.

(2) $p_1 \cdots p_t x = 0$.

(3) $(\prod_{j \neq i} p_j)(x^{p_i} - x) = 0,\ 1 \leqq i \leqq t$.

7. Prime ideals.

DEFINITION 7.1. Let $\mathscr{P}(A)$ be the set of prime ideals of A. Let $\langle A; e_0, \cdots, e_{n-1}\rangle$ be a P_0-lattice with center B. If $Q \in \mathscr{P}(B)$ and $1 \leqq k \leqq n-1$, let $P_k(Q) = \{x \in A\colon x$ has a mon. rep. $\bigvee_i b_i e_i$ such that $b_k \in Q\}$. It was proved in [13, Th. 1.5] that either $P_k(Q) \in \mathscr{P}(A)$ or $P_k(Q) = A$ (the latter possibility was not mentioned). If $P_k(Q) \neq A$, then $P_k(Q) \cap B = Q$ since if $b \in Q$ then $b = \bigvee_i b e_i \in P_k(Q)$ and prime ideals in B are maximal ideals. If $P \in \mathscr{P}(A)$, then P is said to be of *type* k if k is the smallest integer such that $e_k \notin P$. Since $e_{k-1} = e_1 \vee \cdots \vee e_{k-1} \in P_k(Q)$, $P_k(Q)$ is of type $\geqq k$.

LEMMA 7.2. *If P is a prime ideal of type k in A and $Q = P \cap B$, then*

$$P = P_k(Q) = \{x\colon \textit{for every mon. rep. } \textstyle\bigvee_i b_i e_i \textit{ of } x,\ b_k \in Q\}\ .$$

Proof. If $x \in A$ has a mon. rep. $\bigvee_i b_i e_i$ with $b_k \in Q$ then $x \leqq e_{k-1} \vee b_k \in P$. If $x \in P$ and $\bigvee_i b_i e_i$ is any mon. rep. of x, then $b_k e_k \in P$ and $e_k \notin P$, so that $b_k \in Q$.

THEOREM 7.3. *The prime ideals of a P_0-lattice $\langle A; e_0, \cdots, e_{n-1}\rangle$ lie in disjoint maximal chains with at most $n-1$ members.*

Proof. By 7.1, each prime ideal of A is of the form $P_k(Q)$. If $P_k(Q_1) \subseteqq P_j(Q_2)$, then $Q_1 = P_k(Q_1) \cap B \subseteqq P_j(Q_2) \cap B = Q_2$, so that $Q_1 = Q_2$. It is obvious that $P_k(Q) \subseteqq P_{k+1}(Q)$. This proves the theorem.

LEMMA 7.4. *If $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_0-lattice with center B, and $Q \in \mathscr{P}(B)$, then $P_k(Q) = \{x\colon$ for some $b \in Q,\ x \leqq e_{k-1} \vee b\}$. Also*

$P_{k+1}(Q) = P_k(Q)$ *if and only if* $e_k \in P_k(Q)$.

Proof. If $x \in P_k(Q)$, there exists a mon. rep. $\bigvee_i b_i e_i$ of x such that $b_k \in Q$. Also $x \leqq e_{k-1} \vee b_k$. If $x \leqq e_{k-1} \vee b$ and $b \in Q$ then $x \in P_k(Q)$ since $e_{k-1} \vee b = \bigvee_{j=1}^{k-1} e_j \vee \bigvee_{j=k}^{n-1} be_j$. Suppose $e_k \in P_k(Q)$. If $x \in P_{k+1}(Q)$, then $x \leqq e_k \vee b$ for some $b \in Q$, hence $x \in P_k(Q)$. Thus $P_{k+1}(Q) = P_k(Q)$. Conversely if $P_{k+1}(Q) = P_k(Q)$, then $e_k \in P_k(Q)$ since $e_k \in P_{k+1}(Q)$.

THEOREM 7.5. *Let $\langle A; e_0, \cdots, e_{n-1}\rangle$ be a P_0-lattice with center B, and let I_k be the ideal $\{b \in B: be_k \leqq e_{k-1}\}$ in B, $1 \leqq k \leqq n-1$. Then the following are equivalent:*

(i) *Every chain in $\mathscr{P}(A)$ has fewer than $n-1$ elements.*

(ii) *For every $Q \in \mathscr{P}(B)$, there exists $b \in Q$ and an integer $k \geqq 1$ such that $e_k \leqq e_{k-1} \vee b$.*

(iii) $I_1 \vee \cdots \vee I_{n-1} = B$.

(iv) *A has a chain base with fewer than n elements.*

Proof. (i) implies (ii): If $Q \in \mathscr{P}(B)$, then either $P_{n-1}(Q) = A$ or $P_k(Q) = P_{k+1}(Q)$ for some $k < n-1$. Hence by 7.4, $e_k \in P_k(Q)$ for some k, $1 \leqq k \leqq n-1$, and therefore there exists $b \in Q$ such that $e_k \leqq e_{k-1} \vee b$.

(ii) implies (iii): If $I_1 \vee \cdots \vee I_{n-1} \neq B$, there exists $Q \in \mathscr{P}(B)$ such that $Q \supseteqq I_1 \vee \cdots \vee I_{n-1}$. There exists $b \in Q$ and k such that $e_k \leqq e_{k-1} \vee b$. But then $\bar{b} \in I_k \subseteqq Q$, which is impossible.

(iii) implies (iv): There exist elements $b_k \in I_k$ such that $1 = b_1 \vee \cdots \vee b_{n-1}$. By replacing the b_k by smaller elements, we may assume the b_k are pairwise disjoint. Let $f_0 = 0$ and

$$f_k = e_k \vee e_{k+1} \bigvee_{j=1}^{k} b_j\,, \qquad 1 \leqq k \leqq n-2\,.$$

Then $f_n \leqq f_{n+1}$ and $f_{n-2} = 1$, since $b_1 \vee \cdots \vee b_{n-2} = \bar{b}_{n-1}$ and $b_{n-1} \leqq e_{n-2}$. Now $f_{k-1} \vee f_k \bigvee_{j=k+1}^{n-1} b_j = e_{k-1} \vee \bigvee_{j \neq k} b_j$. Therefore,

$$e_k = f_{k-1} \vee f_k \bigvee_{j=k+1}^{n-1} b_j\,,$$

and so $f_0, \cdots, f_{n-2}$ is a chain base of A.

(iv) implies (i) by 7.3.

THEOREM 7.6. *Let A be a P_0-lattice. Then A is of order n if and only if the maximum number of elements in a chain in $\mathscr{P}(A)$ in $n-1$.*

Proof. This follows from 7.3 and the equivalence of 7.5(i) and 7.5(iv).

DEFINITION 7.7. Let $\mathscr{P}_0(A) = \phi$, and let $\mathscr{P}_{k+1}(A)$ be the set of minimal elements of $\mathscr{P}(A) - \mathscr{P}_k(A)$.

THEOREM 7.8. *Let $\langle A; e_0, \cdots, e_{n-1}\rangle$ be a P_1-lattice with center B. Then for $0 \leqq i \leqq n-2$,*

$$e_i \in \bigcap \mathscr{P}_{i+1}(A) - \bigcup_{j \leqq i} \bigcup \mathscr{P}_j(A) .$$

Proof. By 7.4, $P_1(Q) \neq A$ for all $Q \in \mathscr{P}(B)$. Hence $\mathscr{P}_1(A) = \{P_1(Q): Q \in \mathscr{P}(B)\}$. If $1 \leqq i \leqq n-2$, then $e_i \in P_i(Q)$ if and only if $e_i \leqq e_{i-1} \vee b$ for some $b \in Q$. This in turn is equivalent to $\bar{b} \leqq e_i \rightarrow e_{i-1}$ which is equivalent to $\bar{b} \leqq e_{i-1}$, or $1 \leqq e_{i-1} \vee b$. By 7.4, this is equivalent to $P_i(Q) = A$. Also, $P_i(Q) = P_{i+1}(Q)$ if and only if $e_i \in P_i(Q)$. Therefore $\mathscr{P}_i(A) = \{P_i(Q): P_i(Q) \neq A\}$, and $e_i \notin P_i(Q)$ for all $P_i(Q) \in \mathscr{P}_i(A)$. Since $e_i \in P_{i+1}(Q)$ for all $Q \in \mathscr{P}(B)$, the proof is complete.

LEMMA 7.9. *Let A be a bounded distributive lattice. Suppose $\mathscr{P}(A)$ is a union of disjoint maximal chains and there exists an element $e \in \bigcap (\mathscr{P}(A) - \mathscr{P}_1(A)) - \bigcup \mathscr{P}_1(A)$. Let $A_1 = [e, 1]$. Then $\mathscr{P}_i(A_1) = \{P \cap A_1: P \in \mathscr{P}_{i+1}(A)\}$ for each $i \geqq 1$.*

Proof. If $P \in \mathscr{P}(A) - \mathscr{P}_1(A)$, let $\varphi(P) = P \cap A_1$. Then $\varphi(P) \in \mathscr{P}(A_1)$. If $Q \in \mathscr{P}(A_1)$, let $\psi(Q) = \{x \in A: x \geqq$ an element of $Q\}$. Then $\psi(Q) \in \mathscr{P}(A) - \mathscr{P}_1(A)$ and $\psi\varphi(P) = P$. Thus $\varphi: \mathscr{P}(A) - \mathscr{P}_1(A) \rightarrow \mathscr{P}(A_1)$ is an order isomorphism.

LEMMA 7.10. *Under the hypotheses of 7.9, let B and B_1 be the centers of A and A_1 respectively. Then $B_1 = \{b \vee e: b \in B\}$. If $x \in A$, then there exists $b \in B$ such that $x = b(e \vee x)$.*

Proof. Let $\{D_i: i \in S\}$ be the set of maximal chains in $\mathscr{P}(A)$. The intersection and union of any nonempty subset of D_i is in D_i. Let P_i and Q_i be respectively the smallest and largest member of D_i. Let $V = \{i: P_i \neq Q_i\}$. For $i \in V$, let $R_i = \bigcap \{P \in D_i: e \in P\}$. R_i is the immediate successor of P_i in D_i. We divide the proof of the lemma into several parts.

(a) If $x \in P_i$, there exists y such that $xy = 0$ and $y \notin Q_i$.

Indeed for each j such that $x \notin P_j$, choose $y_j \in P_j - Q_i$. Then every prime ideal in A contains a member of $\{x\} \cup \{y_j: x \notin P_j\}$. Therefore, the filter generated by this set is not proper and so there exists a finite meet y of the y_j such that $xy = 0$. Clearly $y \notin Q_i$.

(b) If $x \notin Q_i$, there exists $y \in P_i$ such that $x \vee y = 1$.

For each j such that $x \in Q_j$ choose $y_j \in P_i - Q_j$. The ideal generated by $\{x\} \vee \{y_j: x \in Q_j\}$ is not proper. Therefore, a finite join

y of the y_j satisfies the requirements.

(c) If $x \notin Q_i$ there exists $y \leqq x$ such that $y \notin Q_i$ and $y \in P_j$ whenever $x \in Q_j$.

By (b) there exists $z \in P_i$ such that $x \vee z = 1$. By (a) there exists $y \notin Q_i$ such that $yz = 0$. If $x \in Q_j$, then $z \notin Q_j$, hence $z \notin P_j$ and so $y \in P_j$. If P is any prime ideal containing x then $P \in D_j$ for some j, and so $x \in Q_j$. This implies $y \in P_j \subseteqq P$. Thus $y \leqq x$.

(d) If $x \notin P_i$, there exists $y \notin Q_i$ such that $ey \leqq x$.

For each j such that $x \in P_j$, choose $y_j \in P_j - Q_i$. If P is a prime ideal containing x but not e, then $P = P_j$ for some j and so $y_j \in P$. This implies that x belongs to the filter generated by $\{e\} \vee \{y_j : x \in P_j\}$. The desired y will be the meet of a finite number of y_j.

(e) If $x \in R_i$, there exists $y \notin Q_i$ such that $xy \leqq e$.

For each j such that $x \notin R_j$ choose $y_j \in P_j - Q_i$. If P is a prime ideal containing e but not x, then $P \supseteqq R_j$ for some j and since $x \notin R_j$, $y_j \in P_j \subseteqq P$. This implies that e belongs to the filter generated by $\{x\} \vee \{y_j : x \notin R_j\}$. The desired y is the meet of a finite number of y_j.

(f) If $x \in B_1$ then for all i, either $x \in R_i$ or $x \notin Q_i$.

Let y be the complement of x in A_1. If $x \in Q_i$ then $y \notin Q_i$ since $x \vee y = 1$. Therefore $y \notin R_i$, hence $x_i \in R_i$ since $xy = e \in R_i$.

(g) If for all i, $x \in P_i$ or $x \notin Q_i$, then $x \in B$.

By (a), for each i such that $x \in P_i$, there exists $y_i \notin Q_i$ such that $xy_i = 0$. No prime ideal contains x and $\{y_i : x \in P_i\}$. There exists a finite join y of the y_i such that $x \vee y = 1$ and clearly $xy = 0$.

(h) If $x \in A$, there exists $y \in B$ such that $x = y(e \vee x)$.

Let $T = \{j : x \in P_j\}$. If $T = S$ then $x = 0$ and $y = 0$. If $T = \phi$ then $x \geqq e$ and $y = 1$ will do. Suppose $T \neq S$, $T \neq \phi$. By (d), for each $i \in S - T$, there exists $y_i \notin Q_i$ such that $ey_i \leqq x$. By (a), for each $j \in T$ there exists $z_j \notin Q_j$ such that $xz_j = 0$. No prime ideal contains $\{y_i : i \in S - T\} \cup \{z_j : j \in T\}$. Therefore, there exist y, z such that $y \vee z = 1$, $ey \leqq x$, and $xz = 0$. This implies $x = xy = xy \vee ey = x(y \vee e)$. If $j \in T$, then $ey \leqq x \in P_j$, $e \notin P_j$ so that $y \in P_j$. If $i \in S - T$. then $z \in P_i$ since $x \notin P_i$ and $xz = 0$. Thus $yz \in P_i$ for all $i \in S$, and so $yz = 0$. Hence $y \in B$.

(i) If $x \in B_1$, there exists $y \in B$ such that $x = y \vee e$.

Let $W = \{j \in V : x \in R_j\}$. If $W = V$, then $x = e$ and $y = 0$. If $W = \phi$ then by (f), $x = 1$ and $y = 1$. Suppose $W \neq V$, $W \neq \phi$. By (c), for each $i \in S - W$ there exists $y_i \leqq x$ such that $y_i \notin Q_i$ and $y_i \in P_j$ for all $j \in W$. By (e), for each $j \in W$, there exists $z_j \notin Q_j$ such that $xz_j \leqq e$. No prime ideal contains $\{y_i : i \in S - W\} \vee \{z_j : j \in W\}$. Therefore, there exist y, z such that $1 = y \vee z$, $xz \leqq e$, $y \leqq x$ and $y \in P_j$ for all $j \in W$. If $i \in V - W$ then $x \notin R_i$ and $e \in R_i$, hence $z \in R_i$ and so $y \notin R_i$. If $i \in S - V$, then $y \in P_i$ or $y \notin Q_i$ since $P_i = Q_i$. Therefore by (g), $y \in B$. Finally $y \vee e \leqq x = xy \vee xz \leqq y \vee e$, and so $x = y \vee e$.

(h) and (i) yield the lemma since it is obvious that $\{b \vee e: b \in B\} \subseteqq B_1$.

THEOREM 7.11. *Let A be a bounded distributive lattice. Suppose $\mathscr{P}(A)$ is a union of disjoint maximal chains with maximum number of elements equal to $n-1$, and for each i, $0 \leqq i \leqq n-2$, there exists an element $e_i \in \bigcap \mathscr{P}_{i+1}(A) - \bigcup_{j \leqq i} \bigcup \mathscr{P}_j(A)$. If we set $e_{n-1} = 1$, then $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice.*

Proof. Clearly $0 = e_0 < e_1 < \cdots < e_{n-2} < 1$. If $n = 2$, then A is a Boolean algebra by Nachbin's theorem [8, p. 76], and the theorem holds. Assume $n > 2$ and the theorem holds for $n-1$. Let $A_1 = [e_1, 1]$. By 7.9, A_1 satisfies the hypothesis for $n-1$. Therefore $\langle A_1; e_1, \cdots, e_{n-1}\rangle$ is a P_1-lattice. Let x be any member of A. By 7.10, $x \vee e_1 = \bigvee_{i=2}^{n-1} (e_1 \vee b_i)e_i$, where $b_i \in B$. Again by 7.10, there exists $b \in B$ such that $x = b(x \vee e_1)$. Therefore $x = be_1 \vee \bigvee_{i=2}^{n-1} bb_ie_i$. Clearly $e_{i+1} \rightarrow e_i = e_i$ in A, for $i \geqq 1$. It remains to show $e_1 \rightarrow 0 = 0$. Suppose $ye_1 = 0$ and $y \neq 0$. There exists a maximal filter F containing y. But $A - F \in \mathscr{P}_1(A)$, and so $e_1 \in F$. Thus $0 \in F$, a contradiction.

THEOREM 7.12. *Let A be a bounded distributive lattice. Then there exists a sequence $e_0, \cdots, e_{n-1}$ such that $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_1-lattice of order n if and only if*

(i) *$\mathscr{P}(A)$ is a union of disjoint maximal chains with maximum number of elements equal to $n-1$, and*

(ii) $\bigcap \mathscr{P}_{i+1}(A) - \bigcup_{j \leqq i} \bigcup \mathscr{P}_j(A) \neq \phi$.

Proof. This follows from 7.6, 7.8, and 7.11.

THEOREM 7.13. *Let A be a bounded distributive lattice. Then there exists a sequence $e_0, \cdots, e_{n-1}$ such that $\langle A; e_0, \cdots, e_{n-1}\rangle$ is a P_2-lattice of order n if and only if conditions* (i) *and* (ii) *of Theorem* 7.12 *hold as well as*

(iii) *There exists an element $c \in A$ such that for all $P \in \mathscr{P}(A)$, $c \in P$ if and only if P is a maximal ideal.*

Proof. By the equivalence of (i) and (iii) in Theorem 4.6, this is a consequence of [11, 4.9].

A characterization of Post algebras A by properties of $\mathscr{P}(A)$ is known [4]. However, we know no such characterization of P_0-lattices. We give an example of a P-algebra A such that $\mathscr{P}(A)$ consists of disjoint maximal chains with at most 2 elements but A is not a P_0-lattice. Let $C = \{0, e, 1\}$ be a 3 element chain and let A be the set of all functions f on an infinite set I to C such that $\{C: f(i) = e\}$ is finite. Since A is a P-subalgebra of a Post algebra of order 3, each chain of prime ideals of A has length at most 2, [7, Th. 7.1]. If $0 = f_0 < f_1 < \cdots < f_{n-1} = 1$ and $S_k = \{i: f_k(i) = e\}$, and if $f = \bigvee_{i=1}^{n-1} b_i f_i$, where b_i are in the center of A, then $\{i: f(i) = e\} \subsetneqq S_1 \cup \cdots \cup S_{n-1}$. Therefore, A does not have a chain base.

REFERENCES

1. R. Balbes and A. Horn, *Stone lattices*, Duke Math. J., **38** (1970), 537-545.
2. R. Braddock, G. Epstein, and H. Yamanaka, *Multiple-valued logic design and applications in binary computeers*, Conference Record of the 1971 Symposium on the Theory and Applications of Multiple-Valued Logic Design, Buffalo, (1971), 13-25.
3. M. Breuer and G. Epstein, *The smallest many-valued logic for the treatment of complemented and uncomplemented error signals*, Conference Record of the 1973 International Symposium of Multiple-Valued Logic, Toronto, (1973), 29-37.
4. C. Chang and A. Horn, *Prime ideal characterization of generalized Post algebras*, Proc. of Symposia in Applied Mathematics, Amer. Math. Soc., **2** (1961).
5. G. Epstein, *The lattice theory of Post algebras*, Trans. Amer. Math. Soc., **95** (1960), 300-317.
6. ———, *An equational axiomatization for the disjoint system of Post algebras*, IEEE Transactions on Computers, C-22 (1973), 422-423.
7. G. Epstein and A. Horn, *P-algebras, an abstraction from Post algebras*, Algebra Universalis, (to appear).
8. G. Grätzer, *Lattice Theory*, San Francisco, 1971.
9. A. Horn, *Logic with truth values in a linearly ordered Heyting algebra*, J. Symbolic Logic, **34** (1969), 395-408.
10. J. Kalman, *Lattices with involution*, Trans, Amer. Math. Soc., **87** (1958), 485-491.
11. T. Katriňák und A. Mitscke, *Stonesche verbände der ordnung n und Postalgebren*, Mathematische Annalen, **199** (1972), 13-30.
12. N. McCoy and D. Montgomery, *A representation of generalized Boolean rings*, Duke Math. J., **3** (1937), 455-459.
13. T. Traczyk, *Axioms and some properties of Post algebras*, Colloquium Mathematicum, **10** (1963), 193-209.

Received March 15, 1974. This work was supported in part by the Office of Naval Research under Contract N00014-67-A-0269-0019, NR 048-299.

UNIVERSITY OF SOUTHERN CALIFORNIA,
INDIANA UNIVERSITY AND
UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

ON THE IRRATIONALITY OF CERTAIN SERIES

P. ERDÖS AND E. G. STRAUS

A criterion is established for the rationality of series of the form $\sum b_n/(a_1, \cdots, a_n)$ where a_n, b_n are integers, $a_n \geqq 2$ and $\lim b_n/(a_{n-1}a_n) = 0$. This criterion is applied to prove irrationality and rational independence of certain special series of the above type.

1. Introduction.

In an earlier paper [2] we proved the following result:

THEOREM 1.1. *If $\{a_n\}$ is a monotonic sequence of positive integers with $a_n \geqq n^{11/12}$ for all large n, then the series*

$$\sum_{n=1}^{\infty} \frac{\varphi(n)}{a_1 a_2 \cdots a_n} \quad \text{and} \quad \sum_{k=1}^{\infty} \frac{\sigma(n)}{a_1 a_2 \cdots a_n} \tag{1.2}$$

are irrational.

We conjectured that the series (1.2) are irrational under the single assumption that $\{a_n\}$ is monotonic and we observed that some such condition is needed in view of the possible choices $a_n = \varphi(n) + 1$ or $a_n = \sigma(n) + 1$. These particular choices do not satisfy the hypothesis $\liminf a_{n+1}/a_n > 0$ but we do not know whether that hypothesis which is weaker than that of the monotonicity of a_n would suffice.

In this note we obtain various improvements and generalizations of Theorem 1.1, in particular by relaxing the growth conditions on the a_n and using more precise results in the distribution of primes.

In § 2 we obtain some general conditions for the rationality of series of the form $\sum b_n/(a_1, \cdots, a_n)$ which are modifications of [2, Lemma 2.29]. In § 3 we use a result of A. Selberg [3] on the regularity of primes in intervals to obtain improvements and generalizations of Theorem 1.1.

2. Criteria for rationality.

THEOREM 2.1. *Let $\{b_n\}$ be a sequence of integers and $\{a_n\}$ a sequence of positive integers with $a_n > 1$ for all large n and*

$$\lim_{n=1} \frac{|b_n|}{a_{n-1}a_n} = 0 . \tag{2.2}$$

Then the series

$$\sum_{n=1}^{\infty} \frac{b_n}{a_1 \cdots a_n} \tag{2.3}$$

is rational if and only if there exists a positive integer B and a sequence of integers $\{c_n\}$ so that for all large n we have

$$Bb_n = c_n a_n - c_{n+1}\,, \qquad |c_{n+1}| < a_n/2\,. \tag{2.4}$$

Proof. Assume that (2.4) holds beyond N. Then

$$\begin{aligned} Ba_1 \cdots a_{N-1} \sum_{n=1}^{\infty} \frac{b_n}{a_1 \cdots a_n} &= \text{integer} + \sum_{n=N}^{\infty} \frac{c_n a_n - c_{n+1}}{a_N \cdots a_n} \\ &= \text{integer} + c_N = \text{integer}\,. \end{aligned}$$

Thus condition (2.4) is sufficient for the rationality of the series (2.3).

To prove the necessity of (2.4) assume that the series (2.3) equals A/B and that N is so large that $a_n \geqq 2$ and $|b_n/(a_{n-1}a_n)| < 1/(4B)$ for all $n \geqq N$. Then

$$\begin{aligned} Aa_1 \cdots a_{N-1} &= Ba_1 \cdots a_{N-1} \sum_{n=1}^{\infty} \frac{b_n}{a_1 \cdots a_n} \\ &= \text{integer} + \frac{Bb_N}{a_N} + \sum_{n=N+1}^{\infty} \frac{Bb_n}{a_N \cdots a_n}\,. \end{aligned} \tag{2.5}$$

If we call the last sum R_N we get

$$\begin{aligned} |R_N| &\leqq \max_{n>N} \frac{|Bb_n|}{a_{n-1}a_n} \sum_{n=N+1}^{\infty} \frac{1}{a_N \cdots a_{n-2}} \\ &< \frac{1}{4} \sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{2}\,. \end{aligned} \tag{2.6}$$

Thus, if we choose c_N to be the integer nearest to Bb_N/a_N and write $Bb_N = c_N a_N - c_{N+1}$ then (2.5) yields that $-c_{N+1}/a_N + R_N$ is an integer of absolute value less than 1 and hence 0, so that

$$\frac{c_{N+1}}{a_N} = R_N = \frac{Bb_{N+1}}{a_N a_{N+1}} + \frac{1}{a_N} R_{N+1} \tag{2.7}$$

or

$$\frac{Bb_{N+1}}{a_{N+1}} = c_{N+1} - R_{N+1}\,. \tag{2.8}$$

From (2.8) it follows that c_{N+1} is the integer nearest to Bb_{N+1}/a_{N+1} and if we write $Bb_{N+1} = c_{N+1}a_{N+1} - c_{N+2}$ we get

$$\frac{Bb_{N+2}}{a_{N+2}} = c_{N+2} - R_{N+2}\,. \tag{2.9}$$

Proceeding in this manner we get the desired sequence $\{c_n\}$.

REMARK. Since (2.2) implies $R_n \to 0$ it follows that for rational values of the series (2.3) we get $c_{n+1}/a_n \to 0$. Thus either $a_n \to \infty$ or $c_n = 0$ and hence $b_n = 0$ for all large n.

COROLLARY 2.10. *Let* $\{a_n\}$, $\{b_n\}$ *satisfy the hypotheses of Theorem* 2.1 *and in addition the conditions that for all large* n *we have* $b_n > 0$, $a_{n+1} \geqq a_n$, $\lim (b_{n+1} - b_n)/a_n \leqq 0$ *and* $\liminf a_n/b_n = 0$. *Then the series* (2.3) *is irrational.*

Proof. According to Theorem 2.1 the rationality of (2.3) implies the existence of a positive integer B and a sequence of integers $\{c_n\}$ so that

$$Bb_n = c_n a_n - a_{n+1}$$

for all large n where $c_{n+1}/a_n \to 0$. Thus

$$\frac{b_{n+1}}{b_n} = \frac{c_{n+1}a_{n+1} - c_{n+2}}{c_n a_n - c_{n+1}} > \frac{(c_{n+1} - \varepsilon)}{c_n a_n} \geqq \frac{c_{n+1} - \varepsilon}{c_n}$$

for all $\varepsilon > 0$ and sufficiently large n. Thus $c_{n+1} > c_n$ would lead to

$$b_{n+1} > \left(1 + \frac{1-\varepsilon}{c_n}\right)b_n > b_n + (1-\varepsilon)\left(a_n - \frac{c_{n+1}}{c_n}\right)/B \tag{2.11}$$

$$> b_n + (1-\varepsilon)^2 a_n/B \,.$$

This contradicts our hypothesis for sufficiently large n. Thus we get $0 < c_{n+1} \leqq c_n$ for all large n and hence b_n/a_n is bounded contrary to the hypothesis that $\liminf a_n/b_n = 0$.

In fact, if we omit the hypothesis $\liminf a_n/b_n = 0$ then we get rational values for the series (2.3) only when $Bb_n = C(a_n - 1)$ with positive integers B, C for all large n.

3. Some special sequences.

THEOREM 3.1. *Let* p_n *be the* nth *prime and let* $\{a_n\}$ *be a monotonic sequence of positive integers satisfying* $\lim p_n/a_n^2 = 0$ *and* $\liminf a_n/p_n = 0$. *Then the series*

$$\sum_{n=1}^{\infty} \frac{p_n}{a_1 \cdots a_n} \tag{3.2}$$

is irrational.

Proof. Since the series (3.2) satisfies the hypotheses of Theorem

2.1 it follows that there is a sequence $\{c_n\}$ and an integers B so that for all large n we have

$$Bp_n = c_n a_n - c_{n+1} . \tag{3.3}$$

For large n an equality $c_n = c_{n+1}$ would imply $c_n \mid B$ and $a_n > p_n$. Since $\{c_n\}$ is unbounded there must exist an index $m \geqq n$ so that $c_m \leqq c_n < c_{m+1}$. But this implies by an argument analogous to (2.11) that

$$p_{m+1} > p_m + a_m/(2B) > \left(1 + \frac{1}{2B}\right)p_m \tag{3.4}$$

which is impossible for large m. Thus we may assume that $c_n \neq c_{n+1}$ for all large n. Now consider an interval $N \leqq n \leqq 2N$. If $c_{n+1} > c_n$ then as in (3.4) we get

$$p_{n+1} > p_n + a_n/(2B) > p_n + \sqrt{p_n}$$

which therefore happens for fewer than $(p_{2N} - p_N)/\sqrt{p_N} < N^{1/2+\varepsilon}$ values in the interval $(N, 2N)$. If $c_{n+1} < c_n$ then we get

$$1 > \frac{c_n a_n - c_{n+1}}{c_{n+1}a_{n+1} - c_{n+2}} > \frac{c_n(a_n - 1)}{c_{n+1}a_{n+1}} > \left(1 + \frac{1}{c_{n+1}}\right)\frac{a_n - 1}{a_{n+1}}$$

so that

$$a_{n+1} > a_n + \frac{a_n - 1}{c_{n+1}} > a_n + 1 . \tag{3.5}$$

Since case (3.5) holds for more than $N/2$ values of n in $(N, 2N)$ we get $a_{2N} > N/2$ and thus for all large n we have $a_n > n/4$, $c_n < p_n/a_n + 1 < \sqrt{n}/4$. Substituting these values in (3.5) we get

$$a_{n+1} > a_n + \sqrt{n} \quad \text{when} \quad c_{n+1} < c_n,\ n \quad \text{large} ; \tag{3.6}$$

so that $a_{2N} > N^{3/2}/2$, contradicting the hypothesis that $\liminf a_n/p_n = 0$.

THEOREM 3.7. *Let $\{a_n\}$ be a monotonic sequence of positive integers with $a_n > n^{1/2+\delta}$ for some positive $\delta > 0$ and all large n. Then the numbers 1, x, y, z are rationally independent. Here*

$$x = \sum_{n=1}^{\infty} \frac{\varphi(n)}{a_1 \cdots a_n} , \qquad y = \sum_{n=1}^{\infty} \frac{\sigma(n)}{a_1 \cdots a_n}$$

and

$$z = \sum_{n=1}^{\infty} \frac{d_n}{a_1 \cdots a_n}$$

where $\{d_n\}$ *is any sequence of integers satisfying* $|d_n| < n^{1/2-\delta}$ *for all large* n *and infinitely many* $d_n \neq 0$.

Proof. Assume that there exist integers A, B, C not all 0 so that setting $b_n = A\varphi(n) + B\sigma(n) + Cd_n$ we get that $S = \sum_{n=1}^{\infty} b_n/(a_1, \cdots, a_n)$ is an integer.

From Theorem 2.1 it follows directly that z is irrational and thus not both A and B can be zero. We consider first the case $A + B \neq 0$ so that without loss of generality we may assume $A + B = D > 0$. Since S satisfies the hypotheses of Theorem 2.1 there exist integers $\{c_n\}$ so that

$$b_n = c_n a_n - c_{n+1} \quad \text{for all large} \quad n .$$

Since $|b_n| < n^{1+\delta/2}$ for all large n we get

$$|c_n| < n^{(1-\delta)/2} \quad \text{for all large} \quad n .$$

Let p_n be the nth prime and set

$$a'_n = a_{p_n},\ b'_n = b_{p_n},\ c'_n = c_{p_n},\ c''_n = c_{p_n+1} ,$$

then

$$b'_n = A(p_n - 1) + B(p_n + 1) + Cd_{p_n} = D_{p_n} + d'_n$$

where

$$d'_n = Cd_{p_n} - A + B \quad \text{with} \quad |d'_n| < n^{(1-\delta)/2} \quad \text{for all large} \quad n .$$

Now

$$b'_n = c'_n a'_n - c''_n$$
$$b'_{n+1} = c'_{n+1} a'_{n+1} - c''_{n+1}$$

so that from

$$\frac{b'_{n+1}}{b'_n} = \frac{Dp_{n+1} + d'_{n+1}}{Dp_n + d'_n} = \frac{p_{n+1}}{p_n} \frac{1 + d'_{n+1}/(Dp_{n+1})}{1 + d'_n/(Dp_n)}$$
$$= \frac{p_{n+1}}{p_n}(1 + o(n^{-(1+\delta)/2}))$$

we get

$$\begin{aligned} \frac{p_{n+1}}{p_n} &= \frac{c'_{n+1}a'_{n+1} - c''_{n+1}}{c'_n a'_n - c''_n}(1 + o(n^{-(1+\delta)/2})) \\ &= \frac{c'_{n+1}}{c'_n} \frac{1 - c''_{n+1}/(a'_{n+1}c'_{n+1})}{1 - c''_n/(a'_n c'_n)}(1 + o(n^{-(1+\delta)/2})) \qquad (3.8) \\ &= \frac{c'_{n+1}}{c'_n}(1 + o(n^{-(1+\delta)/2})) . \end{aligned}$$

Here the last inequality follows from the fact that

$$\left|\frac{c_{n+1}}{c_n}\right| = \left|\frac{(b_{n+1}+c_{n+2})/a_{n+1}}{(b_n+c_{n+1})/a_n}\right| = \frac{|A\varphi(n+1)+B\sigma(n+1)|+O(n^{(1-\delta)/2})}{|A\varphi(n)+B\sigma(n)|+O(n^{(1-\delta)/2})}$$
$$= o(n^{\delta/2}) .$$

From (3.8) we get that $c'_{n+1} > c'_n$ implies

$$p_{n+1} > p_n + \frac{p_n}{c'_n} - p_n^{1/2-\delta/4} > p_n + \frac{1}{2}p_n^{1/2+\delta} \tag{3.9}$$

for all large n.

We now use the following result of A. Selberg [3, Theorem 4].

THEOREM 3.10. *Let $\Phi(x)$ be positive and increasing and $\Phi(x)/x$ decreasing for $x > 0$, further suppose*

$$\Phi(x)/x \to 0 \quad \textit{and} \quad \liminf \log \Phi(x)/\log x > 19/77 \quad \textit{for} \quad x \to \infty .$$

Then for almost all $x > 0$,

$$\pi(x+\Phi(x)) - \pi(x) \sim \frac{\Phi(x)}{\log x} .$$

We now apply this theorem with the choice $\Phi(x) = x^{1/2+\delta}$ to inequality (3.9) and consider the primes $N \leqq p_m < p_{m+1} < \cdots < p_n < 2N$ in an interval $(N, 2N)$ with N large. According to Theorem 3.10 the union of the set of intervals (p_i, p_{i+1}) where p_i, p_{i+1} satisfy (3.9) and $m \leqq i < n$, form a set of total length $< \varepsilon N$ where $\varepsilon > 0$ is arbitrarily small. Also the number of indices i for which (3.9) holds is $o(\sqrt{N})$. Thus by (3.8) and (3.9) we have

$$\frac{c'_n}{c'_m} = \prod_{i=m}^{n-1} \frac{c'_{i+1}}{c'_i} = \prod_{\substack{i=m \\ c'_{i+1}c'_i}}^{n-1} \frac{c'_{i+1}}{c'_i} < \frac{N+\varepsilon N}{N}(1+o(N^{-(+\delta)/2}))^{\sqrt{N}}$$
$$< 1 + 2\varepsilon < 2^{2\varepsilon} .$$

From the monotonicity of a_n it now follows that for any $\varepsilon > 0$ we have

$$|c_n| < n^{\varepsilon} \quad \text{for all large} \quad n. \tag{3.11}$$

Substituting this inequality in (3.9) we get that $c'_{n+1} > c'_n$ would imply

$$p_{n+1} > p_n + \frac{p_n}{c'_n} - p^{1/2+\delta/4} > p_n + \frac{1}{2}p_n^{1-\varepsilon} \tag{3.12}$$

which is impossible for large n when $\varepsilon < 5/12$. Thus $\{c'_n\}$ becomes nonincreasing for large n and hence constant, $c'_n = c$, for large n.

This implies $a_p > p/(c+1)$ for large primes p and by the monotonicity of a_n we get

$$\frac{a_n}{n} > \frac{a_p}{2p} > \frac{1}{4c}$$

where p is the largest prime $\leqq n$.

Now consider the successive equations

$$\begin{aligned} b_p &= ca_p - c_{p+1} \\ b_{p+1} &= c_{p+1}a_{p+1} - c_{p+2} \,. \end{aligned}$$

Thus

$$\begin{aligned} A\varphi(p+1) + B\sigma(p+1) + O(p^{1/2-\delta}) &= c_{p+1}a_{p+1} \\ Dp + O(p^{1/2-\delta}) &= ca_p \end{aligned}$$

for all large primes p. This leads to

$$\left|\frac{A}{D}\frac{\varphi(p+1)}{p+1} + \frac{B}{D}\frac{\sigma(p+1)}{p+1} - \frac{c_{p+1}}{c}\right| < p^{-1/2}\,, \tag{3.13}$$

and hence to the conclusion that the only limit points of the sequence

$$\left\{\frac{A}{D}\frac{\varphi(p+1)}{p+1} + \frac{B}{D}\frac{\sigma(p+1)}{p+1}\,\middle|\, p = \text{prime}\right\}$$

are rational numbers with denominator c. To see that this is not the case, consider first the case $B \neq 0$. Then by Dirichlet's theorem about primes in arithmetic progressions we see that $\sigma(p+1)/(p+1)$ is everywhere dense in $(1, \infty)$. Thus we can choose p so that the distance of $B\sigma(p+1)/D(p+1)$ to the nearest fraction with denominator c is greater that $1/(3c)$ while at the same time $\sigma(p+1)/(p+1)$ is so large that $|A\varphi(p+1)/D(p+1)| < 1/(3c)$, contradicting (3.13). If $B = 0$ we use the fact that $\varphi(p+1)/(p+1)$ is dense in $(0, 1)$ to get the same contradiction.

Finally we must consider the case $A + B = 0$. Here we can go through the same argument as before except that we consider the subsequence $b_{2p} = A\varphi(2p) + B\sigma(2p) + Cd_{2p} = 2Bp + (3B + Cd_{2p}) = 2Bp + O(p^{1/2-\delta})$. As before we get

$$b_{2p} = ca_{2p} - c_{2p+1} \quad \text{for all large primes} \quad p$$

which leads to the wrong conclusion that

$$\left\{\frac{\sigma(2p+1)}{2p+1} - \frac{\varphi(2p+1)}{2p+1}\,\middle|\, p = \text{prime}\right\}$$

has rational numbers with denominator c as its only limit points.

REFERENCES

1. P. Erdös, *Sur certaines series a valeur irrationelle*, Enseignment Math., **4** (1958), 93-100.
2. P. Erdös and E. G. Straus, *Some number theoretic results*, Pacific J. Math., **36** (1971), 635-646.
3. A. Selberg, *On the normal density of primes in small intervals, and the difference between consecutive primes*, Arch. Math. Naturvid., **47** (1943), 87-105.

Received April 16, 1974. This work was supported in part under NSF Grant No. GP-28696.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

MEASURABLE UNIFORM SPACES

ZDENĚK FROLÍK

A uniform space is called $\aleph_0$-measurable if the pointwise limit of any sequence of uniformly continuous functions (real valued) is uniformly continuous. A uniform space is called measurable if the pointwise limit of any sequence of uniformly continuous mappings into any metric space is uniformly continuous.

It is shown that measurable spaces are just metric-fine spaces with the property that the cozero sets form a σ-algebra, or just hereditarily metric-fine spaces.

Metric-fine spaces seem to form a very useful class of spaces; they were introduced by Hager [5], and studied recently by Rice [7] and the author [2], [3]. Separable measurable spaces are studied in Hager [6].

The notation and terminology of Čech [1] is used throughout; for very special terms see Frolík [2]. The main result of the author's [3] is assumed, and [4] may help to understand the motivation.

If X is a uniform space we denote by $\operatorname{coz} X$, zX or BaX accordingly the cozero sets in X (i.e., the sets $\operatorname{coz} f = \{x \mid fx \neq 0\}$ where f is a uniformly continuous function), or the zero sets in X (i.e., the complements of the cozero sets), or the smallest σ-algebra which contains $\operatorname{coz} X$ (equivalently: zX). Since any uniform cover is realized by a mapping into a metric space, the completely coz-additive uniform covers form a basis for the uniformity. Completely coz-additive means that the union of each subfamily is a cozero set.

If X is a uniform space then eX is the set X endowed with the uniformity having the countable uniform covers of X for a basis of uniform covers; eX is a reflection of X in the class of separable uniform spaces (i.e., in spaces Y with $eY = Y$).

We denote by α the usual coreflection into fine uniform spaces. Recall that αX is the set X endowed with the finest uniformity which is topologically equivalent to the uniformity of X. The first theorem is a version of a simple classical result on measurable functions. The equivalence of Conditions 1–5 appears in Hager [6]. This theorem is repeatedly used in the sequel, and therefore an economical proof is furnished.

THEOREM 1. *Each of the following conditions is necessary and sufficient for a uniform space X to be $\aleph_0$-measurable.*

1. *eX is $\aleph_0$-measurable.*

2. $\operatorname{coz} X = zX = BaX$, *and every countable partition ranging in* BaX *is a uniform cover.*

3. *Each countable partition ranging in* BaX *is uniform.*

4. *The countable partitions ranging in* BaX *form a basis for uniform covers of* eX.

5. *A function* $f: X \to R$ *is uniformly continuous if (and only if) the preimages of open sets are the Baire sets in* X.

Proof. It follows immediately from the definition that Condition 1 is necessary and sufficient. Condition 5 implies that X is $\aleph_0$-measurable by the classical result that measurable functions are closed under the operation of taking pointwise limits of sequences ("only if" in Condition 5 is always satisfied). We shall check that each of the Conditions 1–4 implies the subsequent one. Two implications are almost self-evident; namely 2 implies 3, and for 3 implies 4 we must just recall that eX always has a basis consisting of countable covers ranging in $\operatorname{coz} X(BaX)$.

Condition 4 implies Condition 5, because if f is Baire measurable, and if $\mathscr{U}$ is any countable open cover of R, then $f^{-1}[\mathscr{U}]$ is refined by a countable partition ranging in BaX.

It remains to show that Condition 1 implies Condition 2. Assume 1. If G is a cozero set, and if $f \geqq 0$ is a uniformly continuous function with $G = \operatorname{coz} f$, then the characteristic (=indicator) function g of G is a pointwise limit of the uniformly continuous functions

$$f_n = \min(1, m \cdot f) ,$$

and hence g is uniformly continuous by 1. Hence $\operatorname{coz} X = zX$, and hence $\operatorname{coz} X$ is a σ-algebra, and hence $\operatorname{coz} X = BaX$. Now let $\{B_n\}$ be a partition ranging in BaX. Let f_n be the n multiple of the characteristic function of B_n. The limit g of uniformly continuous functions $\sum \{f_n \mid n \leqq k\}$ realizes $\{B_n\}$ in the sense that $\{B_n\} = g^{-1}[U]$ for some uniform cover U of R. This concludes the proof.

THEOREM 2. *For each uniform space* X *let* $M_{\aleph_0}X$ *be the underlying set of* X *endowed with the uniformity having for a basis of uniform covers the covers of the following form:*

$$(*) \qquad \{B_n \cap U_a \mid n \in N,\ a \in A\}$$

where $\{U_a\}$ *is a uniform cover of* X, *and* $\{B_n\}$ *is a partition of* X *ranging in* BaX.

Then:

1. $eM_{\aleph_0}X$ *has for a basis of uniform covers the countable partitions ranging in* BaX.

2. *$M_{\aleph_0}X$ is the meet of X and $M_{\aleph_0}eX$.*

3. *$M_{\aleph_0}X$ is a coreflection of X in the category of $\aleph_0$-measurable spaces.*

Proof. 1. The partitions $\{B_n\}$ are uniform because the cover (*) refines $\{B_n\}$. If $\{V_k\}$ is a countable uniform cover of $M_{\aleph_0}X$, take a cover of the form (*) which refines $\{V_k\}$; we may and shall assume that the union of any subfamily of $\{U_a\}$ belongs to coz X. Put

$$\begin{aligned} C_{kn} &= \bigcup \{B_n \cap U_a \mid B_n \cap U_a \subset V_k\} \\ &= B_n \cap \bigcup \{U_a \mid B_n \cap U_a \subset V_k\} . \end{aligned}$$

Clearly $\{C_{kn}\}$ is a countable cover which ranges in BaX and refines $\{V_k\}$. Now take any partition which refines $\{C_{nk}\}$. This concludes the proof of 1.

2. The assertion 2 follows from 1.

3. Every $M_{\aleph_0}X$ is $\aleph_0$-measurable by Theorem 1 because obviously

$$BaM_{\aleph_0}X = \operatorname{coz} M_{\aleph_0}X = BaX .$$

Let f be a uniformly continuous mapping of an $\aleph_0$-measurable space Y into X. We must show that the mapping $f\colon Y \to M_{\aleph_0}X$ is uniformly continuous. Taking in account the description of $M^{\aleph}{}_0X$, it is enough to show that the preimage under f of any partition $\{B_n\}$ ranging in BaX is a uniform cover of Y, and this follows from Theorem 1 because $f\colon Y \to X$ is self-evidently "Baire measurable".

THEOREM 3. *The sums, quotients and subspaces of $\aleph_0$-measurable spaces are $\aleph_0$-measurable.*

Proof. This follows immediately from Theorem 1.

REMARK. Theorem 3 implies by a purely categorical argument that $\aleph_0$-measurable spaces form a coreflective category, and also the coreflectivity of $\aleph_0$-measurable spaces (established in Theorem 2) implies that the sums and the quotients of $\aleph_0$-measurable spaces are $\aleph_0$-measurable, again by a purely categorial argument.

For separable uniform spaces the next theorem is Hager [6, 6.5].

THEOREM 4. *Each of the following two conditions is necessary and sufficient for a uniform space X to be $\aleph_0$-measurable:*

(1) *Every uniformly continuous function on X factorizes through $M_{\aleph_0}R$.*

(2) *Every uniformly continuous mapping of X into a separable metrizable space S factorizes through $M_{\aleph_0}S$.*

Proof. Since $M_{\aleph_0}$ is a functor Condition (2) is necessary, and clearly (2) implies (1). Condition (1) implies immediately that the pointwise limit of uniformly continuous functions is uniformly continuous.

For the next result we need to recall further definitions. A uniform space X is called metric-fine if for every uniformly continuous mapping f of X into a metric space M the mapping $f\colon X \rightarrow \alpha M$ (see introduction) is uniformly continuous. A uniform space is called (separable metric)-fine if the condition is fulfilled for f's into separable M's. For properties of metric-fine and (separable metric)-fine spaces we refer to Frolík [3]; Hager [5] is a good reference, but it is not enough for our purpose. We need the following description of the coreflections $m_{\aleph_0}X$ and mX of a uniform space X in (separable metric)-fine or metric-fine spaces respectively (see Frolík [3, Theorems 1 and 3]:

The covers of the form

$$\{U_a \cap B_n \mid a \in A,\ n \in N\}$$

form a basis for $m_{\aleph_0}X$, and the covers

$$\{U_a^n \cap B_n \mid n \in N,\ a \in A_n\}$$

form a basis for the uniform covers of mX, where $\{U_a \mid a \in A\}$, $\{U_a^n \mid a \in A_n\}$ are uniform covers of X, and $\{B_n\}$ is a cover of X by elements of $\operatorname{coz} X$; in addition we may assume that all covers are completely $\operatorname{coz} X$-additive.

We also need to know that

$$emX = em_{\aleph_0}X = meX = m_{\aleph_0}eX\ .$$

A uniform space X is called inversion-closed if the set $U(X)$ of all uniformly continuous functions is inversion-closed, and this means, that if $f \in U(X)$ and $fx \neq 0$ for all $x \in X$, then $1/f$ is uniformly continuous.

If X is (separable metric)-fine then X is inversion-closed; this is obvious.

LEMMA 1. *Let Y be an inversion-closed subspace of a uniform space X. For each zero set $Z \subset X - Y$ there exists a zero set $Z' \supset Y$ such that $Z' \cap Z = \phi$. Hence, if Y is a cozero set in X, then Y is a zero set.*

Proof. Take a nonnegative function f in $U(X)$ such that $Z = \{x \mid fx = 0\}$, and let g be the inversion of the restriction of f to Y. Take a uniformly continuous pseudometric d on Y such that f is uniformly continuous on $\langle X, d\rangle$, and g is uniformly continuous on the subspace Y of $\langle X, d\rangle$. The function g extends to a uniformly continuous function g' on the closure Z' of X in $\langle X, d\rangle$; Z' is a zero set in $\langle X, d\rangle$, hence in X. We shall check that $Z' \cap Z = \phi$; if $z \in Z' \cap Z$, then $fz = 0$, and a sequence $\{y_n\}$ in Y converges to z; in $\langle X, d\rangle$, since $fz = 0$ necessarily $fy_n \to 0$; hence the value of the extended g should be $\infty \notin R$, and this contradiction proves the lemma.

REMARK. In the proof of Lemma 1 we used the following simple but useful proposition:

If $Y \subset X$, M is metric, and $g\colon Y \to M$ is uniformly continuous, then there is a uniformly continuous pseudometric d on X (X!) such that g is uniformly continuous on $\langle Y, d\rangle$. (*Proof.* For each n, let u_n be a uniform cover of X such that the trace of u_n on Y refines the inverse image under g of the $1/n$-cover of M. Arrange it so that u_{n+1} star-refines u_n for each n, and let d be the pseudometric associated with the sequence $\{u_n\}$.) The existence of the d in the proof of Lemma 1 now follows. We note that the proposition implies that if $Y \subset X$ and $g\colon Y \to R$ is uniformly continuous, then g has a continuous extension over X: Choose d as above, extend g over the d-closure of Y by uniform continuity, then over all a X by the Tietze-Urysohn Theorem. (If g is bounded, there is a uniformly continuous extension by Katětov's well known theorem.)

THEOREM 5. *The following properties of a uniform space X are equivalent*:

1. *X is $\aleph_0$-measurable.*
2. *X is hereditarily (separable metric)-fine.*
3. *X is (separable metric)-fine, and each subspace is inversion-closed.*
4. *X is (separable metric)-fine, and each cozero subspace of X is inversion-closed.*

Proof. Since $\aleph_0$-measurable is hereditary and implies (separable metric)-fine, Condition 1 implies Condition 2. Next (separable metric)-fine implies inversion-closed, and hence Condition 2 implies Condition 3. Self-evidently Condition 3 implies Condition 4. It remains to show that Condition 4 implies Condition 1. Assume 4. By Lemma 1 we get $\operatorname{coz} X = zX$, hence $\operatorname{coz} X = BaX$. As is noted above, since X is (separable metric)-fine, this implies that X is $\aleph_0$-measurable.

REMARK. For separable spaces, the equivalence of 1 and 2 in Theorem 5 is in Hager [5, 4.2]. We are in a good position to derive several results which are not needed in the sequel, but may help the reader to get better understanding of the spaces used. Again for separable spaces, Propositions 1, 2, 3 and the corollaries appear in Hager [5].

PROPOSITION 1. *The following properties of a subspace Y of a (separable metric)-fine space X are equivalent*:

1. *Y is inversion-closed.*
2. *Y is (separable metric)-fine.*
3. *If $G \supset Y$ is a cozero set, then $Y \subset Z \subset G$ for some zero set Z.*

Proof. By Lemma 1 Condition 1 implies Condition 3, and obviously Condition 2 implies Condition 1. The remaining implication is obtained as follows: If $\{U_n\}$ is a countable cover of Y by cozero sets in Y, then we can take cozero sets G_n in X such that $G_n \cap Y = U_n$, and apply Lemma 1 to Y, the complement Z of $\bigcup \{G_n\}$, and to X. Let G' be the complement of Z'. Clearly all G_n together with G' form a countable cover of X, which consists of cozero sets in X, hence form a uniform cover of X. The $\{U_n\}$ is just the trace of the cover on Y.

COROLLARY. *If $Y \subset X$, then $m_{\aleph_0}Y$ is a subspace of $m_{\aleph_0}X$ if and only if Condition 3 of Proposition 1 holds.*

The following Proposition 2 is a corollary to Corollary.

PROPOSITION 2. *Let Y be a dense subspace of a uniform space X. Then $m_{\aleph_0}Y$ is a subspace of $m_{\aleph_0}X$ if and only if Y is G_δ-dense in X* (i.e., $X - Y$ contains no nonvoid G_δ-set, or equivalently, no nonvoid zero set).

Finally:

PROPOSITION 3. *Let K be a compactification of a topological space X* (completely regular). *The following properties are equivalent*:

1. *K is the Samuel compactification of some metric-fine uniformity on X.*
2. *K is the Samuel compactification of some inversion-closed uniformity on X.*
3. *If G is a cozero set in K, $X \subset G \subset K$, then K is a Čech-Stone compactification of G.*

Proof. Since every metric-fine uniformity is inversion-closed, Condition 1 implies Condition 2. Assume Condition 2, and let g be a bounded continuous function on $G \supset X$, G being a cozero set in K. Pick up a bounded nonnegative continuous function f on K such that $G = \operatorname{coz} f$. The function $f \cdot g$ on G extends to a continuous function h on K; indeed, put $hx = 0$ for x in $K - G$. Thus the restriction of g

to X is the ration of two uniformly continuous functions, namely

$$gx = hx/fx \ ,$$

hence is uniformly continuous, and hence extends to K.

Assume Condition 3, and let us consider the (separable metric)-fine coreflection of the relativization of the uniformity of K to X. We must show that every uniformly continuous bounded function f extends to K, and in view of Condition 3, it is enough to extend f to a cozero set $G \supset X$. Take a countable base $\{U_n\}$ for R and extend each U_n to a cozero set G_n in R; let G be the union of all G_n. Clearly f is uniformly continuous with respect to the relativization of the fine uniformity of G to X, and hence f extends to a continuous function on G. This completes the proof.

COROLLARY. *The Samuel compactification of a uniform space X enjoys the properties in Proposition* 3 *if and only if mX is proximally equivalent to X.*

For more results on rings of uniformly continuous functions we refer to Hager [5].

Now we proceed to measurable spaces which seem to be quite interesting. The first result is a characterization of measurable spaces which will be used to describe the coreflection into measurable spaces, and which connects immediately the theory of measurable spaces with the theory of metric-fine spaces.

THEOREM 6. *A uniform space X is measurable if and only if for any sequence $\{\{U_a^n \mid a \in A_n\}\}$ of uniform covers of X, and for any partition $\{B_n\}$ of X ranging in BaX the cover*

$$(*) \qquad \{B_n \cap U_a^n \mid n \in N,\ a \in A_n\}$$

is uniform.

Proof. First assume that X is measurable, and let (*) be given. We shall realize (*) by a uniformly continuous mapping g into a metric space Y.

Since X is $\aleph_0$-measurable, for each n the cover

$$\mathscr{V}_n = \{B_k \cap U_a^n \mid k \in N,\ a \in A_n\}$$

is uniform, and hence there exists a uniformly continuous mapping f of X into a metric space $\langle M, d\rangle$, which realizes all $\mathscr{V}_n$. We may and shall assume that $d \leqq 1$, and the preimage of the $1/n$-cover of

$\langle M, d\rangle$ under f refines $\mathscr{V}_k$ for $k \leqq n$. In particular, the preimage of the 1-cover of M, d refines $\{B_k\}$. Hence $C_k = f[B_k]$ form a uniformly discrete partition of $\langle M, d\rangle$. Now let Y be the set $N \times M$ endowed with a metric D defined as follows:

$$\begin{aligned} D\langle\langle n, y\rangle, \langle m, z\rangle\rangle &= 1 \text{ if } n \neq m \,, \\ &= \min(1, n.d\langle y, z\rangle) \text{ if } n = m \,. \end{aligned}$$

If we put $d_n = \min(1, n.d)$, then d_n is a metric for M uniformly equivalent to d, and

$$J_n = \{y \longrightarrow \langle n, y\rangle\} : \langle M, d_n\rangle \longrightarrow \langle Y, D\rangle$$

is metric preserving (hence uniform embedding).

Define a sequence $\{h_n\}$ of uniformly continuous mappings of M into Y, and a mapping $h: M \to Y$ (which will not be uniformly continuous in general) as follows:

$$\begin{aligned} gy &= \langle n, y\rangle \quad \text{for } y \text{ in } C_n \,, \\ g_n y &= \langle k, y\rangle \quad \text{for } y \text{ in } C_k \text{ with } k \leqq n \,, \\ &= \langle n, y\rangle \quad \text{for } y \text{ in } C_k \text{ with } k \geqq n \,. \end{aligned}$$

The mappings $g_n: M \to Y$ are uniformly continuous, because

$$\begin{aligned} g_n &= J_k \quad \text{on} \quad B_k \text{ with } k < n \\ g_n &= J_n \quad \text{on} \quad \bigcup \{B_k \mid k \geqq n\} \,. \end{aligned}$$

For each y in M the sequence $\{g_n y\}$ is eventually constant and converges to gy, namely if $y \in C_k$ then $g_n y = gy$ for $n \geqq k$.

Now let $h = g \circ f$, $h_n = g_n \circ f$. The mappings h_n are uniformly continuous, and hence h is uniformly continuous because $\{h_n\}$ converges pointwise to h and X is measurable.

It is easy to check that the preimage of the 1-cover $\mathscr{U}$ of Y under h refines our given cover (*). Indeed,

$$h^{-1}[n \times M] = f^{-1}[C_n] = B_n \,,$$

and if U is the open sphere of radius 1 centered at a point $\langle n, y\rangle$, then $U \subset n \times M$ and $V = J_n^{-1}[U]$ is the open sphere of radius 1 in $\langle M, d_n\rangle$ centered at y, and hence V is the sphere of radius $1/n$ in $\langle M, d\rangle$ centered at y, and hence $f^{-1}[V]$ is contained in some U_a^n. Thus

$$h^{-1}[U] = f^{-1}[J_n^{-1}[U]]$$

is contained in U_a^n, and since $U \subset n \times M$,

$$h^{-1}[U] \subset B_n \cap U_a^n \,.$$

This concludes the proof.

Now assume the condition, and let $\{f_n\}$ be a sequence of uniformly continuous mappings of X into a metric space M, which pointwise converges to a mapping $f: X \to M$. We must show that $f: X \to M$ is uniformly continuous. For each positive number r, and for each n consider the set

$$B_n^r = \{x \mid d\langle f_k x, f_l x\rangle \leqq r \text{ for } k,\ h \geqq n\} .$$

Thus $d\langle fx, f_l x\rangle \leqq r$ for $x \in B_n^r$, $l \geqq n$. Clearly the union of the sequence $\{B_n^r\}$ is X for each r, and each B_n^r belongs to BaX. Now given any positive number ε choose a uniform cover $\{U_a^n \mid a \in A_n\}$ such that the diameter of $f_n[U_a^n]$ is less than $1/3\cdot\varepsilon$ for each a in A_n. Finally put

$$B_n = B_n^r - B_{n-1}^r$$

with $r = 1/3\varepsilon$. Clearly the diameter of each $f[B_n \cap U_a^n]$ is at most ε. By our assumption $\{B_n \cap U_a^n\}$ is a uniform cover, and hence f is uniformly continuous. This concludes the proof.

THEOREM 7. *The sums, subspaces and quotients of measurable spaces are measurable.*

Proof. By a routine argument from Theorem 6.

THEOREM 8. *The following conditions on a uniform space X are equivalent*:

1. *X is measurable.*
2. *X is $\aleph_0$-measurable and metric-fine.*
3. *X is hereditarily (separable-metric)-fine and metric-fine.*
4. *X is hereditarily metric-fine* (i.e., each subspace of X is metric-fine).

Proof. If we compare the characterization of metric-fine spaces recalled above and Theorem 6 we see that Conditions 1 and 2 are equivalent. Conditions 2 and 3 are equivalent by Theorem 5. Finally, obviously Condition 4 implies Condition 3, and is implied by Condition 1 because measurable spaces are hereditary.

It follows from Theorem 7 that measurable spaces are coreflective. Now we shall describe a coreflection measurable spaces and get as *a* byproduct that measurable spaces are coreflective.

THEOREM 8. *For every uniform space X let MX be the set X endowed with the uniformity having for a basis of uniform covers the covers of the form described in Theorem* 6. *Then*:

1. *eMX has for a basis of uniform covers the countable partitions ranging in BaX, and hence eMX is $\aleph_0$-measurable, and $BaX =$*

BaMX.

2. $eMX = eM_{\aleph_0}X = MeX = M_{\aleph_0}eX$.
3. *MX is a coreflection of X in measurable spaces.*

Proof. Let $\{W_k\}$ be a countable cover of MX, and let

$$\{U_a^n \cap B_n \mid n \in N,\ a \in A_n\}$$

be a defining cover which refines $\{W_k\}$. We may and shall assume that $\{U_a^n \mid a \in A_n\}$ are completely coz-additive (such covers form a basis for every uniform space). Put

$$\begin{aligned} C_{kn} &= \bigcup \{U_a^n \cap B_n \mid U_a^n \cap B_n \subset W_k\} \\ &= B_n \cap \bigcup \{U_a^n \mid U_a^n \cap B_n \subset W_k\} \,. \end{aligned}$$

It is easily seen that $\{C_{kn}\}$ is a countable cover which ranges in BaX, and $\{C_{kn}\}$ refines $\{W_k\}$. Thus the countable partitions ranging in BaX form a basis for uniform covers of eMX, hence $BaX = \operatorname{coz} MX = BaMX$, hence eMX is $\aleph_0$-measurable. This proves 1.

It follows from 1 and Theorem 2 that $eMX = eM_{\aleph_0}X$, again by Theorem 2 we have $eM_{\aleph_0}X = M_{\aleph_0}eM$. If X is separable then clearly MX is separable (we may take all $\{U_a^n\}$ in the basis consisting of countable uniform covers, and then the defining covers are countable), and hence $M_{\aleph_0}eX = MeX$. This concludes the proof of 2.

Every space MX is measurable, because it follows from the definition of MX and from 1 that $MMX = MX$, and by Theorem 6 X is measurable if and only if $MX = X$. It remains to show that if $f: Z \to X$ is uniformly continuous and if Z is measurable then $f: Z \to MX$ are measurable. This follows from Theorem 6, and the definition of MX. This concludes the proof.

The next result says that the functor M is metrically determined.

THEOREM 9. *MX is projectively generated by mappings $f: MX \to MP$ where f are uniformly continuous mappings of X into metric spaces P. A uniform space X is measurable if and only if for each uniformly continuous mapping f of X into a metric space P the mapping $f: X \to MP$ is uniformly continuous.*

Proof. The second assertion follows immediately from the first one. The first assertion follows from Theorem 8, because any sequence of uniform covers, and a sequence of Baire sets may be realized in a metric space by a uniformly continuous mapping. To be sure we formulate the fact about the realization of Baire sets in a lemma.

LEMMA 2. *Let $\{B_n\}$ be a sequence of Baire sets in a uniform* *space X. Then there exists a uniformly continuous mapping f into a separable metric space S, and a sequence $\{C_n\}$ of Baire sets in S such that $f^{-1}[C_n] = B_n$ for each n.*

Proof. Take a countable collection $\{U_a \mid a \in A\}$ of cozero sets in X such that all B_n belong to the smallest σ-algebra containing all U_a. We may and shall assume that $A = N$. Take uniformly continuous functions f_n such that

$$U_n = \operatorname{coz} f_n ,$$

and $0 \leqq f_n \leqq 1/2^n$. Then f_n are uniformly continuous, and

$$f\colon X \longrightarrow R^N$$

has the required properties, where f is the reduced product of $\{f_n\}$, i.e., $fx = \{f_n x\}$. This concludes the proof.

The next result describes a nice basis for MX.

THEOREM 10. *The space MX has for a basis of uniform covers the collection of all σ-uniformly discrete (in X) partitions of bounded class in BaX.*

COROLLARY. *A space X is measurable if and only if each σ-uniformly discrete partition of bounded class in BaX is a uniform cover of X.*

We must explain the notion "of bounded class in BaX". We know that BaX is the smallest σ-algebra which contains $\operatorname{coz} X$ (or equivalently, zX). It follows that

$$\begin{aligned} BaX &= \bigcup \{\mathscr{B}_\alpha \mid \alpha < \omega_1\} \\ &= \bigcup \{\mathscr{B}_\alpha^1 \mid \alpha < \omega_1\} \end{aligned}$$

where $\mathscr{B}_0 = \operatorname{coz} X$, $\mathscr{B}_0^1 = zX$, and by induction $\mathscr{B}_\alpha$ ($\mathscr{B}_\alpha^1$, resp.) is obtained from $\bigcup \{\mathscr{B}_\beta \mid \beta < \alpha\}(\bigcup \{\mathscr{B}_\beta^1 \mid \beta < \alpha\})$ by taking all countable intersections (countable unions) or countable unions (countable intersections) according to as α is odd or even.

DEFINITION. A family $\{X_a\}$ is of bounded class in BaX if $\{X_a\}$ ranges in some $\mathscr{B}_\alpha \cup \mathscr{B}_\alpha^1$; the smallest α is called the class of $\{X_a\}$.

Proof of Theorem 10. Let $\{X_a^n \mid n \in N,\ a \in A\}$ be a σ-discrete partition of bounded class, say α, in BaX. Put $B_n = \bigcup \{X_a^n \mid a \in A_n\}$. The sets B_n are of class at most $\alpha + 1$ because $\{X_a^n \mid a \in A_n\}$ are uniformly discrete. The sets X_a^n are cozero sets in B_n, and they form a uniform cover of the subspace B_n of X. By Theorem 6, $\{X_a^n\}$ is a uniform cover of MX.

It remains to show that these covers form a basis. By A. H. Stone Theorem every uniform cover $\mathscr{U}$ of every uniform space X has a uniformly σ-discrete refinement $\mathscr{V} = \bigcup \{V_k\}$; $\mathscr{V}$ is not necessarily uniform, but it is a uniform cover of MX by Theorem 6 (in fact it is a uniform cover of mX, which is the coreflection in metric-fine spaces); indeed put $C_n = \bigcup \{\mathscr{V}_n\}$, $B_n = C_n - \bigcup \{C_k \mid k < n\}$. Now if $\{U_a^n \cap B_n\}$ is a typical defining cover of MX, we may replace each cover $\{U_a^n \mid a \in A\}$ by a uniformly (in X) σ-discrete cover $\{V_k^n \mid k \in N\}$, and put $B_{nk} = B_n \cap V_n^k$. Then $\bigcup \{B_{nk} \cap [\mathscr{V}_k^n]\}$ is a uniformly (in X) σ-discrete cover of a bounded class which refines $\{U_a^n \cap B_n\}$. We need a partition; well order $\{(n, k)\}$ according to ω_0, and take the differences as above. This concludes the proof.

In conclusion we show that for mappings of metric-fine (and hence of measurable) spaces uniform continuity depends on two data only: Cozero sets and "σ-discreteness". I do not know whether this property characterizes metric-fine spaces. Recall (we shall not use it) that just metric-fine proximally fine spaces are completely determined by cozero sets, see Frolík [3, Theorem 4]. First let us stress that the only distinction between metric-fine spaces and measurable ones is in cozero sets.

THEOREM 11. *A uniform space X is measurable if and only if* $\operatorname{coz} X = BaX$, *and X is metric-fine.*

Proof. This follows immediately from Theorems 1 and 7.

THEOREM 12. *Assume that X is metric-fine. A mapping f of X into a uniform space Y is uniformly continuous if* (and obviously, only if) *it enjoys the following properties*:

A. *The preimages of cozero sets are cozero sets.*

B. *The preimages of uniformly σ-discrete families are uniformly σ-discrete.*

Proof. Assume that X is metric-fine, and that $f: X \to Y$ satisfies Conditions A and B. To prove that $f: X \to Y$ is uniformly continuous it is enough to show that

$$h = g \circ f: X \longrightarrow Z$$

is uniformly continuous for every uniformly continuous mapping g of Y into a metrizable space Z. If $\mathscr{U}$ is any uniform cover of Z, then by the A. H. Stone Theorem we can take a uniformly σ-discrete open refinement $\mathscr{V} = \bigcup \{\mathscr{V}_n\}$ (not necessarily uniform), and the preimage of $\mathscr{V}$ under h is, in view of Conditions A and B, uniform by Theorem 2 in Frolík [3], which was recalled just after Theorem 5.

REMARK. M. Rice [7] proved independently that a space X is hereditarily metric-fine if and only if the condition in Theorem 6 is satisfied.

REFERENCES

1. E. Čech, *Topological Spaces*, Academia, Prague, 1965.
2. Z. Frolík, *Interplay of measurable and uniform spaces*, Proc. 2nd International Topological Conference in Yugoslavia, Budva, 1972.
3. ———, *A note on metric-fine spaces*, Proc. Amer. Math. Soc.
4. ———, *Topological methods in measure theory and measurable spaces*, Proc. 3rd Prague Symposium in Topology and Applications, Prague, 1971. Academia (Academic Press), Prague, 1972.
5. A. Hager, *Some nearly fine uniform spaces*, Proc. London Math. Soc.
6. ———, *Measurable uniform spaces*, Fund. Math., **77** (1972), 51-73.
7. M. Rice, *Thesis*, Wesleyan University, 1973.
8. ———, *Covering and function theoretic properties of uniform spaces*, Bull. Amer. Math. Soc.
9. J. Isbell, *Uniform spaces*, Math. Surveys 12, Amer. Math. Soc., 1964.
10. A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc., **54** (1948), 977-982.

Received March 2, 1973.

MATHEMATICS INSTITUTE OF ČSAV
AND
UNIVERSITY OF PITTSBURGH

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

CHARACTERS FULLY RAMIFIED OVER A NORMAL SUBGROUP

STEPHEN M. GAGOLA, JR.

Let H be a group and N a normal subgroup. Assume that χ is an irreducible (complex) character of H, and that the restriction of χ to N is a multiple of some irreducible character of N, say θ. Then $\chi_N = e\theta$, and e is called the ramification index. It is easy to see that it always satisfies $e^2 \leqq |H: N|$, and when equality holds, χ is said to be fully ramified over N. It is this "fully ramified case" which will be studied here in some detail. As an application of some of the methods of this paper, we prove the following solvability theorem in the last section. If H has an irreducible character fully ramified over a normal subgroup N and if p^4 is the highest power of p dividing $|H: N|$ for all primes corresponding to nonabelian Sylow p-subgroups of H/N, then H/N is solvable.

1. Fully ramified triples. Groups of type f.r. To simplify notation, say that (H, N, χ) is a fully ramified triple if χ is an irreducible character of H, N is normal in H, and χ is fully ramified over N. It has been conjectured in [13] that H/N is solvable in this case, and some partial results in this direction appear in [12]. We extend this work in Theorem 4.5 below. It is also possible to show that no known simple group can occur as a homomorphic image of H/N, but we will only need to consider a few cases in this paper (see Lemmas 4.1 and 4.3).

Since we are primarily concerned with the factor group H/N, rather than with H itself, the following theorem (due ultimately to I. Schur and A. H. Clifford) is extremely useful.

THEOREM 1.1. *Let H be a group, N a normal subgroup, and χ an irreducible character of H. Let θ be an irreducible constituent of χ_N, and assume $\chi_N = e\theta$* (i.e. *θ is invariant*). *Then, there exists a group H^*, with an irreducible character χ^*, and a normal subgroup N^* having a faithful irreducible character θ^*, such that*

$$\chi^*_{N^*} = e\theta^*$$
$$N^* \text{ is central in } H^*$$
$$\text{and } H/N \cong H^*/N^* \ .$$

Moreover, the isomorphism is "natural" in the sense that if K is any normal subgroup of H containing N, and K^*/N^* corresponds *to K/N, then*:

$$\chi_K = m(\psi_1 + \cdots + \psi_t)$$
$$\text{and } \chi^*_{K^*} = m(\psi^*_1 + \cdots + \psi^*_t)\ ,$$

where $\psi_1, \cdots, \psi_t$(resp. $\psi^*_1, \cdots, \psi^*_t$) are the distinct conjugates of some irreducible constituent of χ_K(resp. $\chi^*_{K^*}$). In particular, (H, K, χ) is a fully ramified triple if and only if (H^*, K^*, χ^*) is a fully ramified triple.

Many other properties hold than those listed above, but they will not be needed. A proof may be found in [9].

If (H, N, χ) is a fully ramified triple in which N is a central subgroup, then it is easy to see that N must be the center, since $|H: N| = \chi(1)^2 \leqq |H: Z(H)| \leqq |H: N|$. Groups with this property have been referred to in the literature as groups of central type, and in view of Theorem 1.1, there is no essential difference between fully ramified triples and groups of central type.

Lemma 2.3 (a) gives a way of constructing new fully ramified triples from old ones. Unfortunately, these new triples need not correspond to groups of central type, even when the original triple does. Because of this, we state our results for triples, rather than groups of central type.

Define a group G to be of type f.r. if G is isomorphic to H/N, for some fully ramified triple (H, N, χ). Groups of type f.r. have been characterized in [12] as those groups G having a factor set α, over the multiplicative group of complex numbers, such that the corresponding twisted group algebra $C[G]_\alpha$ is simple (or equivalently, has center $\cong C$). We shall have no occasion to use this characterization here.

The next theorem may be used to construct examples of fully ramified triples. It will not be needed in any of the later sections, but it does restrict the kinds of properties that hold in groups of type f.r. In particular, if $\mathscr{P}$ is any property of groups which is inherited by subgroups, then the existence of any solvable group not satisfying $\mathscr{P}$ implies the existence of a (solvable) group of type f.r. not satisfying $\mathscr{P}$.

In the following, $\pi(K)$ denotes the set of prime divisors of the order of K.

THEOREM 1.2. *Let G be any solvable group. Then there is a fully ramified triple (H, Z, χ) with G isomorphic to a subgroup of H/Z. Furthermore, such a triple may be chosen with $\pi(H) = \pi(Z) = \pi(G)$, χ faithful, $Z = Z(H)$ and $|Z|$ square-free.*

Proof. Choose $M \lhd G$ with $|G: M| = p$, a prime. By induction, there is a fully ramified triple $(K, Z(K), \zeta)$ with ζ faithful, M isomorphic to a subgroup of $K/Z(K)$, $\pi(K) = \pi(Z(K)) = \pi(M)$, and $|Z(K)|$ squarefree. If $p \nmid |K|$, replace $(K, Z(K), \zeta)$ by $(K \times C, Z(K) \times C, \zeta \divideontimes \lambda)$, where C is a cyclic group of order p, and λ is a faithful linear character of C. We may therefore assume $\pi(K) = \pi(Z(K)) = \pi(G)$. Let

$$W = \{(z_1, \cdots, z_p) \in Z(K) \times \cdots \times Z(K) \mid \prod z_i = 1\} .$$

Then $W \subseteq Z(K \times K \times \cdots \times K)$, and we may form the quotient

$$U = (K \times \cdots \times K)/W .$$

Let $\eta = \zeta \divideontimes \cdots \divideontimes \zeta \in \mathrm{Irr}\,(K \times \cdots \times K)$, and note that $W = \ker \eta$, so we may view $\eta \in \mathrm{Irr}\,(U)$. It is easy to check that $Z(U) = (Z(K) \times \cdots \times Z(K))/W$, and $\eta(1)^2 = |U: Z(U)|$. Also $Z(U) \cong Z(K)$, so its order is square-free and $\pi(U) = \pi(G)$.

Let $\langle b \rangle$ be a cyclic group of order p. Fix an element $z \in Z(K)$ of order p, and construct an automorphism a of $U \times \langle b \rangle$ as follows:

$$((x_1, x_2, \cdots, x_p)W, b^i) \xrightarrow{a} ((z^i x_p, x_1, x_2, \cdots, x_{p-1})W, b^i) ,$$

for $x_1, \cdots, x_p \in K$ and $0 \leqq i < p$. It is easy to check that a is well defined, and is an automorphism of order p. Using this automorphism, construct the usual semi-direct product $H = (U \times \langle b \rangle) \rtimes \langle a \rangle$. Notice $Z(H) = Z(U)$.

Extend $\eta \in \mathrm{Irr}\,(U \times \langle b \rangle)$ so that $\ker \tilde{\eta} = \langle b \rangle$. Now $\langle b \rangle$ is not normalized by a, so $\tilde{\eta}$ is not an invariant character. It follows that $\chi = \tilde{\eta}^H$ is irreducible, and $\ker \chi = \mathrm{Core}_H(\langle b \rangle) = 1$.

Now:

$$\chi(1)^2 = (p\tilde{\eta}(1))^2 = p^2\eta(1)^2 = p^2|U: Z(U)| = |H: Z(H)| ,$$

so $(H, Z(H), \chi)$ is a fully ramified triple. By construction, $\pi(G) = \pi(H) = \pi(Z(H))$, and $|Z(H)|$ is square-free.

It remains only to check that G is isomorphic to a subgroup of $H/Z(H)$. From the construction of H, the group $H/Z(H)$ is isomorphic to the direct product of a cyclic group of order p (generated by the image of b in $H/Z(H)$) with the wreath product $(K/Z(K)) \wr \langle a \rangle$. As M is isomorphic to a subgroup of $K/Z(K)$, it follows $M \wr \langle a \rangle$ is isomorphic to a subgroup of $(K/Z(K)) \wr \langle a \rangle$. Finally, $G \leqq M \wr (G/M) \cong M \wr \langle a \rangle$, and this completes the proof. (Elementary properties of the wreath product which were used may be found in [8]. See especially pp. 98–99).

2. **Restriction to normal subgroups.** Let (H, Z, χ) be a fully ramified triple, and K a normal subgroup of H containing Z. More can be said about the irreducible constituents of χ_K than is already contained in Clifford's theorem. The explicit statement is Lemma 2.3 below.

We begin first with a lemma describing what happens when K is not assumed to be normal. If α and β are characters of the same group, write $\alpha \leqq \beta$(or $\beta \geqq \alpha$) if $\beta - \alpha$ is zero or a character.

LEMMA 2.1. *Let (H, Z, χ) be a fully ramified triple and let L be a subgroup of H containing Z. Write*

$$\chi_L = \sum_{i=1}^{t} a_i \zeta_i$$

for positive integers $a_1, \cdots, a_t$ and distinct irreducible characters $\zeta_1, \cdots, \zeta_t$ of L. Let θ denote the unique irreducible constituent of χ_Z so that $\chi_Z = e\theta$ and $e^2 = |H: Z|$. Let $b_i = |L: Z| a_i/e$ for $i = 1, \cdots, t$. Then:

(a) $e\theta^L = |L: Z| \chi_L$.

(b) $\theta^L = \sum_{i=1}^t b_i \zeta_i$. *In particular,* $e \mid |L: Z| a_i$ *for* $i = 1, \cdots, t$.

(c) $\zeta_{iZ} = b_i \theta$ *and* $\zeta_i^H = a_i \chi$ *for* $i = 1, \cdots, t$.

(d) $\sum_{i=1}^t a_i^2 = |H: L|$ *and* $\sum_{i=1}^t b_i^2 = |L: Z|$.

(e) *Suppose $t = 1$. Then (L, Z, ζ_1) is a fully ramified triple, and $\chi_L = a_1 \zeta_1$, while $\zeta_1^H = a_1 \chi$, with $a_1^2 = |H: L|$. Suppose additionally that $L \lhd H$. Then (H, L, χ) is a fully ramified triple.*

Proof. Since $e\chi = \theta^H$, the character χ vanishes off of Z. But $(e\theta^L)_Z = e|L: Z|\theta = |L: Z| \chi_Z$. Thus (a) holds. Conclusion (b) is immediate from (a) and the definition of the coefficients b_i. Now $\zeta_{iZ} \leqq \chi_Z = e\theta$. By Frobenius reciprocity, $\zeta_{iZ} = b_i \theta$. Similarly, $\zeta_i^H \leqq \theta^H = e\chi$, and $\zeta_i^H = a_i \chi$.

Now $(\chi_L)^H = |H: L|\chi$ as both sides vanish on $H - Z$, while on Z they equal $|H: L|e\theta$. Also $(\theta^L)_Z = |L: Z|\theta$ holds, again because θ is invariant. We conclude,

$$|H: L| = ((\chi_L)^H, \chi) = (\chi_L, \chi_L) = \sum_{i=1}^{t} a_i^2$$

and

$$|L: Z| = ((\theta^L)_Z, \theta) = ((\theta^L, \theta^L) = \sum_{i=1}^{t} b_i^2 ,$$

proving (d).

When $t = 1$, then (b), (c) and (d) imply conclusion (e).

A slight variation of the next lemma appears in [12], but is given here for completeness.

LEMMA 2.2. *Let (H, Z, χ) be a fully ramified triple, and let L be a Hall π-subgroup of H for some set π of primes. Thus, LZ/Z is a Hall π-subgroup of H/Z. Write $\chi_Z = e\theta$ where $\theta \in \mathrm{Irr}(Z)$, $e^2 = |H: Z|$. Then χ_{LZ} is a multiple of some unique irreducible character, say ζ, of LZ, and (LZ, Z, ζ) is a fully ramified triple. If $Z = Z(H)$, then $(L, L \cap Z, \zeta_L)$ is also a fully ramified triple.*

Proof. Let $\chi_{LZ} = \sum_{i=1}^{t} a_i\zeta_i$ as in Lemma 2.1, with LZ in place of L. Write $e = e_\pi e_{\pi'}$ where $e_\pi^2 = |LZ: Z|$ and $e_{\pi'}^2 = |H: LZ|$. By Lemma 2.1 (*b*), $e_\pi e_{\pi'} \mid |LZ: Z| a_1$. Therefore $e_{\pi'} \mid a_1$, and in particular, $e_{\pi'} \leqq a_1$. Now use the first equation from Lemma 2.1 (d):

$$\sum_{i=1}^{t} a_i^2 = |H: LZ| = e_{\pi'}^2 \leqq a_1^2 .$$

Thus $t = 1$ and the character $\zeta = \zeta_1$ is the only irreducible constituent of χ_{LZ}. Lemma 2.1(e) shows that (LZ, Z, ζ) is a fully ramified triple.

Finally, if Z is central, then $LZ = L \times Z_1$ where Z_1 is an abelian π' group. Thus ζ_L is irreducible and $\zeta(1)^2 = |LZ: Z| = |L: L \cap Z|$. Also $(\zeta_L)_{L\cap Z} = \zeta(1)\theta_{L\cap Z}$ so $(L, L \cap Z, \zeta_L)$ is a fully ramified triple, and the proof is complete.

If p is a prime and G is a group, let $\mathrm{Syl}_p(G)$ denote the set of Sylow p-subgroups of G. Also, for any integer n, let n_p denote the p-part of n.

LEMMA 2.3. *Let (H, Z, χ) be a fully ramified triple with $Z = Z(H)$ and let K be a normal subgroup of H containing Z. Let R be a subgroup of H containing Z with $R/Z \in \mathrm{Syl}_p(H/Z)$, and let ζ be the unique irreducible constituent of χ_R guaranteed by Lemma 2.2. Write*

$$\chi_K = a(\tau_1 + \cdots + \tau_t)$$
$$\zeta_{R\cap K} = b(\sigma_1 + \cdots + \sigma_s)$$

where the τ_i and σ_j are the distinct conjugates of an irreducible constituent of χ_K and $\zeta_{R\cap K}$ respectively. As in Clifford's theorem, choose the unique $\psi_1 \in \mathrm{Irr}(\mathscr{I}_H(\tau_1))$ and $\tilde{\psi}_1 \in \mathrm{Irr}(\mathscr{I}_R(\sigma_1))$ with

$$\psi_1^H = \chi,\ (\psi_1)_K = a\tau_1$$

and

$$\tilde{\psi}_1^R = \zeta,\ (\tilde{\psi}_1)_{R\cap K} = b\sigma_1 .$$

Then:

(a) $a^2t = |H: K|$ *and* $b^2s = |R: R \cap K|$. *Moreover*, $(\mathscr{I}_H(\tau_1), K, \psi_1)$ *and* $(\mathscr{I}_R(\sigma_1), R \cap K, \psi_1)$ *are both fully ramified triples.*

(b) $b = a_p$, $s = t_p$, $\tau_1(1)_p = \sigma_1(1)$ *and under suitable ordering*, $R \cap \mathscr{I}_H(\tau_1) = \mathscr{I}_R(\sigma_1)$.

(c) *H contains a subgroup T containing K which satisfies* $|H: T| = s$, $T \cap R = \mathscr{I}_R(\sigma_1)$ *and* $TR = H$.

(d) *If H/K is a simple group, then* $\mathrm{Core}_R(\mathscr{I}_R(\sigma_1))$ *is either R or $R \cap K$.*

Proof. (a) By Lemma 2.1 (d), $|H: K| = a^2t$. Now,

$$a^2t = |H: K| = |H: \mathscr{I}_H(\tau_1)| \cdot |\mathscr{I}_H(\tau_1): K| = t|\mathscr{I}_H(\tau_1): K|$$

so $a^2 = |\mathscr{I}_H(\tau_1): K|$. Also $(\psi_1)_K = a\tau_1$ and this means that $(\mathscr{I}_H(\tau_1), K, \psi_1)$ is a fully ramified triple. The rest of (a) now follows by applying the above to the fully ramified triple (R, Z, ζ).

(b) By Lemma 2.1 (b), there are integers u and v so that $\theta^{R\cap K} = u(\sigma_1 + \cdots + \sigma_s)$, while $\theta^K = v(\tau_1 + \cdots + \tau_t)$. Hence, there are non-negative integers $a_1, \cdots, a_t$ so that

$$\sigma_1^K = \sum_{i=1}^{t} a_i\tau_i .$$

Now, $(\tau_1)_{R\cap K} \leqq \chi_{R\cap K} = (\chi_R)_{R\cap K} = (|H: R|^{1/2})\zeta_{R\cap K} = |H: R|^{1/2}b(\sigma_1 + \cdots + \sigma_s)$. (The second equality follows from Lemma 2.2 and Lemma 2.1 (e).) Hence, there are nonnegative integers $b_1, \cdots, b_s$ so that

$$(\tau_1)_{R\cap K} = \sum b_j\sigma_j .$$

Comparing degrees: $|K: R \cap K|\sigma_1(1) = (\Sigma a_i)\tau_1(1)$ and

$$\tau_1(1) = (\Sigma b_j)\sigma_1(1) .$$

The second equation implies $\sigma_1(1) | \tau_1(1)$ and thus, $\tau_1(1)/\sigma_1(1)$ divides $|K: R \cap K|$ by the first equation. But $\sigma_1(1)$ is a power of p and $|K: R \cap K|$ is prime to p. It now follows that $\tau_1(1)_p = \sigma_1(1)$.

From (a) we have $a^2t = |H: K|$ and $b^2s = |R: Z|$. As $|R: Z|$ is the order of a Sylow p-subgroup of H/K, we get $a_p^2t_p = b^2s$. We have already derived that $\chi_{R\cap K} = |H: R|^{1/2}b(\sigma_1 + \cdots + \sigma_s)$. Since $\chi_K = a(\tau_1 + \cdots + \tau_t)$, we have by comparing degrees:

$$|H: R|^{1/2}bs\sigma_1(1) = at\tau_1(1) .$$

Equating p parts:

$$bs\sigma_1(1) = a_pt_p\tau_1(1)_p .$$

But $\tau_1(1)_p = \sigma_1(1)$, so $a_pt_p = bs$. We already had $a_p^2t_p = b^2s$, so $a_p = b$ and $t_p = s$.

The group $\mathscr{I}_R(\sigma_1)$ stabilizes σ_1 and acts on the set of irreducible constituents of $\sigma_1^K = \Sigma a_i \tau_i$. As $R \cap K$ is contained in $\mathscr{I}_H(\tau_i)$ for all i, it follows that all orbits of $\mathscr{I}_R(\sigma_1)$ on $\{\tau_1, \cdots, \tau_t\}$ have p-power size. Clearly $a_i = a_j$ if τ_i and τ_j lie in the same orbit, and we may write $\sigma_1^K = \sum_{\mathscr{O}} a_{\mathscr{O}} (\sum_{\tau \in \mathscr{O}} \tau)$, where the outer sum extends over all orbits, and $a_{\mathscr{O}}$ is the common value of a_i for any $\tau_i \in \mathscr{O}$.

Comparing degrees, $|K: K \cap R|\sigma_1(1) = (\Sigma_{\mathscr{O}} a_{\mathscr{O}} |\mathscr{O}|)\tau_1(1)$. Now $\sigma_1(1) = \tau_1(1)_p$ and $|K: K \cap R|$ is prime to p, so $\Sigma a_{\mathscr{O}} |\mathscr{O}| \not\equiv 0 (\mathrm{mod}\, p)$. Thus, there exists an orbit $\mathscr{O}$ with $a_{\mathscr{O}} |\mathscr{O}| \not\equiv 0 (\mathrm{mod}\, p)$. But this means $\mathscr{O} = \{\tau_j\}$ for some j, and $a_j \neq 0$ (so that $\tau_j \leqq \sigma_1^K$). Choose notation so that $\tau_j = \tau_1$. Hence τ_1 is invariant under $\mathscr{I}_R(\sigma_1)$, and thus

$$\mathscr{I}_R(\sigma_1) \subseteqq \mathscr{I}_H(\tau_1) \cap R \, .$$

From (a), $(\mathscr{I}_H(\tau_1), K, \psi_1)$ is a fully ramified triple, so a_p^2 is the order of a Sylow p-subgroup of $\mathscr{I}_H(\tau_1)/K$. Now $|\mathscr{I}_R(\sigma_1): R \cap K| = b^2 = a_p^2 \geqq |(\mathscr{I}_H(\tau_1) \cap R)K: K| = |\mathscr{I}_H(\tau_1) \cap R: R \cap K|$. But we had

$$\mathscr{I}_R(\sigma_1) \subseteqq \mathscr{I}_H(\tau_1) \cap R \, ,$$

so equality holds, and this completes the proof of (b). In fact the last argument shows slightly more, namely

$$\mathscr{I}_R(\sigma_1)K/K \in \mathrm{Syl}_p\, (\mathscr{I}_H(\tau_1)/K) \, .$$

(c) Let $N = N_H(R \cap K)$. As $(R \cap K)/Z \in \mathrm{Syl}_p\, (K/Z)$, the Frattini argument yields $NK = H$. Now N acts on the irreducible constituents of $\theta^{R \cap K}$. Hence N permutes the set $\{\sigma_1, \cdots, \sigma_s\}$. Now $R \subseteqq N$ and R acts transitively on this set, so N acts transitively. Clearly, $R \cap K$ is in the kernel of this action. Moreover, $(N \cap K)/(R \cap K)$ is a normal subgroup of $N/(R \cap K)$ having order prime to p. The set of characters therefore breaks up into k distinct $N \cap K$-orbits, each containing l elements where $l \mid |N \cap K: R \cap K|$ so $(p, l) = 1$. But $s = kl$ is a power of p, so $l = 1$ and $k = s$. This means $N \cap K$ is contained in the kernel of the action. Hence $S = \mathscr{I}_N(\sigma_1)$ contains $N \cap K$ and has index s in N. Thus, $T = SK$ has index s in H, and $TR = H$ is clear. Finally $T \cap R = S \cap R = \mathscr{I}_R(\sigma_1)$.

(d) In the notation of (c), $N/(N \cap K) \cong H/K$ and so $N/(N \cap K)$ is simple. We may assume $s > 1$, in which case N acts transitively on $\{\sigma_1, \cdots \sigma_s\}$ with kernel $N \cap K$. $\mathrm{Core}_R\, (\mathscr{I}_R(\sigma_1))$ is contained in the kernel, so $\mathrm{Core}_R\, (\mathscr{I}_R(\sigma_1)) \subseteqq R \cap (N \cap K) = R \cap K$.

The following consequence of Lemma 2.3 generalizes a theorem appearing in [11].

COROLLARY 2.4. *Let (H, Z, χ) be a fully ramified triple. If $Z \subseteqq K \lhd H$ and χ is induced from a character on K, then H/K is solvable.*

In particular, if K is solvable, then so is H.

Proof. Continuing with the above notation, let $\chi_K = a(\tau_1 + \cdots + \tau_t)$. Then $t = |H: \mathscr{I}_H(\tau_1)| = |H: K| = a^2t$, so $a = 1$. By the previous lemma, H/K possesses a subgroup of index t_p for every prime divisor of t. The solvability of H/K now follows by Philip Hall's theorem (see p. 662 of [8]).

A special case of the next result appears as Theorem 5 of [12]. It is extremely useful in showing that many simple groups do not occur as homomorphic images of groups of type f.r.

THEOREM 2.5. *Suppose (H, Z, χ) is a fully ramified triple, $Z \subseteq K \lhd H$ and $G = H/K$. Let P be a Sylow p-subgroup of G and assume P is cyclic. Then P has a p-complement M in G. If G is simple, we also have:*

(a) *The prime p is unique,* i.e., *all other Sylow q-subgroups for $q \neq p$ are non-cyclic.*

(b) *P is a self-centralizing T.I. set in G.*

(c) *G acts doubly transitively on the cosets of M.*

Proof. By Theorem 1.1, and the remarks following that theorem, we may assume $Z = Z(G)$, so that Lemma 2.3 becomes applicable. Let $R/Z \in \mathrm{Syl}_p(H/Z)$, and let $\zeta, \sigma_1, \cdots, \sigma_s$, and $\tau_1, \cdots, \tau_t$ be as in Lemma 2.3. Now $\mathscr{I}_R(\sigma_1)$ has a character which is fully ramified over $R \cap K$, and the factor group $\mathscr{I}_R(\sigma_1)/R \cap K$ is cyclic. Thus, all irreducible constituents of $\sigma_1^{\mathscr{I}_{R(\sigma_1)}}$, including the fully ramified one, are extensions of σ_1. (See p. 54 of [3].) This can only happen if $\mathscr{I}_R(\sigma_1) = R \cap K$, so $s = |R/(R \cap K)| = |P|$. (This could also be seen by applying Theorem 1.1 to the group $\mathscr{I}_R(\sigma_1)$.) By Lemma 2.3 (c), $H/K = G$ has a subgroup M of index s, and this is clearly a p-complement in G. Suppose now G is simple. If a Sylow q-subgroup, say Q, of G is cyclic for some other prime q, then Q acts faithfully on the $|P|$ cosets of M, so that $|Q| < |P|$. Interchanging the roles of P and Q yields $|Q| > |P|$, and this contradication establishes (a).

To prove that P is a self-centralizing T.I. set in G, it suffices to show $C_G(\Omega_1(P)) = P$. Let $C = C_G(\Omega_1(P))$. As $P \cap M = 1$ and $PM = G$, we have $C = P(C \cap M)$, so that $C \cap M$ is a p-complement in C. Now $N_C(P)$ acts on P and centralizes $\Omega_1(P)$, and hence centralizes P by Fitting's lemma (see p. 178 of [6]). But then by Burnside's transfer theorem (see p. 419 of [8], also p. 252 of [6]), C has a normal p-complement, which must be $C \cap M$. Now $G = CM$ and $C \cap M$ is normal in C, so the normal closure of $C \cap M$ is contained in M. Hence, $C \cap M \subseteq (C \cap M)^G = 1$, as G is simple. Thus $C = P$, proving (b).

If $|P| = p$, then G must be doubly transitive on the cosets of M by a theorem of Burnside's (see p. 609 of [8]). If $|P| > p$, then P is a B-group (p. 65 of [15]) and it suffices to show that G is primitive on the cosets of M. Suppose G is not primitive, so there exists a subgroup L with $M < L < G$. But then $1 < P \cap L$ and $P \cap L$ is normalized by P. Since $G = PL$, it follows that the normal closure of $P \cap L$ is contained in L, contradicting the simplicity of G. (This last assertion can also be proved by considering the Brauer tree of the principal p-block of G. It can be proved that the principal character can be connected only to a nonexceptional character, and the double transitivity follows.)

3. **Special elements.** Let (H, Z, χ) be a fully ramified triple, and let K be a normal subgroup of H containing Z. Also, let ξ denote an irreducible constituent of χ_K. Information about the group H/K was obtained in the previous section by considering the possible indices for the inertia group of ξ. In this section, we obtain information about the group K/Z by considering elements of K at which ξ does not vanish (for all possible ξ). Under the right conditions, K/Z will have a proper normal subgroup. The main application of the methods of this section are contained in Corollary 3.6.

The following concept first appears in [5], and a slight variation of it appears in [9].

DEFINITION. Let $N \lhd G$, and let $\psi \in \mathrm{Irr}\,(N)$ be invariant under G. For every $x, y \in G$ with $[x, y] = x^{-1}y^{-1}xy \in N$, define the complex number $\langle\langle x, y\rangle\rangle$ as follows. Extend ψ to $\hat{\psi}$ on $\langle N, y\rangle$. Now x normalizes the group $\langle N, y\rangle$, and fixes ψ, so $\hat{\psi}^x$ is another extension of ψ. It follows that $\hat{\psi}^x = \lambda\hat{\psi}$, where λ is a linear character of $\langle N, y\rangle$ with N in its kernel. Moreover, λ is uniquely determined, i.e., is independent of the choice of the extension $\hat{\psi}$. Define $\langle\langle x, y\rangle\rangle$ to be $\lambda(y)$.

The definition above of course depends on ψ. Properties of the map $\langle\langle,\rangle\rangle$ may be found in [9]. In particular, $\langle\langle x, y\rangle\rangle$ is multiplicative in x and y whenever it is defined, and $\langle\langle x, y\rangle\rangle = \langle\langle y, x\rangle\rangle^{-1}$ if $\langle\langle x, y\rangle\rangle$ is defined.

We are now ready to define special elements.

DEFINITION. Let $N \lhd G$ and $\psi \in \mathrm{Irr}\,(N)$, with ψ invariant in G. Let $\langle\langle\,,\rangle\rangle$ be defined as above. Say that $g \in G$ is *special* if $\langle\langle x, g\rangle\rangle = 1$, for all x satisfying $xN \in C_{G/N}(gN)$.

If g is special, then so is every conjugate of g, and every element the of the coset gN. We may therefore speak of the special classes of G/N. The following theorem and its proof appear in [5].

THEOREM 3.1. *Let $N \lhd G$ and $\psi \in \mathrm{Irr}(N)$ be G-invariant. Define special conjugacy classes of G/N as indicated above. Then, the number of distinct irreducible constituents of ψ^G is the same as the number of special classes of G/N.*

Because of Theorem 1.1, the case that $N \subseteq Z(G)$, and ψ is a faithful linear character of N, deserves to be singled out. In this case, the computation of $\langle\!\langle x, y\rangle\!\rangle$, for $x, y \in G$ with $[x, y] \in N$, becomes easier to carry out: Let ψ be an extension of ψ to $\langle N, y\rangle$. This is an abelian group, so $\hat{\psi}$ is linear. Moreover, $\hat{\psi}^x = \lambda\hat{\psi}$ for $\lambda \in \mathrm{Irr}(\langle N, y\rangle)$ with $N \subseteq \ker \lambda$. All characters appearing in this equation are linear, and so we may solve for λ: $\lambda = \hat{\psi}^{-1}\hat{\psi}^x$.

Evaluating at y yields:

$$\begin{aligned}\langle\!\langle x, y\rangle\!\rangle &= \lambda(y) = \hat{\psi}^{-1}(y)\hat{\psi}^x(y) = \hat{\psi}(y^{-1})\hat{\psi}(xyx^{-1}) = \hat{\psi}(y^{-1}xyx^{-1}) \\ &= \hat{\psi}([y, x^{-1}]) = \hat{\psi}([x, y]) = \psi([x, y]) \,.\end{aligned}$$

Thus $\langle\!\langle x, y\rangle\!\rangle = \psi([x, y])$, for all $x, y \in G$ with $[x, y] \in N$. As ψ is faithful, we may identify $\langle\!\langle\,,\rangle\!\rangle$ with $[\,,]$, defined for all pairs of elements satisfying $[x, y] \in N$. In particular, it is easy to see that $x \in G$ is special if and only if $C_G(x) = C_G(xN \bmod N)$.

The following easy consequence of Theorem 3.1 will be useful later:

COROLLARY 3.2. *Let (H, Z, χ) be a fully ramified triple. Then H/Z contains no self-centralizing cyclic subgroups, unless $H = Z$.*

Proof. As usual, we may assume $Z = Z(H)$, and then we may identify $\langle\!\langle\,,\rangle\!\rangle$ with $[\,,]$. If θ is the unique constituent of χ_Z, then χ is the unique constituent of θ^H. By Theorem 3.1, there is only one special class of H/Z, and this must be the class of the identity element. Suppose $\langle gZ\rangle$ is a self-centralizing subgroup of H/Z. Then $[x, g] \in Z$ implies $x \in \langle Z, g\rangle$, and since this last group is abelian, $[x, g] = 1$. But then g is special, and since $\bar{1}$ is the only special class of H/Z, it follows that $g \in Z$. Hence,

$$H = C_H(g) \subseteq \langle g, Z\rangle = Z\,, \quad \text{so } H = Z\,.$$

LEMMA 3.3. *Let $\psi \in \mathrm{Irr}(N)$, where $N \lhd G$ and ψ is faithful. Assume $N \subseteq Z(G)$, so that ψ is invariant in G, and special elements of G are defined. Let χ be a constituent of ψ^G. If g is not special, then $\chi(g) = 0$.*

Proof. For g not special, there exists $x \in G$ with $[x, g] \in N$ and $[x, g] \neq 1$. Since χ is a class function:

$$\chi(g) = \chi(x^{-1}gx) = \chi(gg^{-1}x^{-1}gx) = \chi(g[g, x]) = \chi(g)\psi([g, x]) ,$$

where the last equality follows from the fact that $[g, x]$ is represented as a scalar matrix, with scalar $\psi([g, x])$, in any representation affording χ. But $\psi([g, x]) \neq 1$, as ψ is faithful, so $\chi(g) = 0$.

When $N \subseteqq Z(G)$ and ψ is a faithful character of N, the following gives a stronger relation between constituents of ψ^G and the special classes of G/N than is already implied in Theorem 3.1.

THEOREM 3.4. *Let $N \subseteqq Z(G)$ and $\psi \in \mathrm{Irr}(N)$, with ψ faithful, and let $\chi_1, \cdots, \chi_m$ be the distinct irreducible constituents of ψ^G. Let $g_1, \cdots, g_m$ be any m elements of G. Then, the matrix $(\chi_i(g_j))$ is nonsingular if and only if $g_1, \cdots, g_m$ represent the m distinct special classes of G/N.*

Proof. (Only if) If $(\chi_i(g_j))$ is nonsingular, then certainly for every j, the jth column is nonzero. By the previous lemma, this means that g_j is special. Let — denote the natural map from G to G/N. We need to check that $\bar{g}_1, \cdots, \bar{g}_m$ lie in distinct conjugacy classes. Suppose $\bar{g}_i$ is conjugate to $\bar{g}_j$. Then $x^{-1}g_ix = ng_j$, for some $x \in G$ and $n \in N$. Then, for every k:

$$\chi_k(g_i) = \chi_k(x^{-1}g_ix) = \chi_k(ng_j) = \psi(n)\chi_k(g_j) ,$$

so that the i-th and j-th columns of the matrix differ by the scalar multiple $\psi(n)$. This can only happen if $i = j$, and we are done with this half of the theorem.

(If) Suppose $g_1, \cdots, g_m$ represent the m distinct special conjugacy classes of G/N. Again let - denote the map $G \to G/N$. Then,

$$\begin{aligned}
\delta_{ij} = (\chi_i, \chi_j) &= (1/|G|) \sum_{g \in G} \chi_i(g)\chi_j(g^{-1}) \\
&= (1/|G|) \sum_{g \text{ special}} \chi_i(g)\chi_j(g^{-1}) \\
&= (1/|G|) \sum_{\nu=1}^{m} \sum_{g \sim g_\nu} \chi_i(g)\chi_j(g^{-1}) \\
&= (1/|G|) \sum_{\nu=1}^{m} |\bar{G}: C_{\bar{G}}(\bar{g}_\nu)| \cdot |N| \cdot \chi_i(g_\nu)\chi_j(g_\nu^{-1}) .
\end{aligned}$$

The third equality follows from Lemma 3.3, and the last follows from the fact that $\chi_i(g)\chi_j(g^{-1})$ is constant on cosets of N. We therefore have:

$$\delta_{ij} = (|N|/|G|) \sum_{\nu=1}^{m} \chi_i(g_\nu) |\bar{G}: C_{\bar{G}}(\bar{g}_\nu)| \chi_j(g_\nu^{-1}) .$$

Writing this last identity in matrix form:

$$I = (|N|/|G|)(\chi_i(g_j)) \operatorname{diag}(|\bar{G}: C_{\bar{G}}(\bar{g}_i)|)(\chi_j(g_i^{-1})) ,$$

where I is the $m \times m$ identity matrix. Hence, $(\chi_i(g_j))$ is nonsingular, and we are done.

THEOREM 3.5. *Let $(H, Z(H), \chi)$ be a fully ramified triple with χ faithful. Let $Z = Z(H) \subseteq K \lhd H$ and let R be a subgroup of H containing Z with $R/Z \in \mathrm{Syl}_p(H/Z)$. Finally, let θ be unique constituent of χ_Z, and $g_1, \cdots, g_s$ be representatives of the distinct special classes of $(R \cap K)/Z$, computed with respect to θ. Then:*

(a) *The $s|Z|$ elements, zg_i, for $z \in Z$ and $1 \leqq i \leqq s$, are all special in K, and lie in distinct conjugacy classes of K.*

(b) *If $g \in R \cap K$ is a special element of K, then g is special in $R \cap K$. In particular, g is $R \cap K$-conjugate to a unique element of the form zg_i.*

(c) *If $g, h \in R \cap K$ and g is special in $R \cap K$, then $g \sim_K h$ implies $g \sim_{R \cap K} h$.*

REMARK. The above implies that there is a natural correspondence between conjugacy classes of special elements in $R \cap K$ and conjugacy classes of special elements of K which meet $R \cap K$. The correspondence is given by $\mathscr{L} \mapsto \mathscr{L}^K$, where $\mathscr{L}$ is a conjugacy class of $R \cap K$ consisting of special elements, and $\mathscr{L}^K$ is the unique class of K containing $\mathscr{L}$. The inverse is given by $\mathscr{M} \mapsto \mathscr{M} \cap (R \cap K)$, where $\mathscr{M}$ is a class of special elements of K which meets $R \cap K$.

Proof of Theorem 3.5. Following the notation of Lemma 2.3, let $\sigma_1, \cdots, \sigma_s$ and $\tau_1, \cdots, \tau_t$ be the distinct irreducible constituents of $\chi_{R \cap K}$ and χ_K respectively. We know there are s constituents of $\chi_{R \cap K}$ because there are s special classes in $(R \cap K)/Z$. Let $Z[\tau_1, \cdots, \tau_t]$ denote the additive subgroup of the character ring of K generated by $\tau_1, \cdots, \tau_t$, and similarly define $Z[\sigma_1, \cdots, \sigma_s]$. Let r denote the restriction map from $Z[\tau_1, \cdots, \tau_t]$ to $Z[\sigma_1, \cdots, \sigma_s]$. Since $\chi_{R \cap K} = (\chi_R)_{R \cap K}$, it is clear that r maps $Z[\tau_1, \cdots, \tau_t]$ into $Z[\sigma_1, \cdots, \sigma_s]$. Reducing coefficients mod P, we have the following commutative diagram:

$$\begin{array}{ccc} Z[\tau_1, \cdots, \tau_t] & \xrightarrow{r} & Z[\sigma_1, \cdots, \sigma_s] \\ \downarrow & & \downarrow \\ Z_p[\tau_1, \cdots, \tau_t] & \xrightarrow{\bar{r}} & Z_p[\sigma_1, \cdots, \sigma_s] \end{array}$$

Now R acts on $\{\tau_1, \cdots, \tau_t\}$, and because $R \cap K \lhd R$, R acts on

$\{\sigma_1, \cdots, \sigma_s\}$. The group $R \cap K$ is contained in the kernel of both actions, so the p-group $R/(R \cap K)$ acts on both sets. This action may be extended in the natural way to each of the four additive groups above, so that each such group is an $R/(R \cap K)$-module. The second row of groups may be viewed as $Z_p[R/(R \cap K)]$-modules. All maps in the above diagram are $R/(R \cap K)$-homomorphisms. Since $R/(R \cap K)$ is a p-group acting transitively on $\{\sigma_1, \cdots, \sigma_s\}$, the module $Z_p[\sigma_1, \cdots, \sigma_s]$ contains a unique maximal submodule $M = \{\Sigma l_i\sigma_i \mid l_i \in Z_p \text{ and } \Sigma l_i = 0\}$.

As in the (b) part of Lemma 2.3, write

$$r(\tau_1) = (\tau_1)_{R\cap K} = \sum_{j=1}^{s} b_j\sigma_j .$$

As $\tau_1(1)_p = \sigma_1(1)$, it follows that $\Sigma b_j \not\equiv 0 \bmod p$. Hence

$$\bar{r}(\tau_1) = \sum_{j=1}^{s} \bar{b}_j\sigma_j \notin M ,$$

where $\bar{b}_j$ denotes the residue class of $b_j \bmod p$. Since $\bar{r}$ is a $Z_p[R/(R \cap K)]$-map, it follows that $\bar{r}$ is surjective.

Now define the $t \times s$ matrix $B = (b_{ij})$ as follows:

$$r(\tau_i) = (\tau_i)_{R\cap K} = \sum_{j=1}^{s} b_{ij}\sigma_j .$$

Let $\bar{B} = (\bar{b}_{ij})$ be the matrix B with all entries reduced $\bmod p$. Then $\bar{B}$ is the matrix of $\bar{r}$ using the natural bases. Thus $\bar{B}$, and hence B itself, has rank s. Now

$$(\tau_i(g_j)) = B(\sigma_i(g_j)) ,$$

where $(\sigma_i(g_j))$ is nonsingular by Theorem 3.4, and B has rank s. Thus, $(\tau_i(g_j))$ has rank s, so that its columns are linearly independent. This means that $g_1, \cdots, g_s$ represent distinct special classes in K/Z.

We now have to check that there is no K-fusion among the elements zg_i. Suppose $zg_i \sim_K z'g_j$. The above implies that $i = j$. Now choose τ_k so that $\tau_k(g_i) \neq 0$. Then

$$\theta(z)\tau_k(g_i) = \tau_k(zg_i) = \tau_k(z'g_i) = \theta(z')\tau_k(g_i) ,$$

and so $\theta(z) = \theta(z')$. But θ is faithful because χ is, and so $z = z'$. This proves (a).

Now suppose $g \in R \cap K$ and g is special in K. We have just shown that $g_1, \cdots, g_s$ represent distinct (special) conjugacy classes in K/Z. We may therefore find $\{x_2, \cdots, x_s\} \subseteq \{g_1, \cdots g_s\}$ such that $g = x_1, x_2, \cdots, x_s$ represent distinct conjugacy classes in K/Z, so that the $t \times s$ matrix $(\tau_i(x_j))$ has rank s, by Theorem 3.4. Now $(\tau_i(x_j)) = B(\sigma_i(x_j))$, and this equation implies that the $s \times s$ matrix $(\sigma_i(x_j))$ is non-singular. By Theorem 3.4 again, $x_1 = g$ is special in $R \cap K$. Hence g is conjugate in $R \cap K$ to some element of the form zg_i. Uniqueness of this element is clear, as these elements are not fused in K even. This proves (b).

Suppose now $g, h \in R \cap K$, g is special in $R \cap K$ and $g \sim_K h$. By (a) above, g is special in K and hence so is h. However, $h \in R \cap K$, so by (b) above, h is special in $R \cap K$. From (b) again, g and h are conjugate in $R \cap K$ to elements of the form zg_i and $z'g_j$ respectively, for some $z, z' \in Z$ and $1 \leqq i, j \leqq s$. Hence $zg_i \sim_K z'g_j$, and from (a) we get $z = z'$, $i = j$. Thus, g and h are fused in $R \cap K$, completing the proof of (c).

As an application of the above non-fusion theorem, we have:

COROLLARY 3.6. *Let (H, Z, χ) be a fully ramified triple. Let $K = O^p(H)Z$, and assume that a Sylow p-subgroup of K/Z is abelian. Then (H, K, χ) is a fully ramified triple, and for ψ the unique constituent of χ_K, the triple (K, Z, ψ) is fully ramified.*

Proof. By applying Theorem 1.1, we may assume $Z = Z(H)$ and that χ is faithful. Let $R/Z \in \mathrm{Syl}_p(H/Z)$, and let $\tau_1, \cdots, \tau_t$ and $\sigma_1, \cdots, \sigma_s$ be as in Lemma 2.3. Then $t = s$ as $|H\colon K|$ is a power of p. If $t = 1$, this means that (H, K, χ) is a fully ramified triple, and hence so is (K, Z, τ_1), and we are done.

Suppose then $t = s > 1$. Let $N = N_K(R \cap K)$, and let—denote the natural map $K \rightarrow K/Z$. Thus $\bar{N} = N_{\bar{K}}(\overline{R \cap K})$. Since $s > 1$, there is an element $g \in R \cap K$ which is special in $R \cap K$ and $g \notin Z$. If $x \in N$, then $g^x \in R \cap K$, and clearly $g \sim_K g^x$. By Theorem 3.5 (c), g^x is conjugate in $R \cap K$ to g. But $\overline{R \cap K}$ is abelian, so $\bar{g}^{\bar{x}} = \bar{g}$, and this shows:

$$\bar{1} \neq \bar{g} \in \overline{R \cap K} \cap Z(N_{\bar{K}}(\overline{R \cap K})) .$$

However, this implies $O^p(H)Z = O^p(K)Z < K$, (see p. 253 of [6]). Thus, the case $s > 1$ leads to a contradiction, and the corollary is proved.

4. **A solvability theorem.** The final theorem of this section is a solvability theorem for certain groups of type f.r. In order to prove that theorem, it is first necessary to show that certain groups do not occur as homomorphic images of groups of type f.r.

LEMMA 4.1. *Let G be a simple subgroup of A_9* (the alternating group on 9 letters). *Then G is not a homomorphic image of a group of type f.r.*

Proof. Suppose G is such a homomorphic image. Now $|G|$ divides $9\cdot 8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3 = 2^6\cdot 3^4\cdot 5\cdot 7$.

Suppose $5 \mid |G|$. By Theorem 2.5 with $p = 5$, we get $G \leqq A_5$, and so $G \cong A_5$. This contradicts Theorem 2.5 (a), and so $5 \nmid |G|$. By Burnside's $p^a q^b$ theorem (see p. 131 of [6]), $\pi(G) = \{2, 3, 7\}$. Hence, $7 \mid |G|$, and using Theorem 2.5 with $p = 7$, we get $G \leqq A_7$. Thus $|G|$ divides $7\cdot 6\cdot 4\cdot 3 = 2^3\cdot 3^2\cdot 7$, as $5 \nmid |G|$.

Using Theorem 2.5 (a) again, a Sylow 3-subgroup of G cannot be cyclic. We therefore have $|G| = 2^j\cdot 3^2\cdot 7$, for some j. By Burnside's transfer theorem, $P < N_G(P) \leqq N_{A_7}(P)$, where P is a Sylow 7-subgroup of G. This last group has order 21, so $|N_G(P)| = 21$. By Sylow's theorem, $2^j\cdot 3 \equiv 1 \bmod 7$, and this is the final contradiction.

The next fact which is needed is a purely number theoretic statement, due to G. D. Birkhoff and H. S. Vandiver, which first appeared about the turn of the century.

LEMMA 4.2. *Let a and n be integers both greater than one. Then, except for the following two cases, there exists a prime divisor p of $(a^n - 1)$, satisfying $p \nmid (a^m - 1)$ for all m with $1 \leqq m < n$:*

(I) $n = 2$ *and a is a Mersenne number*, i.e. $a + 1$ *is a power of* 2.

(II) $n = 6$ *and* $a = 2$.

A proof of the above lemma for $n \geqq 3$ may be found in [1], where, in fact, a more general version is given. Of course, the case $n = 2$ is a triviality.

The above lemma is extremely useful, when used in conjunction with Theorem 2.5, in eliminating known simple groups from occuring as factor groups of groups of type f.r. However, in this section, we shall only need the following:

LEMMA 4.3. *Let* $\mathrm{PSL}(2, p^n) \leqq X \leqq \mathrm{P\Gamma L}(2, p^n)$, *where p is a prime, and $p^n \geqq 4$. Then X is not the homomorphic image of any group of type f.r.*

Proof. We first note that $\mathrm{PSL}(2, p^n)$ can have no subgroup of index $q^a \neq 1$, where q^a is a prime power less than p^n. This is true because $\mathrm{PSL}(2, p^n)$ contains a proper subgroup of index $m < p^n$, only in the case $p^n = 9$ and $m = 6$ (see p. 214 of [8]). This proves the statement, as 6 is not a prime power.

Suppose X is a homomorphic image of H/Z, where (H, Z, χ) is a fully ramified triple. Let K be the kernel of this homomorphism, and S the inverse image of $\mathrm{PSL}(2, p^n)$. Then, $Z \subseteqq K \subseteqq S \subseteqq H$, where K and S are normal in H, $H/K \cong X$ and $S/K \cong \mathrm{PSL}(2, p^n)$. Clearly, $|H: S|$ divides $2n$, as $|\mathrm{P\Gamma L}(2, p^n): \mathrm{PSL}(2, p^n)| = n(2, p-1)$.

Suppose there exists a prime q satisfying the following conditions:

(i) $q \mid |\mathrm{PSL}(2, p^n)|$

(ii) $q \neq p$

(iii) q is odd

(iv) $1 + p^n$ is not a power of q

(v) $q \nmid n$.

Then q divides exactly one of $(p^n + 1)$ or $(p^n - 1)$, as q is odd, and a Sylow q-subgroup of S/K is cyclic of order $< p^n$ by (iv). By (v), a Sylow q-subgroup of S/K is also one for H/K, implying that H/K has a q-complement, by Theorem 2.5. But then S/K also has a q-complement, contradicting the first paragraph.

We now prove, under the hypothesis $p^n \geqq 4$, a prime q can always be chosen satisfying (i)–(v) above.

Suppose q is an odd prime dividing $p^n - 1$, but not dividing $p^m - 1$ for any $m < n$ (if $n = 1$, this last condition is vacuously true). Clearly, q satisfies (i)–(iv) above. Now q divides $p^{q-1} - 1$, forcing $n \leqq q - 1$, so that q also satisfies (v). In particular, we are done if $n = 1$, unless $p - 1$ is a power of 2. If $n > 1$, then Lemma 4.2 is applicable (with p in place of a), and any prime satisfying the conclusion of that lemma also satisfies (i)–(v) above. This brings us to one of the following cases:

(a) $n = 1$ and $p - 1$ is a power of 2

(b) $n = 2$ and $p + 1$ is a power of 2

(c) $n = 6$ and $p = 2$.

We consider these cases in turn.

Case (a). Since $p^n \geqq 4$, it follows that $p + 1$ is even, and is not a power of 2. Any odd prime divisor of $p + 1$ satisfies (i)–(v) above, and we are done in this case.

Case (b). Since $p^2 + 1$ is twice an odd number, in this case, let q be an odd prime divisor of $p^2 + 1$. Again, it is readily checked that q satisfies (i)–(v) above.

Case (c). Here $|\mathrm{PSL}(2, p^n)| = 65 \cdot 64 \cdot 63$, and the prime $q = 5$ satisfies the five conditions above.

Let (H, Z, χ) be a fully ramified triple, and assume that H/Z has an abelian Sylow p-subgroup for some prime p. We saw in the previous section (Corollary 3.6) that $(H, O^p(H)Z, \chi)$ is also a fully ramified triple. This suggests the following definition:

DEFINITION. Let Q be a p-group of type f.r. Say that Q is *reductive* if, for every fully famified triple (H, Z, χ) with Q isomorphic to a Sylow p-subgroup of H/Z, the triple $(H, O^p(H)Z, \chi)$ is fully ramified.

By the remarks preceeding the definition, an abelian p-group of type f.r. is reductive. In the following lemma, we extend slightly the class of reductive p-groups of type f.r. The author is unaware of an example of p-group of type f.r. which fails to have this property. We use the classification of groups with dihedral Sylow 2-subgroups in the case $p = 2$ of the following.

LEMMA 4.4. *Let Q be a p-group of order p^4 and of type f.r. Then Q is reductive.*

Proof. Suppose Q is a p-group of order p^4 which is of type f.r., but which is not reductive. Then, there exists a fully ramified triple (H, Z, χ) with a Sylow p-subgroup of H/Z isomorphic to Q, such that the triple $(H, O^p(H)Z, \chi)$ is not fully ramified. By Theorem 1.1, we may assume $Z = Z(H)$. Let $K = O^p(H)Z$ and let $R/Z \in \mathrm{Syl}_p(H/Z)$. Now $K < H$ as (H, K, χ) is not a fully ramified triple. By Corollary 3.6, $(R \cap K)/Z$ is a non-abelian p-group, and so has order $\geqq p^3$. This forces $|(R \cap K)/Z| = p^3$ and $|H: K| = p$. Let $C/Z = ((R \cap K)/Z)' = Z((R \cap K)/Z)$. Using Lemma 2.3 and Theorem 3.1, there are p special classes of $(R \cap K)/Z$. Suppose that some element, say g, of $C - Z$ is special. As gZ is central in $(R \cap K)/Z$, we get $[g, R \cap K] \subseteqq Z$. But g is special, and this means $[g, R \cap K] = 1$, so $g \in Z(R \cap K)$. It is clear that $1, g, g^2, \cdots, g^{p-1}$ represent the p distinct special classes in $(R \cap K)/Z$. Let $x \in R \cap K - C$. Then x is not special. However, $C_{(R \cap K)/Z}(xZ) = \langle xZ, C/Z \rangle \subseteqq C_{R \cap K}(x)/Z \subseteqq C_{(R \cap K)/Z}(xZ)$. This contradicts the fact that x is not special, and proves that the only special element of $(R \cap K)/Z$ which lies in C/Z is the identity.

Consider now the case that p is odd. Let $N = N_K(R \cap K)$ so that $\bar{N} = N/Z = N_{\bar{K}}(\overline{R \cap K})$. As p is odd, $(R \cap K)/Z$ is a regular p-group, being of class 2. It follows from the Hall-Wielandt theorem (see p. 447 of [8]), that $\bar{N}$ controls p-transfer, i.e., $O^p(\bar{K}) \cap \bar{N} = O^p(\bar{N})$. As $O^p(K/Z) = K/Z$, we will obtain a contradiction by proving that $O^p(\bar{N}) < \bar{N}$.

Let V denote the transfer homomorphism from N/Z into $(R \cap K)/C$. The map V is computed by

$$V(gZ) = \prod_{t \in T} (tgt^{-1})C\,, \qquad \text{for } g \in R \cap K\,,$$

where T is a right transversal for $R \cap K$ in N. (We used the fact that $R \cap K \lhd N$.) Now let g be any special element of $R \cap K$ with $g \notin Z$. Thus $g \notin C$, from above. For any $t \in N$, tgt^{-1} is $R \cap K$-conjugate to g, by the last part of Theorem 3.5. Thus, $tgt^{-1}C = gC$ for all $t \in$

T, and $V(gZ) = (gC)^{|T|} = g^{|N:R\cap K|}C$. As $|N: R\cap K|$ is prime to p, we have $\bar{g} \in \bar{N}$-ker$V$. This yields the contradiction $O^p(\bar{N}) < \bar{N}$, and we are done if p is odd.

Suppose now $p = 2$. Then $(R\cap K)/Z$ is non-abelian of order 8, and a nonidentity special elemet, say gZ, of $(R\cap K)/Z$ does not lie in C/Z.

Consider first the case that $(R\cap K)/Z$ is the quaternion group. Again, if $N = N_K(R\cap K)$, the element gZ can only be conjugate to $g^{-1}Z$ in N/Z. This implies $\bar{N}/C_{\bar{K}}(\overline{R\cap K})$ is a 2-group. Clearly, $N_{\bar{K}}(\bar{S})/C_{\bar{K}}(\bar{S})$ is a 2-group for all $\bar{S} < \overline{R\cap K}$, as $\bar{S}$ is cyclic. Thus, $\bar{K}$ has a normal 2-complement by Frobenius' theorem (see p. 253 of [6]). This contradicts $O^2(K)Z = K$, forcing $(R\cap K)/Z$ to be the dihedral group of order 8.

From the classification of groups with dihedral Sylow 2-subgroups, and the fact that $O^2(K/Z) = K/Z$, it follows that K/Z has a factor group isomorphic to Y, where $\mathrm{PSL}(2, p^n) \leqq Y \leqq \mathrm{P}\Gamma\mathrm{L}(2, p^n)$ for some odd prime power $p^n \neq 3$, or $Y = A_7$. From this, it follows that K/Z has exactly one chief factor isomorphic to the simple group S, where $S = \mathrm{PSL}(2, p^n)$, or $S = A_7$. Therefore, H/Z has a chief factor isomorphic to S. Let $Z \subseteqq V \subseteqq U \subseteqq H$, with V and U normal in H, and $U/V \cong S$. Define C by the equation: $C/V = C_{H/V}(U/V)$. Then $C \lhd H$, and $C \cap U = V$. Replacing U and V by UC and C respectively, and continuing this process, if necessary, we may assume $C = V$. The factor group H/V is isomorphic to a group X, which satisfies: $S \leqq X \leqq \mathrm{Aut}(S)$. Since $\mathrm{Aut}(\mathrm{PSL}(2, p^n)) \cong \mathrm{P}\Gamma L(2, p^n)$, Lemma 4.3 forces $S = A_7$. Now, $\mathrm{Aut}(A_7) \cong S_7$, the symmetric group on 7 letters, so that H/Z has a factor group which is either A_7 or S_7. Both of these groups contain a cyclic Sylow 5-subgroup of order 5, but neither group contains a subgroup of index 5. This contradiction to Theorem 2.5 completes the proof of the lemma.

We are now ready to give an application of the above. In [12], G is shown to be solvable if G is of type f.r., and $p^3 \nmid |G|$ for any prime p dividing $|G|$.

THEOREM 4.5. *Let G be a group of type f.r. Assume that G has an abelian Sylow p-subgroup for every prime p satisfying $p^6 \| |G|$. Then G is solvable.*

Proof. Let (H, Z, χ) be a fully ramified triple with $G \cong H/Z$. We proceed by induction on $|H/Z|$, the assertion being trivial if $H = Z$. Suppose that H/Z is not perfect. Then $O^p(H)Z < H$ for some prime p. The hypothesis of the theorem, together with the previous lemma, imply that a Sylow p-subgroup of H/Z is reductive of type f.r.

Therefore, $H/O^p(H)Z$ and $O^p(H)Z/Z$ are of type f.r., and we are done by induction.

Suppose then H/Z is perfect, and let K/Z be a maximal normal subgroup. Then H/K is a non-abelian simple group, and hence has even order by [4]. Also, a Sylow 2-subgroup S of H/K has order $\geqq 4$, as otherwise H/K would have a normal 2-complement.

Suppose S has order 4. Then $H/K \cong \mathrm{PSL}(2, q)$, where q is an odd prime power. But these simple groups are eliminated as possible homomorphic images of H/Z by Lemma 4.3. If $|S| = 8$, then apply Lemma 2.3 for the prime 2. Here $d^2s = 8$, and so $s = 2$ or 8. By the (c) part of that lemma, H/K has a subgroup of index s, which implies $s = 8$. However, this possibility is ruled out by Lemma 4.1. Thus, $|S| \geqq 16$.

If S is non-abelian, then the hypotheses of the theorem imply that $|S| = 16$, and S is isomorphic to a Sylow 2-subgroup of G. By Lemma 2.2, S is of type f.r. However, the only non-abelian groups of order 16 that occur as Sylow 2-subgroups of simple groups are dihedral and semi-dihedral. These types have cyclic self-centralizing subgroups, and by Corollary 3.2, S can have no such subgroup. Therefore, S must be abelian and $|S| \geqq 16$. By Walter's Theorem [14], $H/K \cong \mathrm{PSL}(2, |S|)$. This contradicts Lemma 4.3, and establishes the theorem.

It is possible to show that no known simple group can be a factor group of a group of type f.r. This strongly suggests that a group of type f.r. cannot be perfect. It would be desirable to have a proof of this fact, since it would represent a major step in proving that groups of type f.r. are solvable.

I wish to express my appreciation to Professor I. M. Isaacs for his invaluable advice during the preparation of this paper. In particular, he pointed out the short proof of Lemma 4.3. I originally had a much longer character theoretic proof that did not use the number theoretic result of Lemma 4.2. Finally, I would like to thank the referee for carefully reading the first draft, and for making a number of good suggestions for this revised version.

REFERENCES

1. G. D. Birkhoff and H. S. Vandiver, *On the integral divisors of* $a^n - b^n$, George David Birkhoff: Collected Mathematical Papers, vol. 3, Amer. Math. Soc., New York (1950), 145-152.

2. W. Feit, *The Current Situation in the Theory of Finite Simple Groups*, Actes, Congres intern. math., Tome 1 (1970), 55-93.

3. ———, *Characters of Finite Groups*, Benjamin, New York, 1967.

4. W. Feit, and J. G. Thompson, *Solvability of groups of odd order*, Pacific J. Math., **13** (1963), 775-1029.

5. P. X. Gallagher, *Group characters and normal Hall subgroups*, Nagoya Math. J., **21** (1962), 223-230.
6. D. Gorenstein, *Finite Groups*, Harper and Row, New York, 1968.
7. D. Gorenstein and J. Walter, *The Characterization of finite groups with dihedral Sylow* 2-*subgroups, I, II, III. J.* Algebra, **21** (1965), 85-151, 218-270, 334-395.
8. B. Huppert, *Endliche Gruppen I*, Springer-Verlag, Berlin-Heidelberg-New York, 1967.
9. I. M. Isaacs, *Characters of solvable and symplectic groups*, Amer. J. Math., vol. **XCV**, no. 3 (1973), 594-635.
10. F. R. DeMeyer, *Groups with an irreducible character of large degree are solvable*, Proc. Amer. Math. Soc., **25** (1970), 615-617.
11. ———, *Irreducible characters and solvability of finite groups*, Pacific J. Math., **41** (1972), 347-353.
12. F. R. DeMeyer and G. J. Janusz, *Finite groups with an irreducible representation of large degree*, Math. Zeit., **108** (1969), 145-153.
13. S. Matsumoto and N. Iwahori, *Several remarks on projective representations*, J. Fac. Sci. Univ. Tokyo, **10** (1964), 129-146.
14. J. H. Walter, *The characterization of finite groups with abelian Sylow* 2-*subgroups*, Ann. of Math., **89** (1969), 405-514.
15. H. Wielandt, *Finite Permutation Groups*, Academic Press, New York-London, 1964.

Received April 30, 1973 and in revised form July 30, 1974. This paper constitutes a partial fulfillment of the requirements for the Ph. D. degree at the University of Wisconsin, Madison.

MICHIGAN STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

OPERATOR VALUED ROOTS OF ABELIAN ANALYTIC FUNCTIONS

FRANK GILFEATHER

In this paper, all spaces are separable Hilbert spaces and all operators are bounded linear transformations. Questions involving the structure of an operator for which an analytic function of it is normal or which satisfies a polynomial with certain operator coefficients have been considered and studied separately. Using von Neumann's reduction theory, a unified approach to these and similar questions can be given. This method yields generalizations of the cases which has been previously investigated, including structure results for n-normal operators. Through reduction theory of von Neumann algebras, the study of structural questions for a particular orerator is reduced to the properties of the often simpler, reduced operators. In all of the applications presented in this paper, the reduced operators will simply involve algebraic operators.

In § 1, we introduce and study analytic functions $\psi(z)$, defined on a complex domain $\mathscr{D}$ and taking values in a commutative von Neumann algebra $\mathscr{A}$. Such a function will be called an *abelian analytic* function; and where there is any question, we shall specify the algebra $\mathscr{A}$. Using the direct integral decomposition of $\mathscr{A}$ into factors, we obtain the decomposition of ψ into a normal family of scalar valued analytic functions on $\mathscr{D}$ indexed by a real variable. The main results in this section will be to show that the zeros of the scalar valued analytic functions can be chosen to be Borel functions of the real variable. We shall restrict our attention to a class of abelian analytic functions, called *locally nonzero*, so that each scalar valued analytic function in the corresponding normal family has no subdomain on which it is identically zero.

An operator T in the commutant $\mathscr{A}'$ of $\mathscr{A}$ is called a *root* of an abelian analytic function ψ, if $\sigma(T)$, the spectrum of T, is contained in $\mathscr{D}$ and $\psi(T)=0$ where $\psi(T)$ is to be defined in the usual B^* algebraic manner or in an equivalent way using the direct integral decomposition of $\mathscr{A}$ into factors. Section 2 develops the structure for roots of locally nonzero abelian analytic functions. The main result, Theorem 2.1, states that the root of an abelian analytic function is "piecewise" a spectral operator of finite type. The structure theorem shows that roots of abelian analytic functions have hyperinvariant subspaces or are scalar multiples of the identity.

The remaining two sections of this paper are essentially applications of the structure theorem for roots of abelian analytic functions to several classes of operators and the further use of reduction theory in their study. In § 3, our investigation leads to theorems concerning solutions of

$$(*) \qquad f(A) = N ,$$

where f is an analytic function on a domain containing $\sigma(A)$ and N is a normal operator. The use of reduction theory in the study of $(*)$ was introduced by the author in [9], and solutions of $(*)$ have been previously studied by many authors with various restrictions on f, A, or N. The most complete investigation of the solutions of $(*)$ has been done by C. Apostol in the setting of the theory of generalized spectral operators, however, his results are of a quite different nature from those given here [1]. If we set $\psi(z) = f(z) - N$, then ψ becomes an analytic abelian function and a solution A of $(*)$ is just a root of ψ. Hence, we may apply our methods and results; and in doing so, we are able to obtain two structure theorems for A. If there is no subdomain of on which f is identically zero, then f will be called *locally nonzero*. We show that whenever A is a solution of $(*)$ where f' is locally nonzero and, of course, where $\sigma(A)$ is contained in $\mathscr{D}$, then it follows that A is the direct sum of two operators; the first, A_1, which is algebraic and the second, A_2, which is "piecewise" similar to a normal operator. In the latter situation, the summand A_2 and the corresponding normal operator have the same spectrum. Under certain conditions, we may conclude that the solution A of $(*)$ is "piecewise" similar to a normal solution N_0 of $(*)$ and that A and N_0 have the same spectrum. We also give a decomposition of certain operators satisfying $(*)$ into direct summands each of which satisfy certain operator valued polynomials. Thus, we are able to generalize results obtained previously by C. Apostol, H. Radjavi, and P. Rosenthal and others [1, 10–13, 15, 16, 18].

The structure of operators satisfying certain operator valued polynomials is studied in § 4. An important class of such operators are the n-normal operators ($n \times n$ matrices of commuting normal operators). An n-normal operator A satisfies a normal valued polynomial of degree n by virtue of the Hamilton-Cayley Theorem; and moreover, the coefficients of the polynomial are in the center of the von Neumann algebra generated by A. N. Dunford has studied these operators primarily from the viewpoint of when they were spectral operators [6]. Since operators in a type I_n von Neumann algebra are also n-normal, they naturally occur in the study of operator algebras. Also the structure and existence of hyperinvariant subspaces for certain n-normal operators have been investigated by various authors [3-5, 12, 13, 15]. We may then apply the theorems in § 1 to n-normal operators showing that they are "piecewise" similar to spectral operators and obtaining conditions for similarity which are compatible to those given in [6]. Whenever an operator A satisfies a monic polynomial of degree less or equal to two with coefficients in the center of the von Neumann algebra generated by A, we can use reduction theory to obtain a complete structure theorem for it. This result will generalize results in [3, 16] and is closely connected to the work of A. Brown on binormal operators (2-normal) [2, 11].

Finally in § 4, we give some sufficient conditions for a root of an abelian analytic function to be a spectral operator and, more specifically, a scalar type (similar to a normal operator) operator. For the n-normal case, our results complement those given by N. Dunford [6]. Also, we give some examples based on an example introduced by J. Stampfli of a 2-normal operator whose square is normal yet it is not similar to a normal square root of its square [18].

The essential component of von Neumann reduction theory is the concept of the direct integral decomposition of an algebra. For the details of the direct integral decomposition of a von Neumann algebra, we refer to [17]; however, we shall introduce some basic notations and results here. Let μ be the completion of a finite positive regular measure defined on the Borel sets of a separable metric space Λ, and let e_n, $1 \leqq n \leqq \infty$ be a collection of disjoint Borel sets of Λ with union Λ. Let $H_1 \subseteqq H_2 \subseteqq \cdots \subseteqq H_\infty$ be a sequence of Hilbert spaces, with H_n having dimension n and H_∞ being separable. By

$$H = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$$

we shall denote the space of weakly μ-measurable functions from Λ into H_∞ such that $f(\lambda) \in H_n$, if $\lambda \in e_n$, and $\int_\Lambda \|f(\lambda)\|^2 \mu(d\lambda) < \infty$. The space H is a Hilbert space, and we shall denote the element $f \in H$ determined by the vector valued function $f(\lambda)$ as $\int_\Lambda \oplus f(\lambda)\mu(d\lambda)$.

An operator A on H is said to be *decomposable* if there exists a μ-measurable operator valued function $A(\lambda)$ so that $(Af)(\lambda) = A(\lambda)f(\lambda)$ for $f \in H$. The operator A is denoted by

$$A = \int_\Lambda \oplus A(\lambda)\mu(d\lambda) .$$

Furthermore, every von Neumann algebra $\mathscr{A}$ on a separable space is spatially isomorphic to an algebra of decomposable operators on a direct integral of Hilbert spaces, such that the von Neumann algebra

$\mathscr{A}(\lambda)$ generated by $\{A(\lambda)\}$, where $A \in \mathscr{A}$, is a factor μ-a.e. Finally, we use the fact that if $A = \int_\Lambda \oplus A(\lambda) d\mu(\lambda)$ generates $\mathscr{A}$, then $A(\lambda)$ generates the von Neumann algebra $\mathscr{A}(\lambda)$ μ-a.e. Whenever in our use of this decomposition, there is no confusion over the space Λ, we shall suppress it.

If A is an operator, we shall denote by $R(A)$, $R(A)'$, and $Z(A)$, respectively, the von Neumann algebra generated by A, the commutant of $R(A)$ and the center of $R(A)$. N. Suzuki has introduced the notion of a primary operator. One calls an operator A *primary*, in case $R(A)$ is a factor; i.e., $Z(A)$ is just the scalar multiples of the identity. Let A be defined on a separable Hilbert space and let $H = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$ be the direct integral decomposition of H related to $R(A)$ for which the algebra $R(A)(\lambda)$ is a factor μ-a.e., then this decomposition is unique in the sense of [17; I. 6]. Thus, the operator A is decomposed as $A = \int_\Lambda \oplus A(\lambda)\mu(d\lambda)$, where $A(\lambda)$ is primarily μ-a.e., and we shall refer to this particular decomposition as the *primary decomposition* of A. We shall call a projection *central* for T if it is in $Z(T)$. Finally, we shall let $R(z; A)$ denote $(zI - A)^{-1}$.

1. **Abelian analytic functions.** In this section, we shall develop the notion of an abelian analytic function and investigate its properties. Let $\mathscr{A}$ be an abelian von Neumann algebra and $\psi(z)$, an $\mathscr{A}$ valued analytic function on a domain $\mathscr{D}$ in the complex plane, then ψ is called an *abelian analytic function* with domain $\mathscr{D}$. For the usual facts about B^* valued analytic functions, we refer to [7; III, 14].

Given an abelian von Neumann algebra $\mathscr{A}$, we may decompose it into a direct integral of factors. That is, H is unitary equivalent to a direct integral of Hilbert spaces $\int_\Lambda \oplus H(\lambda)\mu(d\lambda)$, and this induces a spatial isomorphism between $\mathscr{A}$ and the diagonal operators on $\int_\Lambda \oplus H(\lambda)\mu(d\lambda)$. Thus, $H = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$; and for $A \in \mathscr{A}$, there is a unique $g \in L_\infty(\Lambda, \mu)$, so that $A = \int_\Lambda \oplus g(\lambda)I(\lambda)\mu(d\lambda)$, where $I(\lambda)$ is the identity operator on $H(\lambda)$ [17; I, 2.6].

Let ψ be an abelian analytic function and $\mathscr{A}$ the corresponding von Neumann algebra with $\int_\Lambda \oplus H(\lambda)\mu(d\lambda)$ the decomposition of H given above. Since $\psi(z)$ belongs to $\mathscr{A}$ for each z, we have

$$\psi(z) = \int_\Lambda \oplus \psi(z, \lambda)I(\lambda)\mu(d\lambda) \tag{1.1}$$

where $(\psi z, \lambda)$ corresponds via the isomorphism mentioned above to $\psi(z)$. We first give the relationship between the analyticity of $\psi(z)$ and that of $\psi(z, \lambda)$.

PROPOSITION 1.1. *If $\psi(z)$ is an abelian analytic function with domain $\mathscr{D}$, then $\psi(z, \lambda)$, given by* (1.1), *is analytic on $\mathscr{D}$ for almost all λ and $\|\psi(z, \lambda)\|_\infty$ is uniformly bounded on compact subsets of $\mathscr{D}$. Conversely, let $\psi(z, \lambda)$ be a family of functions defined on $\mathscr{D} \times \Lambda$, where $\mathscr{D}$ is a complex domain. If $\psi(z, \lambda)$ is analytic in z for almost all λ on the domain $\mathscr{D}$ and if $\psi(z, \lambda) \in L_\infty(\Lambda, \mu)$ with $\|\psi(z, \cdot)\|_\infty$ uniformly bounded on compact subsets of $\mathscr{D}$, then $\psi(z)$, given by* (1.1), *is an abelian analytic function with domain $\mathscr{D}$.*

Proof. We assume that ψ is an abelian analytic function on $\mathscr{D}$ and that $z_0 \in \mathscr{D}$. The series $\psi(z) = \sum N_n((z - z_0)^n/n!)$ converges with N_n given by Cauchy's formula is in $\mathscr{A}$ and z is in some neighborhood S_0 of z_0. If $N_n = \int_\Lambda \bigoplus g_n(\lambda)I(\lambda)\mu(d\lambda)$, then for z fixed in S_0, $\psi(z)(\lambda) = \sum_n g_n(\lambda)((z - z_0)^n/n!)I(\lambda)$ for almost all λ. Hence, by the convergence properties of power series, we may conclude that $\psi(z, \lambda)$ is analytic in a neighborhood of z_0 and hence on $\mathscr{D}\mu$ a.e.

Conversely, we assume that $\psi(z, \lambda)$ belongs to $L_\infty(\Lambda, \mu)$ and $\|\psi(z, \cdot)\|_\infty$ is bounded for z in compact subsets of $\mathscr{D}$. For z_0 in $\mathscr{D}$, let $\psi(z, \lambda) = \sum_n g_n(\lambda)((z - z_0)^n/n!)$ be the power series expansion in a neighborhood S_0 of z_0. Since the functions $\{g_n\}$ are given by Cauchy's formula and $\psi(z, \cdot)$ is measurable, we conclude that $\{g_n\}$ are measurable. We are done if we can show that $g_n \in L_\infty(\Lambda, \mu)$. That, however, also follows from Cauchy's formula and using the hypothesis that $\|\psi(z, \cdot)\|_\infty$ are uniformly bounded on compact subsets of $\mathscr{D}$.

REMARK. If it is the case that $\psi(z, \lambda)$ is independent of λ, then the proposition is trivial. For example, if $\psi(z) = f(z)I$, then $\psi(z)(\lambda) = f(z)I(\lambda)$ almost everywhere. In order to save the repetitiousness of deleting a set of measure zero from every argument, whenever $\psi(z)$ is an abelian analytic function on a domain $\mathscr{D}$, we will always assume that $\psi(z, \lambda)$ is analytic on a domain containing $\mathscr{D}$.

The main result in this section will show that the zeros of $\psi(z, \lambda)$ can be chosen in a μ measurable way. Such a result constitutes a generalization of the key lemmas in the study of n-normal operators by N. Dunford [6; XV, 10] and is also related to the Theorem 1 in [5].

For this problem to be well defined, we must make a restriction so that $\psi(z, \lambda)$ is not identically zero on some subdomain of $\mathscr{D}$. We shall call an abelian analytic function ψ *locally nonzero* if for every convergent sequence $\{z_n\}$ in $\mathscr{D}$ with $z_n \to z_0$ in $\mathscr{D}$ then $\bigcap_n \mathscr{N}(\psi(z_n)) = \{0\}$ ($\mathscr{N}(A)$ denotes the nullspace of the operator A). For scalar valued functions, this is the usual definition of locally nonzero. To see this, we just let H be one dimensional, then $\psi(z)$ is just a scalar valued function and $\mathscr{N}(\psi(z_n)) \neq \{0\}$ means that $\psi(z_n) = 0$. The following lemmas establish the relationship between $\psi(z)$ and $\psi(z, \lambda)$ with respect to this property.

LEMMA 1.2. *An abelian analytic function ψ is locally nonzero if and only if $\psi(\cdot, \lambda)$ is locally nonzero for almost all λ.*

Proof. First assume that ψ is not locally nonzero. That is, there exists a nonzero $x \in H$ and a sequence $\{z_n\}$ in $\mathscr{D}$ converging to z_0 in $\mathscr{D}$, so that $\psi(z_n)x = 0$. If $E_1 = \{\lambda \in \Lambda \mid x(\lambda) \neq 0\}$ and $E_2 = \bigcup_n \{\lambda \mid \psi(z_n, \lambda)x(\lambda) \neq 0\}$, then $E = E_1 \backslash E_2$ is a set of positive measure on which $\psi(\cdot, \lambda)$ is not locally nonzero.

Conversely, if $\psi(\cdot, \lambda)$ is not locally nonzero for λ in a set E of positive measure, then we can show that $\psi(z)$ is not locally nonzero. For this, we let $\psi(z, \lambda)$ be zero on the subdomain $\mathscr{D}_\lambda$ if $\lambda \in E$. Since the domain of analyticity of $\psi(z, \lambda)$ contains $\mathscr{D}$, each $\mathscr{D}_\lambda$ contains one of the subdomains of $\mathscr{D}$; and thus, there is a subset F of E with positive measure so that $\bigcap_{\lambda \in F} \mathscr{D}_\lambda \supset \mathscr{D}_0$, a subdomain of $\mathscr{D}$. Therefore, $\psi(z, \lambda) = 0$ for $\lambda \in F$ and $z \in \mathscr{D}_0$. Let $z_n \to z_0$ in $\mathscr{D}_0$ and $x \in H$ so that $\{\lambda \mid x(\lambda) \neq 0\} = F$, then $x \in \bigcap \mathscr{N}(\psi(z_n))$. This completes the proof of this lemma.

Let a locally nonzero abelian analytic function ψ be decomposed as in (1.1). The following theorem shows that the zeros of the functions $\psi(\cdot, \lambda)$ restricted to a compact subset of $\mathscr{D}$ can be made measurable.

THEOREM 1.3. *Let $\psi(z, \lambda)$ be given by* (1.1) *with domain $\mathscr{D} \times \Lambda$. If D is a bounded subdomain of $\mathscr{D}$ with $\bar{D} \subset \mathscr{D}$, then there exist disjoint Borel sets E_i, $i = 0, 1, \cdots$ with the measure of $\Lambda \backslash \bigcup_{i=0}^{\infty} E_i$ zero and for $\lambda \in E_j$, the analytic function $\psi(\cdot, \lambda)$ has exactly j zeros counted to their multiplicities in D. Moreover, there exist Borel functions $\{r_i\}_{i=1}^{\infty}$ so that if $\lambda \in E_j$, then $r_i(\lambda)$ $1 \leqq i \leqq j$ are those zeros.*

Proof. Since the number of zeros of an analytic function inside a desk is given by an integral formula, it is easy to see that if $n(\lambda)$ denotes the number of zeros counted to multiplicity of $\psi(z, \lambda)$ contained in D, then $S_k = \{\lambda \mid n(\lambda) \geqq k\}$ is Borel subset of Λ. Hence, if we may set $E_k = S_k \backslash S_{k+1}$, then E_k is a Borel set; and it follows that $\Lambda \backslash \bigcup_{i=0}^{\infty} E_i$ has measure zero. We shall fix n and define r_i on E_n; and this will be clearly sufficient to complete the proof.

Henceforth, we are assuming that $E_n = \Lambda$, $1 \leqq n < \infty$, and, the mapping ψ on $D \times \Lambda$ is a Borel measurable map from the product space into the complex numbers. The projection of $\{(z, \lambda) \mid \psi(z, \lambda) = 0\}$ onto Λ is Λ (a.e.) and by the Principle of Measurable Choice one finds a Borel function $r_1: \Lambda \to D$ so that $(r_1(\lambda), \lambda)$ is in the null space of ψ, that is, $\psi(r_1(\lambda), \lambda) = 0$ for all $\lambda \in \Lambda$ [17; I, 4.7]. Consider now the function $\psi(z, \lambda)(z - r_1(\lambda))^{-1} \equiv \phi(z, \lambda)$. By judiciously applying Schwartz's lemma on the modulus of a complex valued function one can show that $\phi(z, \lambda)$ is uniformly bounded in λ on compact subsets in $\mathscr{D}$. Thus by Proposition 1.1 we conclude that ϕ is again an abelian analytic function. Moreover, it is clear that $\phi(\cdot, \lambda)$ has $n - 1$ zeros in D counted to their multiplicity almost everywhere. The proposition now follows with repeated application of the above argument.

The motivation for introducing abelian analytic functions is to study the structure of certain of their operator roots; and in doing so, unify several previous investigations. Whenever $\psi(z)$ is a polynomial with commuting normal coefficients and T is an operator commuting with those coefficients, then $\psi(T)$ has an obvious definition. The definition of $\psi(T)$ we shall now give will be compatable with this usual definition when ψ is a polynomial.

Let ψ be an abelian analytic function on a domain $\mathscr{D}$ with values in the von Neumann algebra $\mathscr{A}$. If $H = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$ is the direct integral decomposition of H corresponding to the decomposition of $\mathscr{A}$ into factors; and if $T \in \mathscr{A}'$, then T is a decomposable operator. That is, T is represented as $T = \int_\Lambda \oplus T(\lambda)\mu(d\lambda)$ where $T(\lambda)$ is an operator on H_λ. Now let $T \in \mathscr{A}'$ and $\sigma(T) \subset \mathscr{D}$. Since $\sigma(T(\lambda)) \subset \sigma(T)$, almost everywhere, the operator $\psi(T(\lambda), \lambda)$ is well defined by the usual functional calculus [7, 11].

To complete the definition of $\psi(T)$, let Γ be an admissible curve for $\psi(T)$ in $\mathscr{D}$. Thus $\psi(T(\lambda), \lambda) = (2\pi i)^{-1}\int_\Gamma R(z; T(\lambda))\psi(z, \lambda)dz$ and $\psi(T(\lambda), \lambda)$ is clearly a measurable operator function. If we can show that it is essentially bounded, then we may define $\psi(T)$ to be the decomposable operator given by $\psi(T)(\lambda) = \psi(T(\lambda), \lambda)$. Now let z_n be a dense set on Γ. Since almost everywhere $\| R(z_n; T(\lambda)) \| \leqq \| R(z_n; T) \|$, we may eliminate a set E of measure zero and have on the complement of E, $\| R(z; T(\lambda)) \| \leqq \| R(z; T) \|$ for all $z \in \Gamma$. By Proposition 1.1, $\| \psi(z, \lambda) \|_\infty \leqq M < \infty$ for all z on Γ and thus $\| \psi(z, \lambda)R(z; T(\lambda)) \| \leqq M$ on the complement of a set of measure zero and for all $z \in \Gamma$. Hence if $k = (2\pi i)^{-1}\int_\Gamma | dz |$, we have that $\| \psi(T(\lambda), \lambda) \| \leqq Mk$, for almost all λ and therefore $\psi(T)$ is a bounded operator on H if it is the decomposable operator defined by $\psi(T)(\lambda) = \psi(T(\lambda), \lambda)$. It is clear that $\psi(T) \in \mathscr{A}'$ since $\psi(T(\lambda), \lambda) \in \mathscr{A}'(\lambda)'$ for each λ. We conclude our remarks on the definition of $\psi(T)$ be noting that we have actually shown that $\psi(T)$ satisfies the conditions of a Fubini type theorem. Alternately $\psi(T)$ may be defined by usual B^*

algebraic techniques as

$$\psi(T) = (2\pi i)^{-1} \int_\Gamma \psi(z) R(z; T) dz \,, \tag{1.2}$$

where $\psi(z)$ is a $\mathscr{A}$ valued analytic function defined on a domain containing $\sigma(T)$ and with $T \in \mathscr{A}'$ and the integral converging in the norm. We may conclude that

$$\begin{aligned} \psi(T) &= \int_\Lambda \oplus (2\pi i)^{-1} \int_\Gamma \psi(z, \lambda) R(z; T(\lambda)) dz \mu(d\lambda) \\ &= (2\pi i)^{-1} \int_\Gamma \int_\Lambda \oplus \psi(z, \lambda) R(z; T(\lambda)) \mu(d\lambda) dz \,, \end{aligned} \tag{1.3}$$

that is, $\psi(T)(\lambda) = \psi(T(\lambda), \lambda)$ almost everywhere.

In the two applications of this theory, we wish to pursue we note that $\psi(T)$ coincides with previously understood definitions. If $\psi(z)$ is the polynomial $\psi(z) = N_n z^n + \cdots + N_1 z + N_0$, with coefficients N_i in an abelian von Neumann algebra, then by (1.3) we see that $\psi(T)$ is just $N_n T^n + \cdots + N_1 T + N_0$. On the other hand, if $\psi(z)$ is a scalar valued analytic function, then by (1.3) we have established that $\psi(T)$ is the usual operator determined by the standard functional calculus [7; VII]. Moreover, in this latter case, the fact that the definition above for $\psi(T)$ and the usual one given by contour integration are the same as a special case of Theorem 1 in [11].

2. Roots of abelian analytic functions.

We shall call T a *root* of the abelian analytic function ψ if $\psi(T) = 0$ where $\psi(T)$ was defined in §1. If ψ has domain of analyticity $\mathscr{D}$ and takes values in the von Neumann algebra $\mathscr{A}$, then, by the definition of $\psi(T)$, we are assuming that $T \in \mathscr{A}'$ and that $\sigma(T) \subset \mathscr{A}$. In this section, we give a structure theorem for all roots of an abelian analytic function and several applications.

We shall state and prove the main theorem after which we shall restate it using the language of spectral operators.

THEOREM 2.1. *Let ψ be a locally nonzero abelian analytic function on $\mathscr{D}$ taking values in the von Neumann algebra $\mathscr{A}$ and let T be a root of ψ. There exists a normal operator S in $\mathscr{A}'$ and a sequence of mutually orthogonal projections $\{P_n\}$ in $\mathscr{A}$ with $I = \Sigma P_n$ so that TP_n is similar to $(S + L_n)P_n$, where L_n is a nilpotent operator $SL_n = L_n S$ and both L_n and the operator which induces the similarity are in $\mathscr{A}'$.*

Proof. In assuming that T is a root of $\psi(z)$ we have that $T \in \mathscr{A}'$. We shall give the structure of T by first decomposing T

into a direct integral of operators via the direct integral of decomposition of $\mathscr{A}$ and then determining the structure of each reduced operator in the decomposition of T.

Let $H = \int_{\Lambda} \oplus H(\lambda)\mu(d\lambda)$ be the decomposition of H corresponding to the primary decomposition of $\mathscr{A}$. Since $T \in \mathscr{A}'$, we may decompose T as $T = \int_{\Lambda} \oplus T(\lambda)\mu(d\lambda)$. Furthermore, by (1.3) if $\psi(T) = 0$, then almost everywhere $\psi(T(\lambda), \lambda) = 0$, where $\psi(z, \lambda)$ is an analytic function in a neighborhood of $\sigma(T(\lambda))$. By Lemma 1.2, the analytic function $\psi(z, \lambda)$ is locally nonzero in $\mathscr{D}$. In fact, by Theorem 1.3, there are disjoint Borel sets E_i, $i = 0, 1, \cdots$, where $\Lambda \backslash \bigcup_{i=0}^{n=\infty} E_i$ has measure zero, and Borel functions $r_i(\lambda)$, $i = 1, \cdots$, so that if $\lambda \in E_k$ then $r_1(\lambda), \cdots, r_k(\lambda)$ are the zeros of $\psi(z, \lambda)$ in $\sigma(T)$ counted to their multiplicities. Since $\{E_i\}$ determine mutually orthogonal projections in $\mathscr{A}$, we may assume without loss of generality that for almost all λ in Λ, $\psi(z, \lambda)$ has k roots in $\sigma(T)$ counted their multiplicities and since $\psi(A(\lambda), \lambda) = 0$ a.e., that $\mu(E_0) = 0$.

It follows from the measurability of $\{r_i(\lambda)\}_{i=1}^{k}$, that the distinct roots of $\psi(z, \lambda)$ as well as their multiplicities can be chosen measurably. Thus we let $z_1(\lambda), \cdots, z_n(\lambda)$ be the distinct roots of $\psi(z, \lambda)$ in $\sigma(T)$ for λ in the Borel set $F_n = \{\lambda \mid \psi(z, \lambda)$ has n distinct roots in $\sigma(T)\}$ and let the multiplicity of $z_i(\lambda)$ be $k_i(\lambda)$. Define $\delta(\lambda) = \min_{i \neq j} | z_i(\lambda) - z_j(\lambda) |$, which is also a Borel function. For each i, we determine the algebraic projections

$$E_i(\lambda) = (2\pi i)^{-1} \int_{\Gamma_i} R(z; T(\lambda))dz \,, \tag{2.1}$$

where Γ_i is the circle centered at $z_i(\lambda)$ of radius $\delta(\lambda)/2$. Since $T(\lambda)$ is an algebraic operator with $\sigma(T(\lambda)) \subset \{z_i(\lambda)\}_{i=1}^{k}$ we have

$$T(\lambda)/E_i(\lambda)H(\lambda) = [z_i(\lambda)I(\lambda) + N_i(\lambda)]/E_i(\lambda)H(\lambda) \,, \tag{2.2}$$

where $N_i(\lambda)$ is nilpotent of order $k_i(\lambda)$. Setting

$$R(\lambda) = \left(\sum_{i=1}^{n} E_i(\lambda)E_i(\lambda)^*\right)^{1/2} , \tag{2.3}$$

then $R(\lambda)$ is invertible on $H(\lambda)$, $R(\lambda)E_i(\lambda)R(\lambda)^{-1} = P_i(\lambda)$ are mutually orthogonal self-adjoint projections with $I(\lambda) = \sum_{i=1}^{n} P_i(\lambda)$, for $\lambda \in F_n$ and

$$R(\lambda)T(\lambda)R(\lambda)^{-1} = \sum_{i=1}^{n} z_i(\lambda)P_i(\lambda) + L(\lambda) \,, \tag{2.4}$$

where $L(\lambda)^k = 0$ and $P_i(\lambda)L(\lambda) = L(\lambda)P_i(\lambda)$ for each i. The form (2.4) is what we desired as our structure theorem. The only drawback to integrating the expression (2.4) over F_n and then taking direct sums is the boundedness of the projections $E_i(\lambda)$ (the boundedness of $R(\lambda)$ and $R(\lambda)^{-1}$ only depend on n and the boundedness of the $E_i(\lambda)$).

It is *not* the case that the projections $E_i(\lambda)$ are in general bounded independent of λ and thus the structure theorem is given in terms of "piecewise" similarity. Let

$$G_m = \{\lambda \in F_n \mid \| E_i(\lambda) \| \leqq m,\ i = 1, 2, \cdots, n\} ,$$

and $g_m(\lambda)$ the characteristic function of the Borel set G_m. Let Q_m be the corresponding projections in given by

$$Q_m = \int_A \oplus\, g_m(\lambda) I(\lambda) \mu(d\lambda)$$

and set $H_m = Q_m H$ and $T_m = T/H_m$. Then $R(\lambda)$, $R(\lambda)^{-1}$ and $L(\lambda)$ are uniformly bounded for $\lambda \in G_m$ and hence we may define

$$R_m = (I - Q_m) + \int_A \oplus\, g_m(\lambda) R(\lambda) \mu(d\lambda) ,$$

$$N_m = \int_A \oplus\, g_m(\lambda) N(\lambda) \mu(d\lambda)$$

and

$$S = \int_A \oplus\, (\sum z_i(\lambda) P_i(\lambda)) \mu(d\lambda) ,$$

where the summation under the integral in S is taken over the number of distinct roots of $\psi(z, \lambda)$ in $\sigma(T)$, for example, n for λ in F_n. Considering all the special conditions on the operators, we have

$$R_m T R_m^{-1} / Q_m H = [S + R_m N_m R_m^{-1}] / Q_m H$$

or if we set $L_m = R_m N_m R_m^{-1}$, then

$$R_m T R_m^{-1} Q_m = (S + L_m) Q_m .$$

Finally, it is clear that $S \in \mathscr{A}'$ is a normal operator, $L_m \in \mathscr{A}'$ and $SL_m = L_m S$.

REMARK. Recently, decomposable operators on a direct integral of Hilbert spaces have been investigated by E. A. Azoff [2]. He has shown that in general, the spectrum of a decomposable operator is measurable. The results in § 1 and this section imply this result for roots of abelian analytic functions, so that Azoff's work is related to certain results in these sections.

The following proposition will give a connection between the spectrum of T and that of the corresponding normal operator S.

This will be useful in the next section where we discuss special abelian analytic functions.

PROPOSITION 2.2. *If T and S are as in Theorem 2.1, then the spectrum of S intersects every connected component of $\sigma(T)$.*

Proof. Let $\mathscr{D}_1$ be a subdomain of $\mathscr{D}$ containing a connected component of $\sigma(T)$ and let $\Gamma = \partial\mathscr{D}_1$ be an admissible curve which also is contained in $\mathscr{D}$. Let $E = (2\pi i)^{-1}\int_\Gamma R(z; T)dz$, then $E \in \mathscr{A}'$ and $E = (2\pi i)^{-1}\int_\Lambda \oplus \int_\Gamma R(z; T(\lambda))dz\mu(d\lambda) = \int_\Lambda \oplus E(\lambda)\mu(d\lambda)$ [11]. Clearly if $\int_\Gamma R(z; T(\lambda))d(z) = 0$ almost everywhere, then $E = 0$. Thus there is a Borel set F so that $E(\lambda) \neq 0$ for $\lambda \in F$ and $\mu(F) \neq 0$. Hence, the set $G = \{\lambda \in F \mid \sigma(T(\lambda)) \cap \mathscr{D}_1 \neq \phi\}$ and consequently for some i the set $G_i = \{\lambda \in F \mid r_i(\lambda) \cap \mathscr{D}_1 \neq \phi\}$ has positive measure. Therefore, $\sigma(S) \cap \mathscr{D}_1$ contains the essential range of z_i restricted to G_i.

REMARK 1. The operator S in the theorem is also a root of $\psi(z)$ as well as each of the operators $S + L_m$. Later we shall see that in special cases where the nilpotent part does not appear, we will then have all roots "piecewise" similar to normal roots.

REMARK 2. The proof of the theorem can be used to construct the normal as well as the nonnormal roots of $\psi(z)$. Thus we establish the fact that certain abelian analytic functions have roots. This is related to work in [4] and [12].

As we stated before the theorem, we may put this result in the context of the theory of spectral operators on a Hilbert space H. Our result in this setting then reads: *Let T be a root of a locally nonzero abelian analytic function. There exists mutually orthogonal projections P_n in $R(T)'$ so that $I = \sum P_n$ and T/P_nH is a spectral operator of finite type.*

Before giving an application of this result, we wish to remark on the roots of abelian polynomial functions vis-a-vis abelian analytic functions. If f is a locally nonzero complex valued analytic function defined on a domain containing $\sigma(T)$, then $f(T) = 0$ implies $p(T) = 0$ for some complex valued polynomial. An analogous result holds for the operator valued analytic functions.

PROPOSITION 2.3. *If T is the root of an abelian analytic function with values in $\mathscr{A}$, then T is the direct sum of roots of monic polynomials with coefficients in $\mathscr{A}$.*

Proof. This follows from the structure theorem if we let $p_N(z, \lambda) = \prod_{i=1}^{N}(z - r_i(\lambda))$, on the set where N is the number of roots of $\psi(z, \lambda)$ in $\sigma(T)$ counted to their multiplicities and the Borel functions $r_i(\lambda)$ are the functions given in Theorem 1.3. Thus by equation (2.1) it follows that $p_N(T_N) = 0$ where T_N is defined in the obvious way.

We might point out the importance that a root T of $\psi(z)$ belong to $\mathscr{A}'$ aside from the fact that the proof of Theorem 2.1 would otherwise fail. In case T is not in $\mathscr{A}'$ essentially nothing can be determined, at least along the lines of our results. Let H be a Hilbert space with orthonormal basis $\{e_n\}$, $n = 0, \pm 1, \pm 2, \cdots$. If U is the bilateral shift of H with respect to this basis and V is the unilateral shift on $\{e_n\}$, $n = 0, 1, 2, \cdots$, and 0 on $\{e_n\}$, $0 = -1, -2, \cdots$, then V satisfies the abelian polynomial $z^2 - Uz = \psi(z)$.

As a corollary to our main theorem, we shall show that roots of abelian analytic functions have hyperinvariant subspaces or are multiples of the identity operator. We shall call a closed subspace M in H *hyperinvariant* for an unbounded operator A, if $\overline{M \cap \mathscr{D}(A)} = M$ ($\mathscr{D}(A)$ is the domain of A and will be taken to be dense), and M is invariant under every bounded operator B which commuted with A in the following sense: $B^{-1}\mathscr{D}(A) \cap \mathscr{D}(A)$ is dense and $AB = BA$ on $B^{-1}\mathscr{D}(A) \cap \mathscr{D}(A)$.

Let A be an unbounded operator with dense domain and T be a bounded operator. We say T is *quasisimilar* to A, if there exist bounded one-to-one operators X and Y, with dense ranges, so that $XH \subset \mathscr{D}(A)$, $AX = XT$, and $TY = YA$ on $\mathscr{D}(A)$. The following lemma extends to the unbounded case a useful tool for proving the existence of hyperinvariant subspaces.

LEMMA 2.4. *Let T be quasisimilar to an unbounded operator A. If A has nontrivial hyperinvariant subspaces, then T has nontrivial hyperinvariant subspaces.*

Proof. The proof is similar to the usual proof for the bounded case [13; Theorem 2.1].

Combining this lemma and Theorem 2.1, we have the following result, the proof of which is straightforward and it omitted.

THEOREM 2.5. *Let T be a root of an abelian analytic function. If T is not a multiple of the identity, then T has nontrivial hyperinvariant subspaces.*

3. Solutions to $f(T)$ normal. In this section we develop the structure of the operator roots T of the equation

$$f(T) = N\,, \tag{3.1}$$

where $f(z)$ is a complex valued analytic function on a domain $\mathscr{D} \supset \sigma(T)$ and N is a normal operator. Certain results are known as was mentioned in the introduction; in particular, (3.1) has been studied with various restrictions on f. If we set $\psi(z) = f(z) - N$, then ψ is a locally nonzero abelian analytic function on a domain $\mathscr{D}$ if and only if f' is locally nonzero on $\mathscr{D}$ (f' is locally nonzero is also expressed as f is *locally nonconstant*). Thus we may apply the results of the previous sections to solutions of equation (3.1) whenever f is locally nonconstant. The von Neumann algebra generated by $\{\psi(z) \mid z \in \mathscr{D}\}$ is abelian and in fact, just $R(N)$, the von Neumann algebra generated by N and I. Hence, if T has spectrum in $\mathscr{D}$ and $f(T) = N$, then T commutes with N, so by the Fuglede theorem $T \in R(N)'$ and hence T satisfies the condition in the hypothesis of Theorem 2.1. Moreover, matters are even made simpler in this section if when we apply our results we let $\mathscr{A} = Z(T)$ as then we are utilizing the primary decomposition for T. Thus in this section, unless otherwise stated, $\mathscr{A} = Z(T)$ where T is a solution of (3.1).

To aid in our characterization, we shall use the notion of semi-similarity, which is motivated by the use of a related concept by A. Feldzamen for spectral operators [9]. We call A and B *semi-similar* if there exists a sequence of mutually orthogonal self-adjoint projections $\{P_i\}$ commuting with A and B so that $I = \Sigma_i P_i$ and for each i, there exists an invertible operator S_i on $P_i H$, so that $S_i^{-1} A S_i = B \mid P_i H$. That is, there is a "complete" family of reducing subspaces for A and B, so that A is similar to B on each of these subspaces. Let A and B be semi-similar as above. By considering first the operator $X = \Sigma_i \| S_i \|^{-1} S_i P_i$ on H and then $Y = \Sigma_i \| S_i^{-1} \|^{-1} S_i^{-1} P_i$, we have that $AX = AB$ and $YA = BY$, where X and Y are quasiaffinities [14]. Thus this notion of semi-similarity implies the notion of quasi-similarity which is used by various authors to describe certain operators.

THEOREM 3.1. *Let f be a locally nonconstant analytic function on a domain $\mathscr{D}$ and let N be a normal operator. If T is an operator with $\sigma(T) \subset \mathscr{D}$ and $f(T) = N$, then there is a central projection P of T so that*

$$T = T_0 \oplus T_1\,,$$

where $T_0 = T \mid PH$ and $T_1 = T \mid (I - P)H$, T_0 is semi-similar to a normal operator N_0, $\sigma(N_0) = \sigma(T_0)$ and N_0 is a normal solution to

$f(\cdot) = N \mid PH$. *Finally, T_1 is an algebraic operator with $f(T_1) = 0$.*

Proof. Let $H = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$ be the decomposition of H so that $T = \int_\Lambda \oplus T(\lambda)\mu(d\lambda)$ is the primary decomposition of T. Since $N \in Z(T)$, $N = \int_\Lambda \oplus g(\lambda)I(\lambda)\mu(d\lambda)$, where $g \in L_\infty(\Lambda, \mu)$, and moreover $f(T(\lambda)) = g(\lambda)I(\lambda)$ almost everywhere [11].

Let $\lambda \in E_0$ if and only if $f(z) - g(\lambda)$ has only zeros of multiplicity one in $\sigma(T)$. If we let g_0 be the characteristic function of the set E_0, $P = \int_\Lambda \oplus g_0(\lambda)I(\lambda)\mu(d\lambda)$, then $T_0 = T/P$ is easily seen to be semi-similar to a normal operator $N_0 = N/P$ using Theorem 2.1.

On the complement of E_0, the function $f(z) - g(\lambda)$ has as least one multiple root. Since f' is locally nonzero there are only a finite number of distinct zeros of f' in $\sigma(T)$. Let $z_1, \cdots, z_k$ be the zeros of f' in $\sigma(T)$. Now a multiple root of $f(z) - g(\lambda)$ must be one of the numbers $z_1, \cdots, z_k$. Let F_i be the measurable set of λ in Λ for which $f(z) - g(\lambda)$ has the multiple root z_i. Then $E_i = F_i - \bigcup_{j<i} F_j$ are disjoint measurable sets so that $\Lambda = \bigcup_0^k E_i$. If $\lambda_1, \lambda_2 \in E_j (j > 0)$, then $f(z) - g(\lambda_1)$ and $f(z) - g(\lambda_2)$ both have the root z_i and therefore, g is constant on each $E_j (j > 0)$. If $g(\lambda) = a_i$ on $E_i (i > 0)$, then $T(\lambda)$ satisfies the equation $f(z) - a_i$ for λ in E_i and it follows that $T(\lambda)$ satisfies a complex polynomial $p_i(z)$ for $\lambda \in E_i (i > 1)$. Thus if $P_1 = I - P_0$ and $T_1 = T \mid P_1H$, $p(T_1) = 0$ for $p = p_1 \cdots p_k$.

From Theorem 2.1 it is clear that $\sigma(N_0) \subset \sigma(T_0)$, in fact, z belongs to the essential range of $z_i(\lambda)$ given in (2.4) for some i if and only if z is in $\sigma(N)$ and such a z is in $\sigma(T)$. Conversely, we shall show that $\sigma(N_0) \supset \sigma(T_0)$. Let $N/P_0H = N_1$, then we are considering $f(T_0) = N_1$ and T_0 is semi-similar to N_0. Let $z_0 \in \sigma(T_0)$ and $\varepsilon > 0$ be given. Denote by S_0 a ball of radius r less than ε, centered at z_0 with $\bar{S}_0 \subset \mathscr{D}$, and with $f(z) - f(z_0) \neq 0$ on $\bar{S}_0$ except for $z = z_0$. Let $f(z_0) = z_1$, then by the spectral mapping theorem $z_1 \in \sigma(N_1)$ and by the local mapping theorem, there exists a neighborhood S_1 of z_1 and S_2 of z_0 contained in S_0, so that $f(\bar{S}_2) = \bar{S}_1$.

Let $E(\cdot)$ be the spectral measure for N_1, then $E(S_1)$ is not zero since $z_1 \in \sigma(N_1)$. Also $E(S_1) \in Z(T_0)$ so we denote T_{01} to be $T/E(S_1)H$ and similarly N_{01} and N_{11}. Thus, $f(T_{01}) = N_{11}$ and N_{01} is the normal operator semi-similar to T_{01} given by Theorem 2.1. Since $\sigma(N_{11}) \subset \bar{S}_1$, by the spectral mapping and local mapping theorems we have that S_2 must contain a component of $\sigma(T_{01})$. By Proposition 2.2 there is a z_2 in $\sigma(N_{01}) \subset \sigma(N_0)$ so that $|z_2 - z_0| < \varepsilon$. Since ε was arbitrary, we may conclude that $\sigma(T_0) \subset \sigma(N_0)$ and the proof is complete.

Whenever f' has no zeros on $\sigma(T)$ then a theorem of C. Apostol has shown that T is similar to a normal solution of (3.1) [1]. A generalization of that result will be given in Proposition 4.5. If, however, f' has zeros but $(f')^{-1}(0) \cap \sigma_p(T)$ is empty, then the operator T_1 does not occur need to in the above theorem and we have the following corollary.

COROLLARY 3.2. *Let f be a locally nonzero analytic function on a domain $\mathscr{D}$ and let N be a normal operator. If T is an operator with $\sigma(T) \subset \mathscr{D}$, $f(T) = N$ and $\sigma_p(T) \cap (f')^{-1}(0) = \phi$, then there exists a normal operator N_0 with $\sigma(N_0) = \sigma(T)$, $f(N_0) = N$ and T is semi-similar to N_0.*

Prior to C. Apostol's work, it was shown by J. Stampfli that whenever A^n is normal and A is invertible, then A is similar to a normal operator [18]. It easily follows from Stampfli's result that whenever $0 \notin \sigma_p(A)$, then A is semi-similar to an nth root of N. This result is also an application of the above corollary where, of course, $f(z) = z^n$.

REMARK. That $\sigma(T_0) = \sigma(N_0)$ in Theorem 3.1 also follows the result of C. Apostol, C. Foias, and I. Colojoara when we have first shown that T_0 and N_0 are quasisimilar. For the first author proves that solutions of (3.1) are generalized scalar operators and the later authors have shown that quasisimilarity between decomposable operators preserves the spectrum. Since decomposable operators possess hyperinvariant subspaces, it follows from C. Apostol's results that solutions to (3.1) have hyperinvariant subspaces. However, this fact is also immediate by applying Theorem 2.5 to solutions of (3.1).

The following theorem and corollary generalize existing theorems and are obtained by placing some condition on $f(z)$. We shall only briefly indicate their proofs.

THEOREM 3.3. *Let T satisfy* (3.1) *and let $\{z_i\}_{i=1}^k$ be the zeros of $f'(z)$ in $\sigma(T)$ with multiplicities $\{n_i\}_{i-1}^k$. Assume that for each i there exists a neighborhood N_i of z_i so that there are at most m elements in $N_i \cap \sigma(T) \cap f^{-1}(z)$ for each z in $\sigma(N)$. Then there exists an orthogonal projection P in $R(T)'$ so that*

$$T = T_0 \oplus T_1 ,$$

where $T_1 = T/PH$ is algebraic and satisfies $p(z) = \prod_{i=1}^k (z - z_i)^{n_i}$ and T_0 is similar to an operator S_0 which satisfies a monic abelian polynomial of degree at most m.

Proof. The proof is similar to the proof of Theorem 3.1 in that T_1 is the same operator in each case. Here because of the restriction on the spectrum we divide $\sigma(T_0(\lambda))$ into at most k distinct pieces so that each contains at most m points of $\sigma(T_0(\lambda))$ and each is of multiplicity one. From such a decomposition the theorem will follow.

COROLLARY 3.4. *Let T^n be normal where $\sigma(T)$ lies in m sectors of the plane, each of width at most $2\pi/n$, then T is similar to the direct sum of a nilpotent operator T_0 and an operator T_1, which satisfies a polynomial of degree m with coefficients in the center of the von Neumann algebra generated by T_1.*

4. Operators satisfying an abelian polynomial.

In this section, we give several results in the study of opetators which satisfy

$$p(A) = 0 , \tag{4.1}$$

where $p(z)$ is a monic polynomial with coefficients which are commuting normal operators and A commutes with the coefficients. In view of Proposition 2.3, this problem subsumes the study of roots of abelian analytic functions. First, we shall discuss in some detail the results obtained whenever the polynomial is of degree two, and give results related to Corollary 3.2. As mentioned in the introduction, N. Dunford has studied n-normal operators from the viwpoint of when they were spectral operators. We relate our work to those results and to later works of T. Hoover [13] and H. Radjavi and P. Rosenthal [15, 16]. For example, several authors have shown that whenever A is n-normal, then A is a scalar multiple of the identity operator or A has nontrivial hyperinvariant subspaces. These results also follow from Theorem 2.5.

Recently, H. Radjavi and P. Rosenthal have given a characterization of operators satisfying certain polynomials of degree 2. Specifically, they have studied solutions to $z^2 + az = N$, where N is a normal operator [16]. The following theorem generalizes their results and a similar result of H. Behncke [3].

THEOREM 4.1. *Let T be a root of $p(z)$ where the degree of p is less than or equal to 2 and the coefficients of $p(z)$ are in $Z(T)$. Then there exists a central projection P of T, so that*

$$T = T_0 \oplus T_1$$

where $T_0 = T/PH$ and $T_1 = T/(I - P)H$, T_0 is normal, T_1 is unitarily equivalent to an operator of the form

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

on $K \oplus K$, *where* B, C, D *are commuting normal operators on* K. *Moreover,* $\sigma(B) \cup \sigma(D) = \sigma(T_1)$ *and* C *is positive definite.*

The proof of Theorem 4.1 will follow from a direct integral reduction of T and the next lemma. Recall that an operator is called *primary* if the von Neumann algebra it generates is a factor. The following lemma has a direct elementary proof. However, it does follow from A. Brown's nonelementary work [4] and we cite that as a proof.

LEMMA 4.2. *Let* A *be a primary operator on* H $(\dim H > 2)$. *If* $A^2 + bA + c = 0$ *for complex numbers* b *and* c, *then* A *is unitarily equivalent to*

$$\begin{bmatrix} \gamma I & \beta I \\ 0 & \alpha I \end{bmatrix},$$

on $K \oplus K$, *where* $\{\gamma, \alpha\} = \sigma(A) = \{1/2(-b \pm (b^2 - 4c)^{1/2})\}$ *and* $\beta = (\rho^2 - |\alpha - \gamma|^2)^{1/2}$, *where* $\rho = \|A - \alpha I\|$.

Proof of Theorem 4.1. Let $T = T_0 \oplus T_1$ be the unique central decomposition of T by projection P so that T_0 is normal and T_1 is completely nonnormal. If T satisfies $T^2 + TN_1 + N_2 = 0$, then $T_1^2 + T_1L_1 + L_2 = 0$ where $L_i = N_i/(I - P)H$ and $L_i \in Z(T_i)$ $(i = 1, 2)$. We decompose $H_1 = (I - P)H$ by the primary decomposition of T_1. Thus $H_1 = \int_\Lambda \oplus H(\lambda)\mu(d\lambda)$ and

$$T_1 = \int_\Lambda \oplus\, T_1(\lambda)\mu(d\lambda)\ ,$$

where $T_1(\lambda)$ is a primary operator defined on H_λ. Moreover, there exist bounded Borel functions f_1 and f_2 on Λ so that for $i = 1, 2$,

$$L_i = \int_\Lambda \oplus f_i(\lambda)I(\lambda)\mu(d\lambda)\ .$$

Therefore, we may conclude that

$$T_1(\lambda)^2 + f_1(\lambda)T_1(\lambda) + f_2(\lambda)I(\lambda) = 0$$

almost everywhere. From our proposition, $T_k(\lambda)$ is unitarily equivalent to

$$\begin{bmatrix} g(\lambda)I_\lambda & h(\lambda)I_\lambda \\ 0 & k(\lambda)I_\lambda \end{bmatrix}$$

on $K_\lambda \oplus K_\lambda$ where I_λ is the identity operator on K_λ, where $g, h,$ and

k are measurable, $h(\lambda) > 0$ and the projection $P(\lambda)$ onto the subpace $K_\lambda \oplus 0$ is measurable. We let $Q(\lambda) = I(\lambda) - P(\lambda)$ and then $P(\lambda)T_1(\lambda)P(\lambda) = g(\lambda)P(\lambda)$, $P(\lambda)T_1(\lambda)Q(\lambda) = 0$, $P(\lambda)T_1(\lambda)Q(\lambda) = h(\lambda)P(\lambda)$ and $Q(\lambda)T_1(\lambda)Q(\lambda) = k(\lambda)Q(\lambda)$ and the result follows.

REMARK 1. That $N_1, N_2 \in Z(A)$ is not essential to Theorem 4.1. The same conclusion holds if A is any root of a locally nonzero abelian polynomial of degree less or equal to 2. We need only decompose A as in Theorem 2.1 and thus have $g(\lambda)A(\lambda)^2 + h(\lambda)A(\lambda) + k(\lambda)I(\lambda) = 0$ almost everywhere. By Theorem 4.1, there exists a projection $Q(\lambda)$ measurable with respect to λ, so that $A(\lambda)Q(\lambda) = r_1(\lambda)Q(\lambda)$, $P(\lambda)A(\lambda)P(\lambda) = r_2(\lambda)Q(\lambda)$ where $P(\lambda) = I(\lambda) - Q(\lambda)$ and $Q(\lambda)A(\lambda)P(\lambda) = c(\lambda)Q(\lambda)$ where $c(\lambda)$ is a positive operator on $H(\lambda)$. The more general result now follows.

REMARK 2. A. Brown called 2-normal operators binormal and H. Gonsher called them J_2 operators [4, 12]. Hence, Theorem 4.1 implies that: A is a binormal operator if and only if A is a zero of a locally nonzero abelian polynomial of degree less than or equal to 2. For a discussion of the unitary invariant of these operators we refer the reader to [2].

We can obtain various known theorems as special cases of the preceeding theorems. For example, we can generalize Theorem 3 in [16] with the following corollary.

COROLLARY 4.3. *Let $T^n = N$, where N is normal and let $\sigma(T)$ lie in two sectors of the plane each with width less than $2\pi n^{-1}$. Then there are mutually orthogonal central projections P_0, P_1, and P_2 of T with $I = P_0 + P_1 + P_2$ and*

$$T = T_0 \oplus T_1 \oplus T_2$$

where $T_0 = T/P_0H$ is nilpotent of order n, $T_1 = T/P_1H$ is normal and $T_2 = T/P_2H$ is unitarily equivalent to

$$\begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

where B, C, and D are commuting normal operators with C positive definite.

Proof. Let P_0 be the central projection so that T/P_0H is normal and $T/(I - P_0)H$ is completely nonnormal. If we apply Corollary 3.6 to $T/(I - P_0)H$ we can obtain P_1 and P_2 so that T/P_1H is algebraic and in fact $T^n/P_1H = 0$ and T/P_2H satisfies a monic polynomial of degree 2 with coefficients in $Z(T/P_2H)$. Using Theorem 4.1 we now conclude the complete structure of T.

In Theorem 2.1 we see that if the root functions are different almost everywhere, then the operator zero is semi-similar to a normal zero. We use this observation in the following result concerning solutions of an abelian polynomial of degree 2 which will be useful. It differs from the preceding results in that it utilizes semi-similarity.

PROPOSITION 4.4. *Let T satisfy an abelian polynomial of degree 2. Then there exists unique central decomposition of T into*

$$T_0 \oplus T_1$$

so that T_0 is unitarily equivalent to the commuting sum of a normal operator and a nilpotent operator of index 2. The operator T_0 has no reducing subspace on which it is similar to a normal operator and T_1 is semi-similar to a normal operator.

Proof. We let the root functions be $\{r_i(\lambda)\}_{i=1}^{2}$ and set $M = \{\lambda \mid r_1(\lambda) = r_2(\lambda)\}$. If g is the characteristic function of M, then $P = \int_{\Lambda} \oplus\, g(\lambda) I(\lambda) \mu(d\lambda)$ is a central projections for T. We let T_0 be the completely nonnormal part of T/P and the proposition follows from the fact that on the complement of M, $r_1(\lambda) \neq r_2(\lambda)$ almost everywhere.

In the case of operators satisfying an abelian analytic function, we always have by Theorem 2.1 that they are piecewise similar to spectral operators. The question naturally arises as to when are they spectral. This has been studied by both N. Dunford and C. Apostol for the special cases they considered respectively [1, 6]. The following sufficient condition follows easily from the proof of Theorem 2.1.

PROPOSITION 4.5. *Let T be a root of a locally nonzero analytic abelian function ψ which has root functions $\{r_i(\lambda)\}_{i=1}^{m}$ in $\sigma(T)$ satisfying $\prod_{i \neq j} |r_i(\lambda) - r_j(\lambda)| \geqq \delta > 0$ almost everywhere. Then T is similar to a normal root of ψ.*

Proof. The root functions are given by Theorem 1.3 and under the assumption $\prod_{i \neq j} |r_i(\lambda) - r_j(\lambda)| > 0$ almost everywhere we have no multiple roots. Furthermore, the projections given by equation (2.1) are just $E_i(\lambda) = p_i(T(\lambda))$ where $p_i(z) = \prod_{j \neq i} (z - r_j(\lambda))(r_i(\lambda) - r_j(\lambda))^{-1}$ and are essentially bounded under the hypothesis on $\{r_i(\lambda)\}$.

In fact, a necessary and sufficient condition can be given in case $\prod_{i \neq j} (r_i(\lambda) - r_j(y)) \neq 0$ almost everywhere.

PROPOSITION 4.6. *If T is a solution of an abelian analytic function with $\prod_{i \neq j}(r_i(\lambda) - r_j(\lambda)) \neq 0$ almost everywhere, then T is a scalar type operator if and only if $\prod_{i \neq j_0}(r_{i_0}(\lambda) - r_i(\lambda))^{-1} \| T(\lambda) - r_i(\lambda) \|$ is essentially bounded for $1 \leqq i_0 \leqq n$.*

REMARK. The theorem of J. Stampfli for T^n normal and T invertible as well as S. Foguel's theorem and C. Apostol's theorem for $p(T)$ normal and $p'(z) \neq 0$ on $\sigma(T)$ and $f(T)$ normal and $f'(z) \neq 0$ on $\sigma(T)$ respectively, follow from these propositions.

Unfortunately, these conditions are not sufficient as we shall see below. In the case of an operator T satisfying a second degree monic polynomial with coefficients in $Z(T)$, we can given necessary and sufficient for that T be similar to a normal solution of the polynomial.

THEOREM 4.7. *Let T satisfy a monic second degree polynomial with coefficients in $Z(T)$. If $T = \int_A \oplus T(\lambda)\mu(d\lambda)$ is the primary decomposition of T, $\{r_i(\lambda)\}_{i=1}^2$ are the root functions of the polynomial and $\rho(\lambda) = \| T(\lambda) - r_1(\lambda) \|$, then T is a spectral type operator of nilpotent index 2 if and only if $\{\rho(\lambda) \mid r_1(\lambda) - r_2(\lambda) \mid^{-1}: r_1(\lambda) \neq r_2(\lambda)\}$ is essentially bounded.*

Proof. This follows from Propositions 4.4 and 4.6.

We shall give an example which yields some of the results in N. Dunford's work. Let $H = L_2(0, 1) \oplus L_2(0, 1)$ and M_f denote the multiplication operator on $L_2(0, 1)$ for $f \in L^\infty(0, 1)$. If

$$A = \begin{bmatrix} M_f & M_g \\ M_h & M_k \end{bmatrix},$$

where $f, g, h, k \in L^\infty(0, 1)$, then clearly A satisfies a second degree monic polynomial $z^2 - N_1 z + N_2$ where the coefficients

$$N_1 = \begin{bmatrix} M_{f+k} & 0 \\ 0 & M_{f+k} \end{bmatrix}$$

and

$$N_2 = \begin{bmatrix} M_{fk-gh} & 0 \\ 0 & M_{fk-gh} \end{bmatrix}.$$

Thus, if we take the direct integral decomposition determined by Lebesgue measure on $[0, 1]$ and $H(\lambda) = C^2$, then N_1, N_2 are obviously diagonal operators and A decomposes with $A(\lambda) = \begin{bmatrix} f(\lambda) & g(\lambda) \\ h(\lambda) & k(\lambda) \end{bmatrix}$. Then as in Proposition 4.4, there is a Borel set M so that if g is the characteristic function on M, then A is decomposed by $\int_A \oplus g(\lambda)I(\lambda)\mu d(\lambda)$ into $A_1 \oplus A_2$ so that A_1 is a spectral operator of order 2 and A_2 is semi-similar to a normal operator. By Theorem 4.7, A is a spectral operator iff $\{\|A(\lambda) - r_1(\lambda)\| \, |r_1(\lambda)|^{-1} \colon \lambda \in \Lambda - M\}$ is essentially bounded. This later condition is equivalent (following the notation in [4]) to

$$\frac{((f(\lambda)) - k(\lambda))^2 + g(\lambda)^2 + h(\lambda)^2}{\delta(\lambda)^2} \colon \lambda \in \Lambda - M$$

being essentially bounded where $\delta(\lambda) = ((f(\lambda) - k(\lambda))^2 + 4g(\lambda)h(\lambda))^{1/2}$. Note that $\delta(\lambda) = 0$ on M which parallels the treatment in [4, 6; XI].

Finally, we given an example first introduced by J. Stampfli [17] to show that sequare roots of normal operators need not be spectral. Let

$$A_f = \begin{bmatrix} M_t & M_f \\ 0 & M_{-t} \end{bmatrix}$$

on $H = L_2(0, 1) \oplus L_2(0, 1)$ where $f \in L^\infty(0, 1)$. Then A_f^2 is normal for each f, however A_f is a spectral operator (in fact scalar type operator) if and only if $|t^{-1}f(t)|$ is essentially bounded. Hence, the example of J. Stampfli follows. The operator

$$\begin{bmatrix} M_t & I \\ 0 & M_{-t} \end{bmatrix}$$

is the square root of a normal operator which is not a spectral operator.

We close by remarking on several areas of further research involving these methods and theorems. The theorems in §§1 and 2 can be modified in case $\psi(z)$ takes values in certain commutative algebras of spectral operators; however, the nilpotent operators become quasinilpotent and are not necessarily of finite type. Most of the theorems can be obviously modified if the normal operators are replaced by commuting scalar type operators whenever similarity or semi-similarity is involved. Some results in this direction have been obtained and further work is in progress.

REFERENCES

1. C. Apostol, *Spectral decompositions and functional calculus*, Rev. Roumaine Math. Pures Appl., **13** (1968), 1481-1528.

2. E. Azoff, *Spectrum and direct integral*, Trans. Amer. Math. Soc., to appear.

3. H. Behncke, *The structure of certain nonnormal operators*. II, Indiana Univ. Math. J., **22** (1972), 301-308.

4. A. Brown, *Unitary equivalence of binormal operators*, Amer. J. Math., **76** (1954), 414-434.

5. D. Deckard and C. Pearcy, *On matrices over the rings of continuous complex-valued functions on a Stonian space*, Proc. Amer. Math. Soc., **14** (1963), 322-328.

6. N. Dunford, *A spectral theory for certain operators on a direct sum of Hilbert spaces*, Math. Ann., **162** (1966), 294-330.

7. N. Dunford and J. Schwartz, *Linear Operators, Parts* I, II, III, Wiley Interscience (1957, 1963, 1971), New York.

8. I. Colojoara and C. Foias, *Theory of Generalized Spectral Operators*, Gordon and Breach (1968), New York.

9. A. N. Feldzammen, *Semi-similarity invariants for spectral operators on Hilbert space*, Trans. Amer. Math. Soc., **100** (1961), 277-324.

10. S. Foguel, *Algebraic functions of normal operators*, Israel J. Math., **6** (1968), 199-201.

11. F. Gilfeather, *On a functional calculus for decomposable operators and applications to normal, operator-valud functions*, Trans. Amer. Math. Soc., **176** (1973), 369-383.

12. H. Gonsher, *Spectral theory for a class of nonnormal operators*, Canada J. Math., **8** (1956), 449-461.

13. T. Hoover, *Hyperinvariant subspaces for n-normal operators*, Acta. Soc. Math., **32** (1971), 109-119.

14. B. Sz. Nagy and C. Foias, *Analyse Harmonique des Operateurs de l'espace de Hilbert* (1967), Budapest.

15. H. Radjavi and P. Rosenthal, *Hyperinvariant subspaces for spectral and n-normal operators*, Acta. Sci. Math., **32** (1971), 121-125.

16. ———, *On roots of normal operators*, J. Math. Anal. and App., **34** (1971), 653-665.

17. J. Schwartz, *W^* Algebras*, Gordon and Breach (1967), New York.

18. J. G. Stampfli, *Roots of scalar operators*, Proc. Amer. Math. Co., **13** (1962), 796-798.

Received April 10, 1973 and in revised form January 9, 1974. This work was partially supported by the National Science Foundation under grant, NSF-31435.

UNIVERSITY OF NEBRASKA-LINCOLN

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

BEST APPROXIMATION BY A SATURATION CLASS OF POLYNOMIAL OPERATORS

D. S. GOEL, A. S. B. HOLLAND,
C. NASIM, AND B. N. SAHNEY

The problem of determining a saturation class has been considered by Zamanski, Sunouchi and Watari and others. Zamanski has considered the Cesaro means of order 1 and Sunouchi and Watari have studied the Riesz means of type n. The object of the present paper is to extend these results by considering Nörlund means which include the above-mentioned results as particular cases.

1. Let $\{p_n\}$ be a sequence of positive constants such that

$$P_n = p_0 + \cdots + p_n \longrightarrow \infty \quad \text{as} \quad n \longrightarrow \infty .$$

A given series $\sum_{n=0}^{\infty} d_n$ with the sequence of partial sums $\{S_n\}$ is said to summable (N, p_n) to d, provided that

$$\begin{aligned} N_n\left[\sum_{l=0}^{\infty} d_l\right] &= \frac{1}{P_n}\sum_{k=0}^{n} P_{n-k} d_k \\ &= \frac{1}{P_n}\sum_{k=0}^{n} p_{n-k} S_k \longrightarrow d\,, \quad \text{as} \quad n \longrightarrow \infty\,, \end{aligned} \tag{1.1}$$

and N_n are called the Nörlund operators.

Let

$$\frac{1}{2}a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx) \equiv \sum_{k=0}^{\infty} A_k(x) \tag{1.2}$$

be the Fourier series associated with a continuous periodic function $f(x)$, with period 2π.

We define

$$N_n(x) \equiv N_n(f; x) \equiv \frac{1}{P_n}\sum_{k=0}^{n} P_{n-k} A_k(x) \tag{1.3}$$

and the norm

$$\|f(x) - N_n(x)\| \equiv \max_{0 \leq x \leq 2\pi} |f(x) - N_n(x)| \,.$$

If there exists positive nonincreasing function $\phi(n)$ and a class of functions K, with the following properties:

(I) $\|f(x) - N_n(x)\| = o(\phi(n)) \Longrightarrow f(x)$ is constant,

(II) $\|f(x) - N_n(x)\| = O(\phi(n)) \Longrightarrow f(x) \in K$

and

(III) $f(x) \in K \Longrightarrow ||f(x) - N_n(x)|| = O(\phi(n))$,
then the Nörlund operators are saturated with the order $\phi(n)$ and the class K.

In this paper we prove that the above method of summations is saturated with the order p_n/P_n and that the class K consists of all continuous functions f such that $\tilde{f} \in Lip\ 1$, where $\tilde{f}$ is the conjugate function of f. By definition

$$\tilde{f}(x) = \frac{1}{2\pi} \int_0^{\pi} [f(x+t) - f(x-t)] \cot \frac{1}{2} t\, dt\ ,$$

if the integral converges absolutely for all x and if

$$\int_0^{\pi} |f(x+t) - f(x-t)| \cot \frac{t}{2} dt$$

is an integrable function.

The problem of determining a saturation class by considering $(C, 1)$ means of the Fourier series of $f(x)$ has been considered by Zamanski [6]. Sunouchi and Watari [4] have considered the problem by taking (R, λ, k) means of the Fourier series. Some of these results were later extended by Sunouchi [3] and others [2, 5].

2. We shall prove the following theorem.

THEOREM. *Let $\{p_n\}$ be a sequence of positive constants satisfying the following conditions,*

$$(2.1) \qquad \frac{p_{n-k}}{p_n} \longrightarrow 1 \quad as \quad n \longrightarrow \infty \quad for\ a\ fixed \quad k \leqq n\ ,$$

and

$$(2.2) \qquad \sum_{k=0}^{n} |p_{n-k} - p_{n-k-1}| = O(p_n) \quad where \quad [p_{-1} = 0]\ .$$

Then the operators N_n are saturated with order p_n/P_n and the class of all continuous functions f for which $\tilde{f} \in Lip\ 1$.

The following lemmas are required for the proof of the theorem.

LEMMA 2.1. *If*

$$||f(x) - N_n(x)|| = o\left[\frac{p_n}{P_n}\right]$$

then f is a constant.

Proof. From (1.3) we obtain

$$\begin{aligned}\frac{1}{\pi}\int_{-\pi}^{\pi} N_n(x)\cos rx\,dx &= \frac{1}{\pi}\int_{-\pi}^{\pi}\sum_{k=0}^{n}\frac{P_{n-k}}{P_n}A_k(x)\cos rx\,dx\\ &= \frac{1}{\pi}\sum_{k=0}^{n}\frac{P_{n-k}}{P_n}\int_{-\pi}^{\pi}A_k(x)\cos rx\,dx\\ &= \frac{P_{n-r}}{P_n}a_r\,.\end{aligned}$$

Thus,

$$\begin{aligned}a_r - \frac{P_{n-r}}{P_n}a_r &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos rx\,dx - \frac{1}{\pi}\int_{-\pi}^{\pi} N_n(x)\cos rx\,dx\\ &= \frac{1}{\pi}\int_{-\pi}^{\pi}\cos rx\,[f(x) - N_n(x)]dx\,,\end{aligned}$$

hence

$$\left|a_r - \frac{P_{n-r}}{P_n}a_r\right| \leqq \|f(x) - N_n(x)\|\frac{1}{\pi}\int_{-\pi}^{\pi}1\cdot dx = o\left[\frac{p_n}{P_n}\right].$$

Consequently

$$a_r\left\{\frac{p_n + \cdots + p_{n-r+1}}{p_n}\right\} = o(1)\,, \tag{2.3}$$

and since $p_r > 0$ for all r, we have $(p_n + \cdots + p_{n-r+1})/p_n \geqq 1$ for $r \geqq 1$.

Thus from (2.3) it follows that $a_r = 0$, for each $r \geqq 1$. Similarly we can show that $b_r = 0$ for each $r \geqq 1$. Hence $f(x) = 1/2a_0$, a constant.

LEMMA 2.2. *If*

$$\|f(x) - N_n(x)\| = O\left[\frac{p_n}{P_n}\right]$$

and condition (2.1) *is satisfied, then* $\tilde{f}(x) \in Lip\ 1$.

Proof. It can be shown without much difficulty that if

$$\|f(x) - N_n(x)\| = O\left[\frac{p_n}{P_n}\right],$$

then

$$\left\|\sum_{k=1}^{N}\frac{p_n + \cdots + p_{n-k+1}}{p_n}A_k(x)\left[1 - \frac{k}{N+1}\right]\right\| = O(1),\ N \leqq n\,.$$

Taking the limit as $n \longrightarrow \infty$, and using condition (2.1), we obtain

$$\left\| \sum_{k=1}^{N} kA_k(x)\left[1 - \frac{k}{N+1}\right]\right\| = O(1) . \tag{2.4}$$

The left hand side of the above equation represents the $(C, 1)$ mean of the series

$$\sum_{k=1}^{\infty} - kA_k(x) .$$

Since $-kA_k(x) = B'_k(x)$, where $\sum_{k=1}^{\infty} B_k(x) \equiv \sum_{k=1}^{\infty} (b_k \cos kx - a_k \sin kx)$ is the conjugate series of (1.2), then (2.4) is equivalent to

$$\|\tilde{\sigma}'_N(f)\| < M$$

which implies that $\tilde{f}(x) \in Lip\ 1$, [1].
($\tilde{\sigma}_N(f)$ represents the $(C, 1)$ mean of the conjugate series.)

LEMMA 2.3. *Assume* $\tilde{f} \in Lip\ 1$. *If the sequence* $\{p_n\}$ *satisfies condition* (2.2), *then*

$$\|f(x) - N_n(x)\| = O\left[\frac{p_n}{P_n}\right] .$$

Proof. Since, by definition

$$\tilde{S}_n(\tilde{f}, x) = \frac{1}{\pi}\int_0^{\pi}[\tilde{f}(x, t) - \tilde{f}(x - t)]\frac{\cos\dfrac{t}{2} - \cos\left[n + \dfrac{1}{2}\right]t}{2\sin\dfrac{t}{2}}dt$$

where $\tilde{S}_n(\tilde{f}, x)$ denotes the partial sums of the conjugate series associated with $\tilde{f}(x)$, we have

$$\begin{aligned}
N_n(\tilde{S}_n(\tilde{f}, x)) &= \frac{1}{P_n}\sum_{k=0}^{n} p_{n-k}\tilde{S}_k(\tilde{f}, x) \\
&= \frac{1}{P_n}\sum_{k=0}^{n} p_{n-k}\frac{1}{2\pi}\int_0^{\pi}[\tilde{f}(x + t) - \tilde{f}(x - t)]\cot\frac{1}{2}t\, dt \\
&\quad - \frac{1}{P_n}\sum_{k=0}^{n} p_{n-k}\frac{1}{2\pi}\int_0^{\pi}[\tilde{f}(x + t) - \tilde{f}(x - t)]\frac{\cos\left[k + \dfrac{1}{2}\right]t}{\sin\dfrac{1}{2}t}dt .
\end{aligned}$$

Since the function $\tilde{f}(x) \in Lip\ 1$, $-f + (1/2)a_0$ is identical to $\tilde{\tilde{f}}$, therefore

$$f(x) - N_n(f, x) = \frac{1}{2\pi}\int_0^{\pi}[\tilde{f}(x + t) - \tilde{f}(x - t)]K_n(t)dt , \tag{2.5}$$

where

$$K_n(t) = \frac{1}{P_n \sin \frac{1}{2}t} \sum_{k=0}^{n} p_{n-k} \cos\left[k + \frac{1}{2}\right]t \,.$$

Now by partial summation

$$\begin{aligned} K_n(t) &= \frac{1}{2P_n \sin^2 \frac{1}{2}t} \sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \sin(k+1)t \\ &= \frac{1}{P_n}\left\{\frac{2}{t^2} + O(1)\right\} \sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \sin(k+1)t \\ &= \frac{2}{P_n t^2} \sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \sin(k+1)t + O\left[\frac{p_n}{P_n}\right], \end{aligned}$$

by hypothesis. Since $\tilde{f}(x)$ is certainly bounded, the right hand side of (2.5) becomes

$$(2.6)\quad \frac{1}{\pi P_n}\int_0^\pi [\tilde{f}(x+t) - \tilde{f}(x-t)]\frac{1}{t^2}\left\{\sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \sin(k+1)t\right\}dt + O\left[\frac{p_n}{P_n}\right].$$

Let us write

$$F_n(t) = \frac{1}{P_n}\int_t^\pi \frac{1}{u^2}\left\{\sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \sin(k+1)u\right\}du \,.$$

Since $\tilde{f}(u) \in Lip\ 1$, it is an indefinite integral of a bounded function, say $\tilde{f}'(u)$. Further, since $\tilde{f}(x+t) - \tilde{f}(x-t) = O(t)$, as $t \to 0$, while for fixed n, $F_n(t) = O(\log(1/t))$, we can integrate (2.6) by parts to obtain

$$\frac{1}{\pi}\int_0^\pi [\tilde{f}'(x+t) + \tilde{f}'(x-t)]F_n(t)dt + O\left[\frac{p_n}{P_n}\right],$$

noting that the integrated term vanishes at both limits. The absolute value of this above expression is now,

$$(2.1)\qquad O\left\{\int_0^\pi |F_n(t)|\,dt\right\} + O\left[\frac{p_n}{P_n}\right] \text{ since } \tilde{f}' \text{ is bounded}\,.$$

Now

$$\begin{aligned} F_n(t) &= \frac{1}{P_n}\sum_{k=0}^{n} (p_{n-k} - p_{n-k-1}) \int_t^\pi \frac{\sin(k+1)u}{u^2}du \\ &= \frac{1}{P_n}\sum_{k=0}^{n} (p_{n-k} - p_{n-k-1})(k+1) \int_{(k+1)t}^{(k+1)\pi} \frac{\sin \nu}{\nu^2}d\nu \,. \end{aligned}$$

However,

$$\int_{(k+1)t}^{(k+1)\pi} \frac{\sin \nu}{\nu^2} dv = \begin{cases} O(\log 1/(k+1)t) & \text{if} \quad (k+1)t < 1 \\ O(1/(k+1)^2 t^2) & \text{if} \quad (1(k+1)t \geqq 1 \,. \end{cases}$$

Hence

$$\int_0^\pi |F_n(t)|\,dt = O\left\{\frac{1}{P_n} \int_0^\pi \left[\sum_{\substack{(k+1)<1/t \\ k\geqq 0}} |p_{n-k} - p_{n-k-1}|(k+1)\log(1/(k+1)t) \right.\right.$$

$$\left. + \sum_{\substack{(k+1)\geqq 1/t \\ k\leqq n}} |p_{n-k} - p_{n-k-1}| 1/(k+1)t^2 \right] dt$$

$$= O\left\{\frac{1}{P_n} \sum_{k=0}^{n} |p_{n-k} - p_{n-k-1}| \left[\int_0^{1/(k+1)} (k+1)\log(1/(k+1)t)\,dt \right.\right.$$

$$\left.\left. + \int_{1/(k+1)}^{\pi} \frac{1}{(k+1)t^2} dt \right]\right\} .$$

Further,

$$\int_0^{1/(k+1)} \log(1/(k+1)t)\,dt = \int_0^1 \log\left(\frac{1}{u}\right) du = \text{constant}$$

and

$$\int_{1/(k+1)}^{\pi} \frac{1}{(k+1)t^2} dt < M \text{ (constant)} ,$$

therefore

$$\int_0^\pi |F_n(t)|\,dt = O\left\{\frac{1}{P_n} \sum_{k=0}^{n} |p_{n-k} - p_{n-k-1}|\right\} = O\left[\frac{p_n}{P_n}\right]$$

from (2.2).

Thus (2.7) and hence (2.6) is $O[p_n/P_n]$.
Consequently from (2.5), we have that

$$||f(x) - N_n(f, x)|| = O\left[\frac{p_n}{P_n}\right]$$

which proves the lemma.

The proof of the theorem now follows from Lemmas 2.1, 2.2, and 2.3.

The authors wish to thank Dr. B. Kuttner of the University of Birmingham, for his very helpful suggestions.

REFERENCES

1. J. Favard, *Sur la saturation des procédés de sommation*, J. de Math., **36** (1957), 359-372.

2. K. Ikeno and Y. Suzuki, *Some remarks of saturation problem in the local approximation*, Tôhoku Math. J., **20** (1968), 214-233.

3. G. Sunouchi, *On the class of saturation in the theory of approximation I, II, III*, Tôhoku Math. J., **12** (1960), 339-344; **13** (1961), 112-118; 320-328.
4. G. Sunouchi and C. Watari, *On determination of the class of saturation in the theory of approximation of Functions*, Proc. Japan Acad., **34** (1958), 477-481.
5. Y. Suzuki, *Saturation of local approximation by linear positive operators*, Tôhoku Math. J., **17** (1965), 210-221.
6. M. Zamanski, *Classes de saturation de certaines procédés d'approximation des séries de Fourier des functions continues*, Ann. Sci Ecole Normale Sup., **66** (1949), 19-93.

Received March 2, 1973.

THE UNIVERSITY OF CALGARY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

PUISEUX SERIES FOR RESONANCES AT AN EMBEDDED EIGENVALUE

JAMES S. HOWLAND

Let $H(\kappa) = T + \kappa B^*A$ be a self-adjoint perturbation of the self-adjoint operator T, and suppose that T has an eigenvalue λ_0 of finite multiplicity m embedded in its continuous spectrum. If the operator

$$Q(z) = A(T - z)^{-1}B^*$$

is bounded and can be continued meromorphically across the axis at λ_0, the asymptotic spectral concentration of the family $H(\kappa)$ at λ_0 is determined by the poles of

$$\kappa A(H(\kappa) - z)^{-1}B^* = I - [I + \kappa Q(z)]^{-1} . \tag{1}$$

These "resonances" can be expanded in a series of fractional powers of κ, and therefore have a unitarily invariant significance for the family $H(\kappa)$. An example shows that nonanalytic series may indeed occur; however, if a resonance is an actual eigenvalue of $H(\kappa)$ for all sufficiently small *real* κ, its series is analytic. Because the resonances cannot lie on the first sheet when κ is real, these series must have a special form. In the generic case, they yield, as the lowest order approximation to the imaginary parts of the resonances, the famous Fermi's Golden Rule. The case when λ_0 is embedded at a branch point of (1) is studied by means of a simple example.

To outline briefly, Puiseux expansions are obtained in § 1, and their special form is noted (c.f. [15, Theorem 4.2]). In § 2, a study of these series for perturbations which remove the degeneracy at λ_0 leads to Fermi's Golden Rule. The discussion of spectral concentration in § 3 relies heavily on the arguments of [3], particularly on a grouping of the resonances into "clusters" which act asymptoticly as a single simple pole. The examples appear in § 4. The appendix contains a technical result which simplifies not only Theorem 3.1 but also [3, Theorem 2.1] (c.f. [3, p. 156; Note (1)]). The results proved here were announced in [4].

Simon [14, 15] has recently discussed a similar problem for N-body Hamiltonians with dilatation analytic interactions. It is of particular interest that the Balslev-Combes technique which he employs reduces the problem to that of an *isolated* eigenvalue of a *non-self-adjoint* operator. This gives an interesting insight into the occurrence of Puiseux series, and suggests that, in the general case, resonance series can be viewed as perturbation series for an isolated eigenvalue of a suitable non-self-adjoint operator. Simon considers eigenvalues of arbitrary finite multiplicity, and not, as erroneously remarked in [4], only simple multiplicity.

Eigenvalues embedded at "thresholds" are not considered by Simon. Mathematically, a threshold may be variously described as (i) a branch point of an appropriate function, (ii) a point where the absolutely continuous part of T changes multiplicity, or (sometimes) (iii) an end point of the spectrum of T. The unperturbed eigenvalue in the second example of § 4 is a threshold in all three senses. A slightly revised Golden Rule is shown to apply to this case.

Let us conclude this introduction with an observation about the invariant significance of "resonances". It is tempting, at first glance, to call a point Λ a resonance of the self-adjoint operator H if the continuation of some matrix element $((H-\zeta)^{-1}f, f)$ across the spectrum of H has a pole at Λ. However, this definition is worthless; for if H is the multiplication

$$Hf(x) = xf(x) \qquad -\infty < x < \infty$$

(which is essentially the general case in which continuation is possible), then given any point Λ in the lower half-plane, there is a rational function $f(x)$ for which the continuation of

$$((H-\zeta)^{-1}f, f) = \int (x-\zeta)^{-1}|f(x)|^2 dx$$

has a pole at Λ. The "resonances" considered by various authors are always something more than this—poles of an S-matrix [11], of an integral operator [13], or (as here) of an operator-valued function. Accordingly, the definition of "resonance" is referred to some structure in addition to the operator H—such as outgoing subspaces, the representation of H as a differential operator, or a decomposition $H = T + AB^*$.

While something of this sort is necessary in general, in the case of an analytic perturbation $H(\kappa)$ of an embedded eigenvalue, a *unitarily invariant* significance can be attached to a Puiseux series $\Lambda(\kappa)$ of "resonances" in the weak sense which we have scorned above. There is of course additional structure here, too: the analyticity of the families $H(\kappa)$ and $\Lambda(\kappa)$.

To be precise, suppose that $H(\kappa)$ is an analytic family [6, Chapter VII] of closed operators, self-adjoint for real κ, with essential spectrum independent of κ. Let λ_0 be an eigenvalue of $H(0)$ and assume that for some vector f

$$((H(\kappa)-\zeta)^{-1}f, f)$$

has a continuation $F(\zeta, \kappa)$ to a meromorphic function of (ζ, κ) for $|\kappa| < \delta$ and $|\zeta - \lambda_0| < \delta$. Assume further that

$$\Lambda(\kappa) = \lambda_0 + \beta\kappa^{n/p} + \cdots \qquad \beta \neq 0$$

is a pole of $F(\zeta, \kappa)$ for each κ. Since for small κ, the term $\beta\kappa^{n/p}$ dominates those which follow it, $\Lambda(\kappa)$ *will be in the upper half-plane for* κ *in certain sectors of the complex plane, and will therefore be an eigenvalue of* $H(\kappa)$, because of the assumed invariance of the essential spectrum. Thus the same analytic family $\Lambda(\kappa)$ represents a "resonance" for some values of the perturbation parameter, and an actual eigenvalue of $H(\kappa)$ for others. Put differently, the resonances are continuations in κ of eigenvalues of $H(\kappa)$, and have, therefore, a unitarily invariant significance for the family $H(\kappa)$.

1. Puiseux series. The following assumptions will be made throughout this article. For proofs of the various assertions, see [2, 7, and 10].

Let $\mathscr{H}$ and $\mathscr{H}'$ be separable Hilbert spaces. Let T be a self-adjoint operator on $\mathscr{H}$ with resolvent $G(z) = (T - z)^{-1}$, and let A and B be closed, densely defined operators from $\mathscr{H}$ to $\mathscr{H}'$ such that $\mathscr{D}(T) \subset \mathscr{D}(A) \cap \mathscr{D}(B)$ and

$$(1.1) \qquad (Ax, By) = (Bx, Ay) \text{ for every } x, y \in \mathscr{D}(A) \cap \mathscr{D}(B) .$$

Suppose that for every $z \in \rho(T)$, the operator $AG(z)B^*$, which is defined on $\mathscr{D}(B^*)$, has a bounded extension $Q(z)$ to $\mathscr{H}'$, and that $I + Q(z)$ is invertible for some $z \in \rho(T)$. Then, for sufficiently small real κ, there is a self-adjoint extension $H(\kappa)$ of $T + \kappa B^*A$ the resolvent of which is

$$(1.2) \qquad R(z, \kappa) = G(z) - \kappa[BG(\bar{z})]^* [I + \kappa Q(z)]^{-1} AG(z)$$

whenever $z \in \rho(T)$ and $I + \kappa Q(z)$ has a bounded inverse. In particular, $H(0) = T$ and $R(z, 0) = G(z)$. We shall write $H(\kappa) = \int \lambda dE_\kappa(\lambda)$. If $\mathscr{M}(A^*)$ denotes the smallest reducing subspace of T which contains $\mathscr{R}(A^*)$, then $\mathscr{M} = \mathscr{M}(A^*) \cap \mathscr{M}(B^*)$ reduces both $H(\kappa)$ and T and $H(\kappa) = T$ on $\mathscr{M}^\perp$. Only the parts of $H(\kappa)$ and T in $\mathscr{M}$ are of interest in perturbation theory.

Let Ω be a neighborhood of a point λ_0 of the real axis, and $\Omega^\pm = \{z \in \Omega: \pm \operatorname{Im} z > 0\}$. Assume that $Q(z)$ has a continuation $Q^\pm(z)$ from $\Omega^\pm$ to Ω, which is analytic on Ω except for a simple pole at λ_0 with residue of finite rank m. The part of T in $\mathscr{M}$ is then absolutely continuous in $\Omega \cap \boldsymbol{R}$, except for an eigenvalue λ_0 of finite multiplicity equal to m. Since $Q^+(z)$ and $Q^-(z)$ do not in general agree on Ω, *the eigenvalue* λ_0 *is in general embedded in the absolutely continuous spectrum of* T.

If we now write

$$Q^{\pm}(z) = Q_c^{\pm}(z) + (\lambda_0 - z)^{-1}F$$

where F has finite rank and $Q_c^{\pm}(z)$ is analytic at λ_0, then $I + \kappa Q_c^{\pm}(z)$ can be inverted by a Neumann series for $|z - \lambda_0| < \delta_1$ and $|\kappa| < \delta_2$ if δ_1 and δ_2 are sufficiently small. Hence, $AR(z, \kappa)B^*$ also has a bounded extension $Q_1(z, \kappa)$ for $\operatorname{Im} z \neq 0$, which has completely meromorphic (meromorphic with finite rank principal parts at all poles [2]) continuations $Q_1^{\pm}(z, \kappa)$ from $\Omega^{\pm}$ to $|z - \lambda_0| < \delta_1$ satisfying

$$\begin{aligned} I - \kappa Q_1^{\pm}(z, \kappa) &= [I + \kappa Q^{\pm}(z)]^{-1} \\ &= \{I + \kappa(\lambda_0 - z)^{-1}[I + \kappa Q_c^{\pm}(z)]^{-1}F\}^{-1}[I + \kappa Q_c^{\pm}(z)]^{-1} . \end{aligned} \tag{1.3}$$

The poles of $Q_1^{\pm}(z, \kappa)$ need not be real, but for real κ do not lie in $\Omega^{\pm}$; *they are the resonances of this perturbation problem.*

THEOREM 1.1. *There is an analytic function $\Delta(z, \kappa)$ on a polydisc $\{(z, \kappa): |z - \lambda_0| < \delta_1, |\kappa| < \delta_2\}$ such that*

(a) *For $|\kappa| < \delta_2$, $\Delta(z, \kappa)$ has exactly m zeros $z_1(\kappa), \cdots, z_m(\kappa)$ (repeated according to multiplicity) in $|z - \lambda_0| < \delta_1$, which are precisely the poles of $Q_1^{+}(z, \kappa)$ in $|z - \lambda_0| < \delta_1$. For $\kappa = 0$, $z_j(0) = \lambda_0$ $(j = 1, \cdots, m)$.*

(b) *If for some real κ, $z_j(\kappa)$ is real, then $z_j(\kappa)$ is an eigenvalue of $H(\kappa)$ of multiplicity equal to the multiplicity $m_j(\kappa)$ of $z_j(\kappa)$ as a zero of $\Delta(z, \kappa)$.*

This result was proved in [2, § 5], except for analyticity of $\Delta(z, \kappa)$ which is clear from the construction of $\Delta(z, \kappa)$ (see equation (2.2) below). However, we have omitted the hypothesis of [2] that $Q(z)$ is *compact*. This can be done; for in [2] compactness was used only for two things: (a) to prove that $I + \kappa Q^{\pm}(z)$ has a completely meromorphic inverse, and (b) to prove, by references to [10], that $H(\kappa)$ is self-adjoint for real κ. However, we have argued above that (a) holds here, while (b) holds for κ sufficiently small [10, p. 59].

Note that [2] $F = AP_0[BP_0]^*$.

We shall now show that the resonances can be grouped into *cycles*, so that each of the p elements of a cycle is one of the values of a series expansion in powers of $\kappa^{1/p}$. Such series are known as *Puiseux series* [9, p. 130]. For their application to perturbation theory, see [6; Chapters II and VII].

THEOREM 1.2. *The resonances $z_1(\kappa), \cdots, z_m(\kappa)$ may be labeled so* *that each* $z_j(\kappa)$ *has a Puiseux series expansion in* κ. *If*

$$z_j(\kappa) = \lambda_0 + \alpha_1 \omega^j \kappa^{1/p} + \alpha_2 \omega^{2j} \kappa^{2/p} + \cdots \qquad (j = 1, \cdots, p) \tag{1.4}$$

is a given Puiseux cycle of resonances, where ω *is a primitive* pth *root of unity, then either the series has the form*

$$z_j(\kappa) = \lambda_0 + \alpha_p \kappa + \cdots + \alpha_{2np} \kappa^{2n} + \alpha_{2np+1} \omega^j \kappa^{2n+1/p} + \cdots \tag{1.5}$$

where $\lambda_0, \alpha_p, \cdots, \alpha_{(2n-1)p}$ *are real and* $\operatorname{Im} \alpha_{2np} < 0$, *or* $p = 1$ *and all the coefficients* α_n *are real.*

Moreover, the multiplicity $m_j(\kappa)$ *is independent of* κ *for* $\kappa \neq 0$ *and sufficiently small, and is the same for each element* $z_j(\kappa)$ *of a given Puiseux cycle.*

In particular, if $z_j(\kappa)$ belongs to a Puiseux cycle with $p \geqq 2$, then $z_j(\kappa)$ is not real for all sufficiently small real $\kappa \neq 0$. Thus any actual embedded eigenvalues of $H(\kappa)$ are analytic.

COROLLARY 1.3. *For real* $\kappa \neq 0$ *sufficiently small, the multiplicity of point eigenvalues in the interval* $(\lambda_0 - \delta_1, \lambda_0 + \delta_1)$ *is independent of* κ. *If for some* j, $z_j(\kappa)$ *is real for all sufficiently small* κ, *then* $z_j(\kappa)$ *is analytic in* κ.

Proof of Theorem 1.2. Since $\Delta(z, 0) = (\lambda_0 - z)^m$, the Weierstrass Preparation Theorem [1, p. 188] yields that

$$\Delta(z, \kappa) = [(z - \lambda_0)^m + g_{m-1}(\kappa)(z - \lambda_0)^{m-1} + \cdots + g_0(\kappa)] F(z, \kappa)$$

where $g_0, \cdots, g_{m-1}$ and F are analytic, $F(\lambda_0, 0) \neq 0$ and $g_0(0) = \cdots = g_{m-1}(0) = 0$. Thus $z_1(\kappa), \cdots, z_m(\kappa)$ are the zeros of a polynomial in z with coefficients analytic in κ, namely $\Delta(z, \kappa)/F(z, \kappa)$. Hence, (c.f. [6, pp. 63-66]) $z_1(\kappa), \cdots, z_m(\kappa)$ are algebroidal functions having at most an algebraic singularity at $\kappa = 0$, and must therefore have Puiseux series expansions. The statement about multiplicities is part of this theory.

Since $H(\kappa)$ is self-adjoint for real κ, $R(z, \kappa)$, and hence $Q_1^+(z, \kappa)$, is analytic for $\operatorname{Im} z > 0$, so that in the cycle (1.4), one has $\operatorname{Im} z_j(\kappa) \leqq 0$ for real κ, and each $j = 1, \cdots, p$. Therefore, the first term of (1.4) with a nonreal coefficient must have negative imaginary part for all real κ and $j = 1, \cdots, p$. But this can only happen for an even integer power κ^{2n} where, moreover, $\operatorname{Im} \alpha_{2np} < 0$. If all coefficients $\alpha_n \omega^{jn}$ are real, then because of the factor ω^{jn}, we can only have $p = 1$ or 2. However, if $p = 2$ and $\alpha_n \omega^{jn} \kappa^{n/2}$ is the first nonzero term with n odd, then changing κ into $-\kappa$ introduces a factor i, so that by proper choice of j, the imaginary part of this term can be made positive. Since this cannot occur, we must have $p = 1$.

REMARK. With perhaps a mild additional hypothesis, stationary scattering theory [8] shows that, for real κ, the absolutely continuous parts of $H(\kappa)$ and T in $(\lambda_0 - \delta_2, \lambda_0 + \delta_2)$ are unitarily equivalent.

2. Fermi's golden rule. In the simple case in which the perturbation B^*A removes the degeneracy at λ_0, calculation of the resonances up to terms of order κ^2 leads to the venerable Golden Rule for the line widths $\Gamma_j(\kappa)$. In order to discuss this, we must recall the construction of $\Delta(z, \kappa)$ [2, § 5].

It was proved in [2, p. 329; Theorem 3.1] that the residue of $Q^+(z)$ at λ_0 is $-AP_0[BP_0]^*$, where P_0 is the orthogonal projection onto $\ker(T - \lambda_0)$. Hence the operator

$$Q_c^+(z) = Q^+(z) - (\lambda_0 - z)^{-1}AP_0[BP_0]^* , \tag{2.1}$$

which corresponds to the continuous part of T near λ_0, is analytic on Ω. According to [2, p. 335; Theorem 5.1]

$$\Delta(z, \kappa) = (\lambda_0 - z)^m \det [I + [I + \kappa Q_c^+(z)]^{-1}\kappa(\lambda_0 - z)^{-1}AP_0[BP_0]^*] .$$

Using the formula $\det(I + ST) = \det(I + TS)$ [6, p. 162; Problem 4.17] gives

$$\Delta(z, \kappa) = (\lambda_0 - z)^m \det \{I + [BP_0]^*[I + \kappa Q_c^+(z)]^{-1}\kappa(\lambda_0 - z)^{-1}AP_0\} . \tag{2.2}$$

Now, A and B are one-one on $\mathscr{R}(P_0)$ and $\mathscr{R}([BP_0]^*) = \mathscr{R}(P_0)$ [2, p. 331]. We may therefore write (2.2) as a determinant on $\mathscr{R}(P_0)$, and then the factor $(\lambda_0 - z)^m$ may be taken inside the $m \times m$ determinant to yield

$$\begin{aligned} &\Delta(z, \kappa) \\ &\quad = \det \{(\lambda_0 - z)I_m + \kappa[BP_0]^*AP_0 - \kappa^2[BP_0]^*Q_c^+(z)AP_0 + O(\kappa^3)\} \end{aligned} \tag{2.3}$$

uniformly in z, where I_m is the identity on $\mathscr{R}(P_0)$ and $[I + \kappa Q_c^+(z)]^{-1}$ has been expanded in a Neumann series.

The operator $V_0 = [BP_0]^*AP_0$ maps $\mathscr{R}(P_0)$ into itself, and is essentially the compression of the perturbation B^*A to $\mathscr{R}(P_0)$. Using (1.1), we find that for $x, y \in \mathscr{H}$

$$\begin{aligned} (V_0x, y) &= ([BP_0]^*AP_0x, y) = (AP_0x, BP_0y) = (BP_0x, AP_0y) \\ &= ([AP_0]^*BP_0x, y) = (V_0^*x, y) \end{aligned}$$

which means that V_0 is self-adjoint on $\mathscr{R}(P_0)$. Therefore, with respect to a suitable orthonormal basis $\phi_1, \cdots, \phi_m$ of $\mathscr{R}(P_0)$, V_0 has a diagonal matrix

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix} .$$

The perturbation B^*A is said *to remove the degeneracy* at λ_0 iff the eigenvalues $\lambda_1, \cdots, \lambda_m$ to V_0 are all distinct. If $X(z)$ denotes the matrix with entries

$$X_{ij}(z) = -(Q_c^+(z)A\phi_i, B\phi_j)$$

then writing (2.3) with respect to the basis $\phi_1, \cdots, \phi_m$ yields finally

$$\Delta(z, \kappa) = \det\{(\lambda_0 - z)I_m + \kappa D + \kappa^2 X(z) + O(\kappa^3)\} \tag{2.4}$$

uniformly in z on a neighborhood of λ_0.

THEOREM 2.1. *If B^*A removes the degeneracy at λ_0, then $z_j(\kappa)$ is analytic ($j = 1, \cdots, m$) and*

$$z_j(\kappa) = \lambda_0 + \kappa\lambda_j + \kappa^2 X_{jj}(\lambda_0) + O(\kappa^3) . \tag{2.5}$$

Taking the imaginary part of (2.5) for real κ, we obtain formally

$$\begin{aligned} \Gamma_j(\kappa) &= -\operatorname{Im} z_j(\kappa) = -\kappa^2 \operatorname{Im}(Q_c^+(\lambda_0)A\phi_j, B\phi_j) + O(\kappa^3) \\ &= -\kappa^2 \operatorname{Im}(R_c(\lambda_0 + i0)V\phi_j, V\phi_j) + O(\kappa^3) \\ &= (2i)^{-1}\kappa^2([R_c(\lambda_0 - i0) - R_c(\lambda_0 + i0)]V\phi_j, V\phi_j) + O(\kappa^3) \end{aligned}$$

and hence finally

$$\Gamma_j(\kappa) = \pi\kappa^2(\delta_c(T - \lambda_0)V\phi_j, V\phi_j) + O(\kappa^3) \tag{2.6}$$

where $V = B^*A = A^*B$, $R_c(z) = R(z) - (\lambda_0 - z)^{-1}P_0$, and

$$\delta_c(T - \lambda) = (2\pi i)^{-1}[R_c(\lambda - iO) - R_c(\lambda + iO)] .$$

Formula (2.6) is *Fermi's Golden Rule.*

Proof of Theorem 2.1. We already know that $z_j(\kappa) = \lambda_0 + O(\kappa)$, and hence $X(z_j(\kappa)) = X(\lambda_0) + O(\kappa)$. If we define

$$\zeta_j(\kappa) = \kappa^{-1}(z_j(\kappa) - \lambda_0) .$$

Then the equation for $\zeta_j(\kappa)$ is, by (2.4),

$$\det\{-\kappa\zeta_j(\kappa)I_m + \kappa D + \kappa^2 X(\lambda_0) + O(\kappa^3)\} = 0 . \tag{2.7}$$

Expanding and dividing by κ^m gives

$$(\lambda_1 - \zeta_j(\kappa)) \cdots (\lambda_m - \zeta_j(\kappa)) + O(\kappa) = 0 . \tag{2.8}$$

Since the polynomial $(\lambda_1 - \zeta) \cdots (\lambda_m - \zeta)$ obtained for $\kappa = 0$ has distinct simple zeros, equation (2.8) has m analytic solutions, one asymptotic to each root as $\kappa \to 0$. Thus we may take

$$\zeta_j(\kappa) = \lambda_j + \beta_j\kappa + O(\kappa^2) \qquad (j = 1, \cdots, m) .$$

Setting $j = 1$ and substituting into (2.7), we find that

$$\det \{\kappa J + \kappa^2 X(\lambda_0) + O(\kappa^3)\} = 0$$

where

$$J = \begin{pmatrix} -\kappa\beta_1 & & & \\ & (\lambda_2 - \lambda_1) - \kappa\beta_1 & & \\ & & \ddots & \\ & & & (\lambda_m - \lambda_1) - \kappa\beta_1 \end{pmatrix} .$$

Expanding (2.7) gives

$$\kappa^{m+1}(\lambda_2 - \lambda_1) \cdots (\lambda_m - \lambda_1)(X_{11}(\lambda_0) - \beta_1) + O(\kappa^{m+2}) = 0$$

so that, in fact,

$$\beta_1 = X_{11}(\lambda_0) .$$

3. Spectral concentration.

The following theorem extends the main result of [3] to embedded eigenvalues.

THEOREM 3.1. *Assume that there exists a subspace $\mathscr{D}$ of $\mathscr{D}(A) \cap \mathscr{D}(B)$ such that $B\mathscr{D} \subset \mathscr{D}(A^*)$, $A\mathscr{D} \subset \mathscr{D}(B^*)$, and which is dense in $\mathscr{D}(A)$ and $\mathscr{D}(B)$ in the respective graph norms. For $j = 1, \cdots, m$ and κ real, choose $\delta_j(\kappa)$ such that $\delta_j(\kappa) = o(1)$ and* Im $z_j(\kappa) = o(\delta_j(\kappa))$ *as $\kappa \to 0$. Let*

$$S(\kappa) = \bigcup_{j=1}^{m} \{t\colon \operatorname{Re} z_j(\kappa) - \delta_j(\kappa) < t < \operatorname{Re} z_j(\kappa) + \delta_j(\kappa)\} .$$

If $H(\kappa) = \int \lambda dE_\kappa(\lambda)$, then

$$P_0 = st - \lim_{\kappa \to 0} \int_{S(\kappa)} dE_\kappa(\lambda) .$$

As shown in the appendix, the additional hypothesis insures that, for real κ, the poles of $Q_1^+(z, \kappa)$ are the complex conjugates of those of $Q_1^-(z, \kappa)$. Thus we did not need to take into account the poles of $Q_1^-(z, \kappa)$ when defining $S(\kappa)$, as was done for the corresponding set J_n in [3, Theorem 2.1]. In order that $\mathscr{D}$ exists, it is sufficient that either A or B be bounded, or that A and B be commuting self-adjoint operators.

Theorem 3.1 has a proof very similar to that of [3, Theorem 2.1], but cannot be deduced directly from that result because the operator $Q_1^+(z, \kappa)$, which corresponds to $Q_1^+(z, n)$ of [3], *tends to zero* as $\kappa \to 0$, and cannot, therefore, satisfy Hypothesis III (b) of [3]. To avoid repeating the lengthy arguments of [3], we shall simply carry the argument along to a point at which the arguments become essentially identical. A considerable study of [3] is therefore necessary to understanding the remainder of this section.

In order to surmount the difficulties posed by nonsimple poles, or poles close together, we shall show that for *real* κ, the resonances $z_1(\kappa), \cdots, z_m(\kappa)$ may be grouped into what we shall call *clusters* in such a way that, as $\kappa \to 0$, the resonances of a single cluster act together as a *single, simple* pole of $Q_1^+(z, \kappa)$, at least insofar as their asymptotic effect on the spectral measure of $H(\kappa)$ is concerned.

The result of our considerations is a rather detailed description of the singular part of $Q_1^+(z, \kappa)$.

In the first two lemmas, κ may be complex.

LEMMA 3.2. *Let $z_j(\kappa)\,(j = 1, \cdots, N)$ be the distinct poles of $Q_1^+(z, \kappa)$. Then $Q_1^+(z, \kappa)$ has the partial fraction expansion*

$$(3.1)\quad Q_1^+(z, \kappa) = \sum_{j=1}^{N} \frac{B_1^{(j)}(\kappa)}{(z - z_j(\kappa))} + \cdots + \frac{B_{m_j}^{(j)}(\kappa)}{(z - z_j(\kappa))^{m_j}} + L(z, \kappa)\ ,$$

where $L(z, \kappa)$ is analytic in z and κ. If $z_j(\kappa)$ has a Puiseux series expansion in powers of $\kappa^{1/p}$, then $B_k^{(j)}(\kappa)\,(k = 1, \cdots, m_j)$ also has an expansion in powers of $\kappa^{1/p}$, and has at most an algebraic pole at $\kappa = 0$.

The proof is a simple adaptation of the argument on pp. 69–70 of [6]. Certain additional facts obtained there do not hold here, since $Q_1^+(z, \kappa)$ is not a resolvent. Analyticity of $L(z, \kappa)$ is proved in the proof of the next lemma.

It follows immediately that for small $\kappa \neq 0$, $B_\kappa^{(j)}(\kappa)$ either vanishes identically or is never zero. Hence, for small $\kappa \neq 0$, *the order m_j of the jth pole $z_j(\kappa)$ of $Q_1^+(z, \kappa)$ is independent of κ.*

If the terms of the singular part of $Q_1^+(z, \kappa)$ in (3.1) are combined, we obtain

$$Q_1^+(z, \kappa) = \frac{P(z, \kappa)}{\Delta(z, \kappa)} + L(z, \kappa)$$

where $P(z, \kappa)$ is a polynomial in z with coefficients having at most an algebraic singularity at $\kappa = 0$, and $\Delta(z, \kappa)$ is the analytic function of z and κ defined in § 1.

LEMMA 3.3. (a) *As* $\kappa \to 0$, $Q_1^+(z, \kappa) \to Q^+(z)$ *uniformly on* $0 < \varepsilon \leqq |z - \lambda_0| \leqq \delta_2$ *for every* $\varepsilon > 0$.

(b) $P(z, \kappa)$, $\Delta(z, \kappa)$, *and* $L(z, \kappa)$ *are all analytic in* z *and* κ. *Moreover,*

$$\lim_{\kappa \to 0} P(z, \kappa) = (z - \lambda_0)^{N-1} AP_0[BP_0]^* . \tag{3.2}$$

Proof. From (1.3) and (2.1) one obtains

$$I - \kappa Q_1^+(z, \kappa) = [I + \kappa(\lambda_0 - z)^{-1}\Gamma(z, \kappa)AP_0[BP_0]^*]^{-1}\Gamma(z, \kappa) \tag{3.3}$$

where

$$\Gamma(z, \kappa) = [I + \kappa Q_c^+(z)]^{-1}$$

is analytic in z and κ, for κ and $z - \lambda_0$ small. Expanding the right side, canceling I on both sides and dividing by κ yields the result. Analyticity of $L(z, \kappa)$ and the coefficients of $P(z, \kappa)$, as well as (3.2) follow from the formulas between equations (2.7) and (2.8) of [3], where the discrete parameter n must be replaced by κ.

Assume now that κ is real, and write

$$z_j(\kappa) = \lambda_j(\kappa) - i\Gamma_j(\kappa) \qquad (j = 1, \cdots, N)$$

where $\lambda_j(\kappa)$ is real and $\Gamma_j(\kappa) \geqq 0$. We shall now describe the grouping of the $z_j(\kappa)$'s into *clusters*. To begin with, we specify that if $\Gamma_j(\kappa) \equiv 0$, then $z_j(\kappa)$ is to form a cluster by itself. Otherwise, $\Gamma_j(\kappa) > 0$ for small $\kappa \neq 0$, and we shall assume now for convenience that

$$\Gamma_j(\kappa) > 0 \qquad (j = 1, \cdots, N) .$$

Then $\Gamma_j(\kappa)$ has a Puiseux series, so that

$$\Gamma_j(\kappa) = a_j\kappa^{p(j)} + \cdots \tag{3.4}$$

where $a_j > 0$ and $p(j)$ is an integer ($j = 1, \cdots, m$). (If κ is complex in (3.4), $\Gamma_j(\kappa)$ is defined, but no longer the imaginary part of $-z_j(\kappa)$.) For $\kappa \neq 0$, choose $\delta_j(\kappa) > 0$ such that

$$\delta_j(\kappa) = o(\kappa^{p(j)-1}) \qquad (j = 1, \cdots, m)$$

while

$$\kappa^{p(j)} = o(\delta_j(\kappa)) \qquad (j = 1, \cdots, m)$$

as $\kappa \to 0$, and consider the intervals

$$J_j(\kappa) = (\lambda_j(\kappa) - \delta_j(\kappa), \lambda_j(\kappa) + \delta_j(\kappa)) .$$

If κ is small, the number of component intervals of

$$J_1(\kappa) \cup \cdots \cup J_m(\kappa) \tag{3.5}$$

is independent of κ, and each component is the union of the intervals $J_j(\kappa)$ corresponding to a certain set of resonances. For the distance between $\lambda_j(\kappa)$ and $\lambda_k(\kappa)$ is of the order of some integral power of κ, and is therefore either much greater or much less than the length of $J_j(\kappa)$. These sets are the *clusters*; they are independent of κ. We shall denote the components of (3.5) by

$$(c_j(\kappa) - \rho_j(\kappa),\ c_j(\kappa) + \rho_j(\kappa)) \qquad (j = 1, \cdots, N)$$

where N is the number of clusters. We shall refer to $c_j(\kappa)$ and $\rho_j(\kappa)$ as the *center* and *radius* of the jth cluster.

It is easily seen that if $\{z_1(\kappa), \cdots, z_{p_1}(\kappa)\}$ is the first cluster, then

$$\lambda_j(\kappa) - c_1(\kappa) = o(\rho_1(\kappa)) \qquad (j = 1, \cdots, p_1)\,. \tag{3.6}$$

For if $\lambda_j(\kappa)$ and $\lambda_k(\kappa)$ belong to the first cluster, the distance between them is much less than either $\delta_j(\kappa)$ or $\delta_k(\kappa)$, neither of which can exceed $\rho_1(\kappa)$. Similarly

$$\rho_i(\kappa) = o(|\,c_1(\kappa) - c_2(\kappa)\,|) \qquad (i = 1, 2) \tag{3.7}$$

because $c_1(\kappa) - c_2(\kappa)$, being determined by the $\lambda_j(\kappa)$'s, is of integral power order, while $\rho_j(\kappa)$, being determined by the $\delta_j(\kappa)$'s is not.

Similar statements hold for other clusters. The interpretation of (3.6) is that the resonances of a cluster are asymptotically very close to the center of the corresponding interval $(c_n - \rho_n, c_n + \rho_n)$, while (3.7) says that distinct components of (3.5) are asymptotically very small compared to their distance apart.

LEMMA 3.4. *For* $\operatorname{Im} z > 0$, *and* $|\,z - \lambda_0\,| \leqq \delta_2$

$$\|\,P(z, \kappa)\,\| \leqq C|\,\Delta(z, \kappa)\,|\,(\operatorname{Im} z)^{-1}$$

where C *is independent of* κ.

Proof. For each κ, the coefficients of $P(z, \kappa)$ are of finite rank, since they are residues of functions with singular parts of finite rank, and are also analytic in κ. The lemma therefore follows by a proof similar to that of equation (2.8) of [3].

The procedures of [3] could now be applied to yield an asymptotic expansion for the singular part $P(z, \kappa)/\Delta(z, \kappa)$ of $Q_1^+(z, \kappa)$. However, we shall be content to remark that for any sequence $\kappa_n \to 0$, the quantities $P(z, \kappa_n)$, $\Delta(z, \kappa_n)$, etc. have precisely the properties of $P_n(z)$, $\Delta_n(z)$ etc. which are used in the proof of [3, Theorem 2.1] from equation (2.10) of [3] onward. The remainder of the proof of Theorem 3.1 follows [3] with essentially no change.

4. Examples. We shall now consider some simple examples which illustrate certain phenomena.

EXAMPLE 1. We shall first give an example in which *a nonanalytic Puiseux series occurs*. Let $\mathscr{H} = L_2(-\infty, +\infty) \oplus \mathbb{C}^2$, and let e_1, e_2 be the usual orthonormal basis of $\mathbb{C}^2$. Define

$$H_0\begin{pmatrix} u(t) \\ \xi \end{pmatrix} = \begin{pmatrix} t & 0 \\ 0 & c \end{pmatrix}\begin{pmatrix} u(t) \\ \xi \end{pmatrix} = \begin{pmatrix} tu(t) \\ c\xi \end{pmatrix}$$

where $u \in L_2(-\infty, +\infty)$, $\xi \in \mathbb{C}^2$ and c is a fixed real number. $H_0 = T$ has absolutely continuous spectrum of simple multiplicity, except for an embedded eigenvalue c of multiplicity $m = 2$. Let $f_1(t)$, $f_2(t)$ be an *orthonormal* pair of functions in $L_2(-\infty, +\infty)$, and define an operator Y from $\mathbb{C}^2$ into $L_2(-\infty, +\infty)$ by

$$Y(\xi_1 e_1 + \xi_2 e_2) = \xi_1 f_1(t) + \xi_2 f_2(t) .$$

The operator Y^* from $L_2(-\infty, +\infty)$ back into $\mathbb{C}^2$ is then

$$Y^* u = \left(\int u(t)\bar{f}_1(t)dt\right)e_1 + \left(\int u(t)\bar{f}_2(t)dt\right)e_2 .$$

We shall consider the perturbed operator

$$H(\kappa) = H_0 + \kappa V$$

where

$$V = \begin{pmatrix} 0 & Y \\ Y^* & \lambda_1 I \end{pmatrix}$$

and $\lambda_1 > 0$. The perturbation V is self-adjoint of rank 4, and its range has the orthonormal basis $\{f_1, f_2, e_1, e_2\}$. If we choose the factorization

$$V = VP = PV$$

where P is the orthogonal projection onto the range of V, then the matrix of

$$Q(z) = V(H_0 - z)^{-1}P$$

with respect to the orthonormal basis f_1, f_2, e_1, e_2 of the range of V is

$$\begin{pmatrix} 0 & (c - z)^{-1}I_2 \\ F(z) & (c - z)^{-1}\lambda_1 I_2 \end{pmatrix}$$

where

$$F(z) = \int (t - z)^{-1} \begin{pmatrix} |f_1(t)|^2 & \bar{f}_1(t)f_2(t) \\ f_1(t)\bar{f}_2(t) & |f_2(t)|^2 \end{pmatrix} dt$$

and I_2 is the 2×2 identity matrix.

If we now assume that $F(z)$ has a meromorphic continuation from the upper half-plane across the axis in a neighborhood of c, then the equation

$$(c - z)^2 \det (I + \kappa Q(z)) = 0$$

for the resonances reduces to

$$\kappa^4 D(z) - \kappa^2 T(z)(c + \kappa\lambda_1 - z) + (c + \kappa\lambda_1 - z)^2 = 0$$

where $T(z)$ and $D(z)$ are the trace and determinant of $F(z)$. Solving for $(c + \kappa\lambda_1 - z)^{-1}$ by the quadratic formula yields

$$z = c + \lambda_1\kappa + \kappa^2 g(z)$$

where

$$g(z) = -\frac{1}{2}(T(z) \pm \sqrt{T^2(z) - 4D(z)}) .$$

For simplicity, let us now take $c = 0$. Then, if the function

$$H(z) = T^2(z) - 4D(z)$$

has a simple zero at $z = 0$, the function $g(z)$ has a Puiseux series expansion

$$g(z) = a_0 + a_1 z^{1/2} + a_2 z + \cdots$$

where $a_1 \neq 0$. It then follows easily from

$$z = \lambda_1\kappa + \kappa^2(a_0 + a_1 z^{1/2} + a_2 z + \cdots)$$

that

$$z = \lambda_1\kappa + a_0\kappa^2 + a_1\lambda_1^{1/2}\kappa^{5/2} + O(\kappa^3)$$

which means that $z(\kappa)$ has a nonanalytic Puiseux series in κ. We shall therefore have obtained the desired example, if we can find $f_1(t)$ and $f_2(t)$ such that $H(z)$ has a *simple* zero at $z = 0$.

To this end, let

$$f_1(t) = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{t^2 + 1}$$

and

$$\begin{aligned} f_2(t) &= (2 - 2\varepsilon)^{-1/2} \operatorname{sgn} t \qquad && 0 < \varepsilon < |t| < 1 \\ &= 0 && \text{otherwise} . \end{aligned}$$

Then f_1 and f_2 are an orthonormal pair, and since they are real,

$$F_{12}(z) = F_{21}(z) .$$

The values of $F_{11}(0)$ and $F'_{11}(0)$ may be computed from

$$F_{11}(z) = -(z + 2i)(z + i)^{-2} \qquad \operatorname{Im} z > 0$$

while due to the fact that $f_2(t)$ vanishes near the origin, the integrals for $F_{12}(0)$ and $F_{22}(0)$, as well as those obtained for $F'_{12}(0)$ and $F'_{22}(0)$ by differentiation under the integral sign are absolutely convergent. In fact, one has

$$F_{22}(0) = (2 - 2\varepsilon)^{-1}\int_{\varepsilon<|t|<1} \frac{dt}{t} = 0$$

and

$$F'_{22}(0) = (2 - 2\varepsilon)^{-1}\int_{\varepsilon<|t|<1} \frac{dt}{t^2} = \varepsilon^{-1} .$$

Similarly,

$$F_{12}(0) = 2(\pi - \pi\varepsilon)^{-1/2}\int_{\varepsilon}^{1} \frac{1}{t^2+1} \cdot \frac{dt}{t}$$

and

$$F'_{12}(0) = 0 .$$

Hence, one computes that

$$\begin{aligned} H(0) &= (F_{11}(0) - F_{22}(0))^2 + 4F^2_{12}(0) \\ &= -4 + 16(\pi - \pi\varepsilon)^{-1}\left\{\int_{\varepsilon}^{1} \frac{1}{t^2+1}\,\frac{dt}{t}\right\}^2 \end{aligned}$$

and

$$\begin{aligned} H'(0) &= 2(F_{11}(0) - F_{22}(0))(F'_{11}(0) - F'_{22}(0)) + 8F_{12}(0)F'_{12}(0) \\ &= -4i(3 + \varepsilon^{-1}) \neq 0 . \end{aligned}$$

It therefore remains to choose ε such that $H(0) = 0$; that is, such that

$$\left(\frac{\pi}{4}\right)^{1/2} = (1 - \varepsilon)^{-1/2}\int_{\varepsilon}^{1} \frac{1}{t^2+1}\,\frac{dt}{t} \equiv \Phi(\varepsilon) .$$

But since $\Phi(\varepsilon)$ is decreasing on $0 < \varepsilon < 1$, $\Phi(0+) = +\infty$, and $\Phi(1-) = 0$, there is a unique ε in the interval $0 < \varepsilon < 1$ satisfying this equation.

Finally, note that the Puiseux series appears here as a degenerate case, since in the usual case when $H(z)$ does not vanish at the origin, $g(z)$ and hence $z(\kappa)$, have two distinct analytic branches.

EXAMPLE 2. An example will now be given of an eigenvalue of multiplicity one embedded at an *end point* of the continuous spectrum, and perturbed by an operator of rank two, which gives rise to a resonance or an eigenvalue which cannot be represented as a Puiseux series. The endpoint appears as a *branch point* of $Q^+(z)$. Branch points of continued quantities occur in Simon's articles [14, 15] as "*thresholds*" for certain processes (that is, the minimum energies at which the processes can occur). His theory excludes eigenvalues embedded at thresholds—with good reason, as this example shows. Most of the thresholds in [14, 15] are *embedded* in a continuous spectrum, rather than at an end point. An example of this along the present lines would be easily constructed. The example is similar to Example 8.3 of [5, p. 581]. The operator $H_0 = T$ on $L_2(0, \infty) \oplus \mathbb{C}$ defined by

$$H_0[u(t), \xi] = [tu(t), 0]$$

has absolutely continuous spectrum $[0, \infty)$ and an eigenvalue at $\lambda_0 = 0$ with eigenvector

$$\phi_0 = [0, 1] .$$

Let $H(\kappa) = H_0 + \kappa V$ where

$$V[u(t), \xi] = [\xi f(t), (u, f) + \lambda_1 \xi] .$$

We assume that $\lambda_1 > 0$ and

$$\int_0^\infty |f(t)|^2 dt = 1 .$$

The perturbation V has rank 2, so the resonances are to be sought as poles of an analytic continuation of the inverse of the matrix $W(z, \kappa)$ of the restriction of $I + \kappa V(H_0 - z)^{-1}$ to the range $\mathscr{R}(V)$ of V. Computing $W(z, \kappa)$ with respect to the orthonormal basis ϕ_0, f of $\mathscr{R}(V)$, one obtains [5; eq. (8.9), p. 581]

$$W(z, \kappa) = \begin{pmatrix} 1 & -\kappa z^{-1} \\ \kappa F(z) & 1 - \kappa\lambda_1 z^{-1} \end{pmatrix}$$

where

$$F(z) = \int_0^\infty |f(t)|^2 (t - z)^{-1} dt .$$

If we assume that $F(z)$ has a continuation $F_+(z)$ from the upper half-plane across the *positive* real axis, then the resonances satisfy the equation

$$z = \kappa\lambda_1 - \kappa^2 F_+(z) . \tag{4.1}$$

(See [5, p. 581], the third equation from the bottom of the page—in which there is an error of sign.)

Now choose

$$|f(t)|^2 = \frac{2}{\pi}\frac{1}{1+t^2} \tag{4.2}$$

so that

$$F(z) = \frac{2i - (2/\pi)\log z - z}{1+z^2}$$

where $0 < \arg z < 2\pi$. The solution of (4.1) then has the asymptotic expansion

$$z(\kappa) = \kappa\lambda_1 + (2/\pi)\kappa^2 \log(\kappa\lambda_1) - 2i\kappa^2 + O(\kappa^3) \tag{4.3}$$

which is not of Puiseux type. For $\kappa < 0$, $z(\kappa)$ lies in the region $0 < \arg z < 2\pi$, and is therefore a negative eigenvalue $\lambda(\kappa)$ of $H(\kappa)$, with the expansion

$$\lambda(\kappa) = \kappa\lambda_1 + (2/\pi)\kappa^2 \log(-\kappa\lambda_1) + O(\kappa^3)\kappa < 0 .$$

For $\kappa > 0$, the continuation $F_+(z)$ of $F(z)$ leads to the solution $z_+(\kappa)$ with $\arg z_+(\kappa) \cong 0$, while if $F_+(z)$ is replaced in (4.1) by the continuation $F_-(z)$ of $F(z)$ from the lower half-plane, one obtains the solution $z_-(\kappa)$ with $\arg z_-(\kappa) \cong 2\pi$. These numbers are complex conjugates. If κ is complex, the first situation essentially prevails, in the sense that the non-self-adjoint operator $H(\kappa)$ has an eigenvalue at $z(\kappa)$ for all sufficiently small κ in any given sector $|\arg\kappa - \pi| \leqq \pi - \delta$, $\delta > 0$.

If instead of (4.2), one chooses

$$|f(t)|^2 = \frac{2}{\pi}\cos(\pi\alpha/2)\frac{t^\alpha}{1+t^2} \tag{4.4}$$

where $-1 < \alpha < 1$, then one obtains, for $\alpha \neq 0$,

$$F(z) = \frac{\cot(\pi\alpha/2) - \csc(\pi\alpha/2) z^\alpha e^{-i\pi\alpha} - z}{1+z^2}$$

where $0 < \arg z < 2\pi$. The solution of (4.2) then has the expansion

$$z(\kappa) = \kappa\lambda_1 - \kappa^2\cot(\pi\alpha/2) + \kappa^{2+\alpha}e^{-i\pi\alpha}\lambda_1^\alpha \csc(\pi\alpha/2) + O(\kappa^3) . \tag{4.5}$$

This has the same general behavior: for $\kappa > 0$, there is an eigenvalue $\lambda(\kappa)$ with expansion

$$\lambda(\kappa) = \kappa\lambda_1 - \kappa^2\cot(\pi\alpha/2) + (-\kappa)^{2+\alpha}\lambda_1^\alpha \csc(\pi\alpha/2) + O(\kappa^3)$$

while for $\kappa > 0$, there is a resonance. A notable feature, however, is that one may obtain a Puiseux series by taking, for example, $\alpha = \pm 1/2$, in which case $W(z, \kappa)$ has only an algebraic singularity at $z = 0$. In fact there are only two sheets, and it is interesting to note that for $\kappa < 0$, these is a pole on the second sheet directly below the eigenvalue $\lambda(\kappa)$.

Let us see what becomes of Fermi's Golden Rule in this case. One has

$$\langle \delta_c(H_0 - \lambda)\bar{V}\phi_0, \bar{V}\phi_0 \rangle = |f(\lambda)|^2 .$$

(See [5, eq. (8.7)]. Note that, in the notation of [5], the V_1 term contributes nothing.) Hence, Fermi's Rule gives

$$\Gamma(\kappa) \cong \pi\kappa^2 |f(\lambda_0)|^2 .$$

Applied to the case $\lambda_0 = 0$ with $f(t)$ given by (4.4), this gives the following results: (a) for $\alpha = 0$

$$\Gamma(\kappa) \cong 2\kappa^2$$

which agrees with (4.3); (b) for $\alpha > 0$

$$\Gamma(\kappa) \cong 0$$

which agrees with (4.5), to order κ^2, but is not informative; (c) for $\alpha < 0$, $\Gamma(\kappa)$ is infinite, which is not surprising because according to (4.5), $\Gamma(\kappa)$ is *not* $O(\kappa^2)$. The Gold from which the Rule is made is apparently mixed with Brass.

If, however, λ_0 is replaced in the Rule by $\lambda_0 + \kappa\lambda_1$, the resulting formula

$$\Gamma(\kappa) \cong \pi\kappa^2 \langle \delta_c(H_0 - \lambda_0 - \kappa\lambda_1) V\phi_0, V\phi_0 \rangle \tag{4.6}$$

is an unalloyed success; for one then obtains

$$\Gamma(\kappa) \cong \pi\kappa^2 |f(\kappa\lambda_1)|^2 \cong 2\lambda_1^\alpha \kappa^{2+\alpha} \cos(\pi\alpha/2)$$

which agrees with (4.5).

APPENDIX. Let T be self-adjoint and suppose that for some pair of vectors f, g the function

$$r(z) = ((T - z)^{-1} f, g)$$

has meromorphic continuations $r_\pm(z)$ across some interval of the real axis. That the poles of $r_-(z)$ need not be the complex conjugates of the poles of $r_+(z)$ may be seen by taking $Tu(t) = tu(t)$ on $L_2(-\infty, +\infty)$ and choosing $f(t) = (t + i)^{-1}$ and $g(t) = (t - i)^{-1}$. Then $r_+(z)$ has a pole at $z = -i$, while $r_-(z)$ vanishes identically.

Similarly, the poles of $Q_1^+(z)$ and $Q_1^-(z)$ are not always conjugate. For $A = (\cdot, f)f$ and $B = (\cdot, g)g$ are bounded and self-adjoint, and $AB = BA = 0$ because f and g are orthogonal. Hence, $H = T + B^*A = T$, and

$$Q_1(z) = Q(z) = (G(z)f, g)(\cdot, g)g = r(z)(\cdot, f)g$$

so that $Q_1^+(z)$ has a pole at $z = -i$ while $Q_1^-(z)$ vanishes identically.

We shall give sufficient conditions that $Q_1^+(z)$ and $Q_1^-(z)$ have conjugate poles. Let T, A, and B satisfy the hypotheses of § 1, and assume that $Q_1^\pm(z)$ defined by

$$I - Q_1^\pm(z) = [I + Q^\pm(z)]^{-1}$$

is meromorphic, and has finite rank principal parts at all its poles. This is true, for example, if κ is small in § 1, or if $Q^\pm(z)$ is compact. Formula (1.2) (with $\kappa = 1$) then defines the resolvent $R(z)$ of an extension H of $T + B^*A$, and $Q_1(z)$ is the extension of $AR(z)B^*$. (It is not clear whether or not H is self-adjoint in this generality, but this is not at issue.) By taking adjoints, [7, eq. (2.2)] one also finds that $BG(z)A^*$ has the compact extension

$$\tilde{Q}(z) = [Q(\bar{z})]^*$$

which has the continuations

$$\tilde{Q}^\pm(z) = [Q^\pm(z)]^* \tag{1}$$

defined on Ω. Similarly, $BR(z)A^*$ leads to $\tilde{Q}_1(z)$ and $\tilde{Q}_1^\pm(z)$.

THEOREM. *In addition to the hypotheses above, suppose that there exists a subspace $\mathscr{D}$ of $\mathscr{D}(A) \cap \mathscr{D}(B)$ such that $B\mathscr{D} \subset \mathscr{D}(A^*)$, $A\mathscr{D} \subset \mathscr{D}(B^*)$, and $\mathscr{D}$ is dense in $\mathscr{D}(A)$ and $\mathscr{D}(B)$ respectively, in the graph norms. If $Q^+(z)$ is analytic at z_0, then $Q_1^+(z)$ is analytic at z_0 iff $\tilde{Q}_1^+(z)$ is analytic at z_0.*

Proof. Let P_A and P_B be the orthogonal projections onto the closures of the ranges of A and B. Then $I - P_B$ projects onto ker B^*, so that

$$P_AQ(z) = Q(z) \quad \text{and} \quad Q(z)[I - P_B] = 0$$

for $\operatorname{Im} z > 0$, and hence by continuation

$$P_AQ^+(z) = Q^+(z) \tag{2}$$

and

$$Q^+(z)P_B = Q^+(z)\,. \tag{3}$$

Observe next that by (1.1),

$$B^*Ax = A^*Bx \qquad x \in \mathscr{D} .$$

Hence, for x, $g \in \mathscr{D}$, and $\operatorname{Im} z > 0$, one has

$$\begin{aligned}(\tilde{Q}_1(z)Bx, Ay) &= (BR(z)A^*Bx, Ay)\\ &= (BR(z)B^*Ax, Ay) = (AR(z)B^*Ax, By)\\ &= (Q_1(z)Ax, By)\end{aligned}$$

where (1.1) was used in the equality next to last. Using that $\mathscr{D}$ is dense in the graphs, and passing to a continuation shows that analyticity of $P_A\tilde{Q}_1^+(z)P_B$ at z_0 is equivalent to analyticity of $P_BQ_1^+(z)P_A$ at z_0.

If we now assume that $Q^+(z)$ and $\tilde{Q}_1^+(z)$ are analytic at z_0, then since (1), together with (2) and (3), implies that

$$\begin{aligned}Q_1^+(z) &= Q^+(z) - [Q^+(z)]^2 + Q^+(z)Q_1^+(z)Q^+(z)\\ &= Q^+(z) - [Q^+(z)]^2 + Q^+(z)P_BQ_1^+(z)P_AQ^+(z)\end{aligned}$$

it follows that $Q_1^+(z)$ is also analytic at z_0. The other implication is proved similarly.

It is evident from the proof that if the ranges of A and B are dense, the assumption that $Q^+(z)$ is analytic at z_0 may be dropped. However, the example above shows that it cannot be dropped in general.

COROLLARY. *If all poles of $Q^+(z)$ are real, then the nonreal poles of $Q_1^+(z)$ and $Q_1^-(z)$ are complex conjugates.*

This follows from (1).

PROPOSITION. *Either of the following conditions suffices for the existence of $\mathscr{D}$.*

(a) *Either A or B is bounded.*

(b) *A and B are commuting self-adjoint operators.*

Proof. If A is bounded, it follows from (1.1) that $A\mathscr{D}(B) \subset \mathscr{D}(B^*)$. Hence, one may take $\mathscr{D} = \mathscr{D}(B)$. Similarly if B is bounded.

Sufficiency of (b) follows easily from [12, p. 358].

REFERENCES

1. S. Bochner and W. T. Martin, *Several Complex Variables*, Princeton University Press, 1948.

2. J. S. Howland, *On the Weinstein-Aronszajn formula*, Arch. Rat. Mech. Anal., **39** (1970), 323-339.

3. J. S. Howland, *Spectral concentration and virtual poles, II*, Trans. Amer. Math. Soc., **162** (1971), 141-156.

4. ———, *Perturbation of embedded eigenvalues*, Bull. Amer. Math. Soc., **78** (1972), 280-283.

5. ———, *Embedded eigenvalues and virtual poles*, Pacific J. Math., **29** (1969), 565-582.

6. T. Kato, *Perturbation Theory for Linear Operators*, Berlin-Heidelberg-New York, Springer, 1966.

7. ———, *Wave operators and similarity for some non-self-adjoint operators*, Math. Ann., **162** (1966), 258-279.

8. T. Kato and S. T. Kuroda, *The abstract theory of scattering*, Rocky Mountain J. Math., **1** (1971), 127-171.

9. K. Knopp, *Theory of Functions, Part II*, New York, Dover, 1947.

10. R. Konno and S. T. Kuroda, *On the finiteness of perturbed eigenvalues*, J. Fac. Sci. Univ. Tokyo Sect. I, **13** (1966), 55-63.

11. P. D. Lax and R. S. Phillips, *Scattering Theory*, Academic Press, New York-London, 1967.

12. F. Riesz and B. Sz-Nagy, *Functional Analysis*, Ungar, New York, 1955.

13. N. A. Shenk and D. Thoe, *Eigenfunction expansions and scattering theory for perturbations of* $-\Delta$, Rocky Mountain J. Math., **1** (1971), 89-125.

14. B. Simon, *Convergence of time dependent perturbation theory for autoionizing states of atoms*, Phys. Letters A., **36** (1971), 23-24.

15. ———, *Resonances in N-body quantum systems with dilatation analytic potentials and the foundations of time-dependent perturbation theory*, Ann. Math., **97** (1973), 247-274.

Received March 23, 1973. Supported by DA-ARO-31-124-71-G182.

UNIVERSITY OF VIRGINIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

LINEAR *GCD* EQUATIONS

DAVID JACOBSON

Let R be a GCD domain. Let A be an $m \times n$ matrix and B an $m \times 1$ matrix with entries in R. Let $c \neq 0$, $d \in R$. We consider the linear GCD equation $GCD(AX + B, c) = d$. Let S denote its set of solutions. We prove necessary and sufficient conditions that S be nonempty. An element t in R is called a solution modulus if $X + tR^n \subseteq S$ whenever $X \in S$. We show that if c/d is a product of prime elements of R, then the ideal of solution moduli is a principal ideal of R and its generator t_0 is determined. When R/t_0R is a finite ring, we derive an explicit formula for the number of distinct solutions (mod t_0) of $GCD\ (AX + B, c) = d$.

1. Introduction. Let R be a GCD domain. As usual GCD $(a_1, \cdots, a_m)$ will denote a greatest common divisor of the finite sequence of elements $a_1, \cdots, a_m$ of R.

Let A be an $m \times n$ matrix with entries a_{ij} in R and let B be an $m \times 1$ matrix with entries b_i in R for $i = 1, \cdots, m$; $j = 1, \cdots, n$. Let $c \neq 0$, d be elements of R. In this paper we consider the "linear GCD equation"

$$GCD(a_{11}x_1 + \cdots + a_{1n}x_n + b_1, \cdots, \\ a_{m1}x_1 + \cdots + a_{mn}x_n + b_m, c) = d \,. \tag{1.1}$$

Letting X denote the column of unknows $x_1, \cdots, x_n$ in (1.1), we shall find it convenient to abbreviate the equation (1.1) in matrix notation by

$$GCD(AX + B, c) = d \,. \tag{1.2}$$

Of course we allow a slight ambiguity in viewing (1.1) as an equation, since the GCD is unique only up to a unit.

Let R^n denote the set of $n \times 1$ matrices with entries in R. We let $S \equiv S(A, B, c, d)$ denote the set of all solutions of (1.1), that is

$$S = \{X \in R^n \mid GCD(AX + B, c) = d\} \,.$$

If S is nonempty, we say that (1.1) or (1.2) is solvable. Note that X satisfies $GCD(AX + B, d) = d$ if and only if X is a solution of the linear congruence system $AX + B \equiv 0 (\mathrm{mod}\ d)$.

We show in Proposition 1 that if (1.1) is solvable, then $d \mid c$, $AX + B \equiv 0 (\mathrm{mod}\ d)$ has a solution and $GCD(A, d) = GCD(A, B, c)$. Here $GCD(A, d) = GCD(a_{11}, \cdots, a_{1n}, \cdots, a_{m1}, \cdots, a_{mn}, d)$ and $GCD(A, B, c) = GCD(A, b_1, \cdots, b_m, c)$. Conversely we show in Proposition 3 that if the above conditions hold and $e = c/d$ is atomic, that is e is a product of prime elements of R, then (1.1) is solvable. (Also see Proposition 4).

Let the solution set S of (1.1) be nonempty. We say that t in R is a solution modulus of (1.1) if given X in S and $X \equiv X'(\mathrm{mod}\, t)$, then X' is in S. We let $M \equiv M(A, B, c, d)$ denote the set of all solution moduli of (1.1). We show in Theorem 2 that M is an ideal of R and if $e = c/d$ is atomic, then M is actually a principal ideal generated by $d/g(p_1 \cdots p_k)$, where $g = GCD(A, d)$ and $\{p_1, \cdots, p_k\}$ is a maximal set of nonassociated prime divisors of e such that for each p_i, the system $AX + B \equiv 0(\mathrm{mod}\, dp_i)$ is solvable. This generator $d/g(p_1 \cdots p_k)$ denoted by t_0 is called the minimum modulus of (1.1).

In § 4 we assume that R/t_0R is a finite ring and we derive an explicit formula for the number of distinct equivalence classes of $R^n(\mathrm{mod}\, t_0)$ comprising S. We denote this number by $N_{t_0} \equiv N_{t_0}(A, B, c, d)$. Let $A' = A/g$ and $d' = d/g$. Let $L = \{X + d'R^n \mid A'X \equiv 0(\mathrm{mod}\, d')\}$ and $L_i = \{X + d'R^n \mid A'X \equiv 0(\mathrm{mod}\, d'p_i)\}$ for $i = 1, \cdots, k$. In Theorem 3 we show that

$$N_{t_0} = |L| \prod_{i=1}^{k} (|R/p_iR|^n - |R/p_iR|^{n-(r_i+s_i)}) \tag{1.3}$$

where r_i is rank $A'(\mathrm{mod}\, p_i)$ and s_i is the dimension of the R/p_iR vector space L/L_i.

The formula (1.3) is applied in some important cases. For example in Corollary 6 we determine N_{t_0} when R is a principal ideal domain.

This paper is an extension and generalization to GCD domains, of the results obtained over the ring of integers $\boldsymbol{Z}$ in [2].

2. Solvability of $GCD\ (AX + B, c) = d$.

PROPOSITION 1. *If $GCD\ (AX + B, c) = d$ is solvable, then the following conditions hold.*

(2.1) (i) $d \mid c$,

(ii) $AX + B \equiv 0(\mathrm{mod}\, d)$ *is solvable,*

(iii) $GCD(A, d) = GCD(A, B, c)$.

Proof. Let X satisfy $GCD(AX + B, c) = d$. Then clearly (i) $d \mid c$ and (ii) $AX + B \equiv 0(\mathrm{mod}\, d)$. Let $AX + B = dU$ where U is an $m \times 1$ matrix with entries u_i for $i = 1, \cdots, m$. Then $GCD(dU, c) = GCD(du_1, \cdots, du_m, c) = d$. Let $g = GCD(A, d)$ and $h = GCD(A, B, c)$. Then $B \equiv 0(\mathrm{mod}\, g)$ as $AX - dU = B$ and $g \mid c$ as $d \mid c$, which shows that $g \mid h$. Also $dU \equiv 0(\mathrm{mod}\, h)$, so that $h \mid GCD(dU, c)$, that is $h \mid d$. Thus $h \mid g$, which proves (iii).

PROPOSITION 2. *Let e in R have the following property*

(I) *$GCD(AX + B, e) = 1$ is solvable whenever $GCD(A, B, e) = 1$. Suppose that $c = de$, $AX + B \equiv 0 (\bmod d)$ is solvable and $GCD(A, d) = GCD(A, B, c)$. Then $GCD(AX + B, c) = d$ is solvable.*

Proof. There exist X' in R^n and V in R^m such that $AX' + B = dV$. Let $g = GCD(A, d)$ and let A' denote the matrix with entries a_{ij}/g and B' the matrix with entries b_i/g for $i = 1, \cdots, m$; $j = 1, \cdots, n$. Then $A'X' + B' = d'V$ where $d' = d/g$. We claim that $GCD(A', V, e) = 1$. For let h be any divisor of $GCD(A', V, e)$. Then $B' \equiv 0 (\bmod h)$ and $h \mid GCD(A', B', c')$ where $c' = d'e$. However, $GCD(A', B', c') = 1$ as $g = GCD(A, B, c)$. Hence h is a unit, that is $GCD(A', V, e) = 1$. So by property (I), there is a Y in R^n such that $GCD(A'Y + V, e) = 1$. Thus $GCD(A(d'Y) + dV, de) = d$ and if we set $X = X' + d'Y$, then $GCD(AX + B, c) = d$, establishing the proposition.

We show in Proposition 3 that if e is atomic, then e satisfies property (I).

We require the following useful lemmas.

LEMMA 1. *Let $e = p_1 \cdots p_k$ be a product of nonassociated prime elements $p_1, \cdots, p_k$ in R. If $GCD(A, B, e) = 1$, then $GCD(AX + B, e) = 1$ is solvable.*

Proof. Let $GCD(A, B, e) = 1$. We use induction on k. Let $k = 1$. If $GCD(B, p_1) = 1$, then $X = 0$ satisfies $GCD(AX + B, p_1) = 1$. Suppose that $B \equiv 0 (\bmod p_1)$. Then $GCD(A, p_1) = 1$. Hence there is a j such that $GCD(a_{1j}, \cdots, a_{mj}, p_1) = 1$. Let X^j in R^n have a 1 in the jth position and o's elsewhere. Then $GCD(AX^j + B, p_1) = GCD(AX^j, p_1) = 1$. Thus $GCD(AX + B, p_1) = 1$ is solvable. Now let $k > 1$ and let $e' = p_1 \cdots p_{k-1}$. By the induction assumption there is X' in R^n such that $GCD(AX' + B, e') = 1$. Let $B' = AX' + B$. We claim that $GCD(Ae', B', p_k) = 1$. If $GCD(A, p_k) = 1$, then $GCD(Ae', B', p_k) = 1$. Suppose that $A \equiv 0 (\bmod p_k)$. If $B' \equiv 0 (\bmod p_k)$, then $B \equiv 0 (\bmod p_k)$, contradicting the hypothesis that $GCD(A, B, e) = 1$. Hence $GCD(B', p_k) = 1$, establishing the claim. So there exists a Y in R^n such that $GCD((Ae')Y + B', p_k) = 1$. Let $X = X' + e'Y$. Then $X \equiv X' (\bmod e')$ yields that $AX + B \equiv B' (\bmod e')$. Thus $GCD(AX + B, e') = 1$ since $GCD(B', e') = 1$. Also

$$GCD(AX + B, p_k) = GCD((Ae')Y + B', p_k) = 1 ,$$

so that $GCD(AX + B, e'p_k) = 1$, completing the proof.

LEMMA 2. *Suppose that e is an atomic element of R.*

(*) *Let* $\{p_1, \cdots, p_k\}$ *be a maximal set of nonassociated prime divisors of* e *such that for each* p_i, *the system* $AX + B \equiv 0(\mathrm{mod}\ dp_i)$ *is solvable.*

Then X *is a solution of* $GCD(AX + B, c) = d$ *if and only if* $GCD(AX + B, de_0) = d$, *where* $c = de$ *and* $e_0 = p_1 \cdots p_k$.

Proof. Since e is atomic, it is clear that we may select a set $\{p_1, \cdots, p_k\}$ as defined in (*). If this set is empty, we let $e_0 = 1$. Suppose that X satisfies $GCD(AX + B, c) = d$. Then there is U in R^m such that $AX + B = dU$ and $GCD(U, e) = 1$. Since $e_0 \mid e$, $GCD(U, e_0) = 1$ and thus $GCD(dU, de_0) = d$, that is, $GCD(AX + B, de_0) = d$.

Conversely let X satisfy $GCD(AX + B, de_0) = d$. Then $AX + B = dU$ and $GCD(U, e_0) = 1$. Suppose there is a prime $p \mid e$ and $U \equiv 0(\mathrm{mod}\ p)$. Then $AX + B \equiv 0(\mathrm{mod}\ dp)$ and the maximal property of the set $\{p_1, \cdots, p_k\}$ shows that p is an associate of some p_i. So $U \equiv 0(\mathrm{mod}\ p_i)$, contradicting that $GCD(U, e_0) = 1$. Hence $GCD(U, p) = 1$ for all primes $p \mid e$ and thus $GCD(U, e) = 1$, that is $GCD(AX + B, c) = d$.

PROPOSITION 3. *Suppose that* $c = de$, $AX + B \equiv 0(\mathrm{mod}\ d)$ *is solvable and* $GCD(A, d) = GCD(A, B, c)$. *If* e *is atomic, then* $GCD(AX + B, c) = d$ *is solvable.*

Proof. Let e be atomic. By Proposition 2 it suffices to show that e satisfies property (I). Thus let $GCD(A_0, B_0, e) = 1$ where A_0 is an $m \times n$ matrix and B_0 is an $m \times 1$ matrix. By Lemma 2, $GCD(A_0X + B_0, e) = 1$ is solvable if and only if $GCD(A_0X + B_0, e_0) = 1$ is solvable where $e_0 = p_1 \cdots p_k$ is a product of nonassociated prime divisors of e. However by Lemma 1, $GCD(A_0X + B_0, e_0) = 1$ is solvable since $GCD(A_0, B_0, e_0) = 1$. Thus (I) holds and $GCD(AX + B, c) = d$ is solvable.

THEOREM 1. *Let* R *be a GCD domain. Consider the following condition*

(II) $GCD(a_1x + b_1, \cdots, a_mx + b_m, c) = 1$ *is solvable if*

$$GCD(a_1, \cdots, a_m, b_1, \cdots, b_m, c) = 1;$$

(i) *If* R *satisfies* (II), *then* $GCD(AX + B, c) = 1$ *is solvable whenever* $GCD(A, B, c) = 1$.

(ii) *If* R *is a Bezout domain such that* $GCD(ax + b, c) = 1$ *is solvable whenever* $GCD(a, b, c) = 1$, *then* R *satisfies* (II).

Proof. (i) Let R satisfy (II). Let $GCD(A, B, c) = 1$ where A

is an $m \times n$ matrix. We prove that $GCD(AX + B, c) = 1$ is solvable by induction of n. For $n = 1$, solvability is granted by the supposition (II). Let $n > 1$ and let A' denote the $m \times (n-1)$ matrix with entries $a_{i,j+1}$ for $i = 1, \cdots, m$; $j = 1, \cdots, n-1$. If $c' = GCD(a_{11}, \cdots, a_{1m}, c)$, then $GCD(A', B, c') = 1$. Hence by the induction assumption, there exist $x_2, \cdots, x_n$ in R such that $GCD(a_{12}x_2 + \cdots + a_{1n}x_n + b_1, \cdots, a_{m1}x_2 + \cdots + a_{mn}x_n + b_m, c') = 1$. If $b'_i = a_{i2}x_2 + \cdots + a_{in}x_n + b_i$ for $i = 1, \cdots, m$, then $GCD(a_{11}, \cdots, a_{m1}, b'_1, \cdots, b'_m, c) = 1$. Thus by (II), there exists x_1 in R such that $GCD(a_{11}x_1 + b'_1, \cdots, a_{m1}x_1 + b'_m, c) = 1$. So if X in R^n has entries $x_1, x_2, \cdots, x_n$, then $GCD(AX + B, c) = 1$, completing the proof of (i).

(ii) Let R be a Bezout domain, that is a domain in which every finitely generated ideal is principal. Suppose that R has the property that $GCD(ax + b, c) = 1$ is solvable if $GCD(a, b, c) = 1$. Let

$$GCD(a_1, \cdots, a_m, b_1, \cdots, b_m, c) = 1 .$$

Let A and B denote the $m \times 1$ matrices with entries $a_1, \cdots, a_m$ and $b_1, \cdots, b_m$ respectively. Then by [3, Theorem 3.5], there exists an invertible $m \times m$ matrix P such that PA has entries $a, 0, \cdots, 0$. Also it is clear that $GCD(PA, PB, c) = 1$. Let PB have entries $b, b'_2, \cdots, b'_m$. Thus by hypothesis, $GCD(ax + b, c') = 1$ is solvable where $c' = GCD(b'_2, \cdots, b'_m, c)$. Hence $GCD(Ax + B, c) = 1$ is solvable, that is R satisfies (II).

As an immediate consequence of the preceding propositions and Theorem 1, we state

PROPOSITION 4. *Let R be a UFD or a Bezout domain such that $GCD(ax + b, c) = 1$ is solvable if $GCD(a, b, c) = 1$. Then $GCD(AX + B, c) = d$ is solvable if and only if $d \mid c$, $AX + B \equiv 0 (\mathrm{mod}\ d)$ is solvable and $GCD(A, d) = GCD(A, B, c)$.*

We remark that we do not know whether there exists a *GCD* domain in which (II) is not valid. Any Bezout domain satisfying (II) is an elementary divisor domain [3, Theorem 5.2].

We conclude this section with the following result.

PROPOSITION 5. *Let R be a Bezout domain. Suppose that* (0) *$GCD(ax + b, c) = 1$ is solvable whenever $GCD(a, b) = 1$ and $a \mid c$. Then $GCD(ax + b, c) = 1$ is solvable whenever $GCD(a, b, c) = 1$.*

Proof. Let $GCD(a, b, c) = 1$. If $a' = GCD(a, c)$, then $GCD(a', b) = 1$ and $a' \mid c$. By the assumption (0), there is x' in R such that $GCD(a'x' + b, c) = 1$. If $u = a'x' + b$, then $a' \mid (u - b)$ and since R is a Bezout domain, there is an x in R such that $ax + b \equiv u (\mathrm{mod}\ c)$.

Thus $GCD(ax + b, c) = 1$ since $GCD(u, c) = 1$.

Let $a \mid c$ and let $\nu: R/cR \to R/aR$ be the epimorphism given by $\nu(r + cR) = r + aR$ for all r in R. Let G(resp. G') denote the group of units of R/cR(resp. R/aR). If $\nu': G \to G'$ is the induced homomorphism, then note that (0) is equivalent to the condition that $\nu'(G) = G'$. (See [5].)

3. The minimum modulus. Let the solution set S of $GCD(AX + B, c) = d$ be nonempty. Then

$$M = \{t \in R \mid X + tR^n \subseteqq S \text{ for all } X \in S\}$$

is the set of solution moduli of $GCD(AX + B, c) = d$.

Note that $c \in M$ for if $X \in S$ and $X \equiv X' (\text{mod } c)$, then $AX + B \equiv AX' + B (\text{mod } c)$, so that $d = GCD(AX' + B, c)$.

It is obvious that $M = R$, that is $S = R^n$ if and only if $d = GCD(A, d) = GCD(A, B, c)$ and $GCD(A/d(X) + B/d, c/d) = 1$ for all X in R^n.

THEOREM 2. *Let R be a GCD domain. Let $GCD(AX + B, c) = d$ be solvable. Let $e = c/d = \prod_{i=1}^{k} e_i$. Let $\hat{e}_i = e_1 \cdots e_{i-1}e_{i+1} \cdots e_k$ for $i = 1, \cdots, k$.*

(1) *M is an ideal of R,*

(2) *$M \supseteqq \bigcap_{i=1}^{k} M_i$ where M_i is the ideal of solution moduli for $GCD(AX + B, de_i) = d$.*

(3) *If each $\hat{e}_i$ satisfies property* (I) *of Proposition* 2, *then $M = \bigcap_{i=1}^{k} M_i$ and M is a principal ideal if each M_i is principal.*

(4) *If e is atomic, then M is a principal ideal generated by $d/g(p_1 \cdots p_k)$ where $g = GCD(A, d)$ and $\{p_i, \cdots, p_k\}$ is defined in* (*) *of Lemma* 2.

Proof.

(1) As S is nonempty, the set M is well-defined and o, c belong to M. Let t_1, t_2 be in M and let $r \in R$. Let $X \in S$ and let $Y \in R^n$. Then $X + t_1Y \in S$ and hence $(X + t_1Y) + t_2(-Y) \in S$, that is $X + (t_1 - t_2)Y \in S$ which shows that $t_1 - t_2 \in M$. Also $X + t_1(rY) \in S$, that is $X + (t_1r)Y \in S$. So $t_1r \in M$ and thus M is an ideal of R.

(2) As $d \mid c$ we let $c = de$. Let S_i denote the solution set of $GCD(AX + B, de_i) = d$ where $e = \prod_{i=1}^{k} e_i$. Then clearly $S = \bigcap_{i=1}^{k} S_i$. Let $t \in \bigcap_{i=1}^{k} M_i$. Let $X \in S$ and let $Y \in R^n$. Then $X + tY \in \bigcap_{i=1}^{k} S_i$ since $X \in \bigcap_{i=1}^{k} S_i$. So $X + tY \in S$, that is $t \in M$, which proves that $M \supseteqq \bigcap_{i=1}^{k} M_i$.

(3) Assume that each $\hat{e}_i$ satisfies property (I). We prove that $M \subseteqq M_i$ for $i = 1, \cdots, k$. As $g = GCD(A, d) = GCD(A, B, c)$, let $A' = A/g$, $B' = B/g$, and $d' = d/g$. Let i be fixed and let $X_i \in S_i$.

Then $A'X_i + B' = d'U$ where $GCD(U, e_i) = 1$. We claim that $GCD(e_iA', U, \hat{e}_i) = 1$. For let h be a divisor of $GCD(e_iA', U, \hat{e}_i)$. Then $A' \equiv 0 (\mathrm{mod}\, h)$ since $GCD(h, e_i) = 1$. Thus $h \mid GCD(A', B', d'e)$, that is $h \mid 1$. So by assumption there exists X' in R^n such that

$$GCD((e_iA')X' + U, \hat{e}_i) = 1 .$$

Let $X = X_i + d'e_iX'$. Then for $j = 1, \cdots, k$,

$$\begin{aligned} &GCD(A'X + B', d'e_j) \\ &\qquad = d'\, GCD((e_iA')X' + U, e_j) = d' . \end{aligned}$$

Hence $X \in \bigcap_{j=1}^k S_j$, that is $X \in S$. Now let $t \in M$ and let $Y \in R^n$. Then $X + tY \in S$ and so $X + tY \in S_i$. However, $X + tY \equiv X_i + tY (\mathrm{mod}\, d'e_i)$ and thus $X_i + tY \in S_i$, that is $t \in M_i$, which proves that $M \subseteq M_i$. So by (2), $M = \bigcap_{i=1}^k M_i$. Moreover, if each M_i is a principal ideal, say $M_i = t_iR$, then $\bigcap_{i=1}^k M_i$ is a principal ideal generated by the $LCM(t_1, \cdots, t_k)$.

(4) Let t be any element of M. We show that $d/g \mid t$ where $g = GCD(A, d)$. First note that S is the solution set of $GCD(A'X + B', d'e) = d'$ where $A' = A/g$, $B' = B/g$, and $d' = d/g$. Let $X \in S$ and let $A'X + B' = d'U$. Then $GCD(A'(X + tY) + B', d'e) = d'$ for all Y in R^n. So $GCD((A't)Y + d'U, d'e) = d'$ and thus $(A't)Y \equiv 0 (\mathrm{mod}\, d')$ for all Y in R^n. Hence $A't \equiv 0(\mathrm{mod}\, d')$ and since $GCD(A', d') = 1$, it follows that $d' \mid t$.

Now suppose that e is atomic. By Lemma 2, S is also the solution set of $GCD(A'X + B', d'e_0) = d'$ where $e_0 = p_1 \cdots p_k$ and $\{p_1, \cdots, p_k\}$ is defined in (*). Thus M is also the ideal of solution moduli of $GCD(A'X + B', d'e_0) = d'$. Let M_i' denote the ideal of solution moduli of $GCD(A'X + B', d'p_i) = d'$ for $i = 1, \cdots, k$. Then Lemma 1 shows that (3) can be applied to yield that $M = \bigcap_{i=1}^k M_i'$. We prove that each M_i' is a principal ideal generated by $d'p_i$. Clearly $d'p_i \in M_i'$ for $i = 1, \cdots, k$. Let i be fixed and let t be any element in M_i'. Then as shown earlier, $d' \mid t$ say $t = d't'$. By (*) there exists X in R^n such that $A'X + B' \equiv 0(\mathrm{mod}\, d'p_i)$. Thus $GCD(A', p_i) = 1$ since $GCD(A', B', d'e) = 1$. So there is a j for which $GCD(A'E_j, p_i) = 1$ where E_j is the $n \times 1$ matrix with 1 in the jth position and o's elsewhere.

Now assume that $GCD(t', p_i) = 1$. Let $X' = X + tE_j$. Then $GCD(A'(X' - X), d'p_i) = d'\, GCD(t'A'E_j, p_i) = d'$ since $GCD(t'A'E_j, p_i) = 1$. So $GCD(A'X' - A'X, d'p_i) = d'$ and thus $GCD(A'X' + B', d'p_i) = d'$ as $B \equiv -A'X (\mathrm{mod}\, d'p_i)$. Hence $GCD(A'(X' + t(-E_j)) + B', d'p_i) = d'$ since $t \in M_i'$. That is $GCD(A'X + B', d'p_i) = d'$ and thus $d'p_i \mid d'$, which contradicts that p_i is a nonunit. So the assumption that $GCD(t', p_i) = 1$ is untenable, that is $p_i \mid t'$. Thus $d'p_i \mid t$ proving that

$M_i' = d'p_iR$. However $M = \bigcap_{i=1}^{k} M_i'$, so that M is a principal ideal generated by the $LCM(d'p_1, \cdots, d'p_k)$, that is M is generated by $d'p_1 \cdots p_k$.

The generator $d'p_1 \cdots p_k$ of M is called the minimum modulus of $GCD(AX + B, de) = d$.

4. The number of solutions with respect to a modulus. Let $GCD(AX + B, c) = d$ be solvable where $e = c/d$ is atomic. If t in R is a solution modulus of $GCD(AX + B, c) = d$, then S consists of equivalence classes of $R^n(\mathrm{mod}\, t)$. If R/tR is also a finite ring, we let $N_t \equiv N_t(A, B, c, d)$ denote the number of distinct equivalence classes of $R^n(\mathrm{mod}\, t)$ comprising S.

For R/tR finite, let $|t| = |R/tR|$ denote the number of elements in R/tR. Note that if $t_0 \mid t$, then each equivalence class of $R^n(\mathrm{mod}\, t_0)$ consists of $|t/t_0|^n = (|t|/|t_0|)^n$ classes of $R^n(\mathrm{mod}\, t)$. Thus if t is a solution modulus and t_0 denotes the mininum modulus of $GCD(AX + B, c) = d$, then $N_t = |t/t_0|^n N_{t_0}$. In Theorem 3, we explicitly determine N_{t_0}.

The following lemma is also of independent interest.

LEMMA 3. *Let R be a GCD domain and suppose that R/dR is a finite ring. Let $p_1, \cdots, p_k$ be nonassociated elements such that R/p_iR is a finite field for $i = 1, \cdots, k$. Let A be an $m \times n$ matrix and let r_i denote the rank of $A(\mathrm{mod}\, p_i)$ for $i = 1, \cdots, k$. Let $\mathscr{L} = \{X \in R^n \mid AX \equiv 0(\mathrm{mod}\, d)\}$ and $L = \{X + dR^n \mid X \in \mathscr{L}\}$. Let $e_0 = \prod_{i=1}^{k} p_i$ and let $\mathscr{L}' = \{X \in R^n \mid AX \equiv 0(\mathrm{mod}\, de_0)\}$ and $L' = \{X + de_0R^n \mid X \in \mathscr{L}'\}$. Let $\mathscr{L}_i = \{X \in R^n \mid AX \equiv 0(\mathrm{mod}\, dp_i)\}$ and $L_i = \{X + dR^n \mid X \in \mathscr{L}_i\}$ for $i = 1, \cdots, k$. Let $H = \{X + e_0R^n \mid X \in \mathscr{L}'\}$ and $H_i = \{X + p_iR^n \mid X \in \mathscr{L}_i\}$ for $i = 1, \cdots, k$. Then*

(1) $$|L'| = |L|\,|H|$$

and

$$|H| = \prod_{i=1}^{k} |H_i| \,.$$

(2) *L/L_i is an R/p_iR vector space of dimension s_i and $|H_i| = |R/p_iR|^{n-(r_i+s_i)}$ for $i = 1, \cdots, k$.*

(3) *$s_i = 0$ if and only if for each X in $\mathscr{L}$ there exists X' in $\mathscr{L}_i$ such that $X' \equiv X(\mathrm{mod}\, d)$.*

(4) *If $GCD(d, p_i) = 1$, then $s_i = 0$.*

(5) *$|L| = 1$ if and only if $n = \mathrm{rank}\, A(\mathrm{mod}\, p)$ for each prime $p \mid d$.*

Proof.

(1) In the obvious way, L, L', and H are R-modules. Let $\sigma: L' \to H$ denote the R-homomorphism defined by $\sigma(X + de_0R^n) = X + e_0R^n$ for all X in $\mathscr{L}'$. Then clearly $\operatorname{Ker}\sigma = \{e_0Y + de_0R^n \mid Y \in \mathscr{L}\}$ so that $L \cong \operatorname{Ker}\sigma$ under the R-isomorphism $\tau: L \to \operatorname{Ker}\sigma$ defined by $\tau(Y + dR^n) = e_0Y + de_0R^n$ for all Y in $\mathscr{L}$. Thus $|L'| = |L||H|$ since $\operatorname{Im}\sigma = H$. We now show that H is isomorphic to $\bigoplus_{i=1}^k H_i$, the direct sum of the R-modules H_i. Let $\gamma: H \to \bigoplus_{i=1}^k H_i$ denote the R-homomorphism defined by $\gamma(X + e_0R^n) = (X + p_1R^n, \cdots, X + p_kR^n)$ for all X in $\mathscr{L}'$. If $X + e_0R^n \in \operatorname{Ker}\gamma$, then $X \equiv 0 (\operatorname{mod} p_i)$ for $i = 1, \cdots, k$, that is $X \equiv 0(\operatorname{mod} e_0)$, which shows that γ is $1 - 1$. To show that $\operatorname{Im}\gamma = \bigoplus_{i=1}^k H_i$, let $X_i \in \mathscr{L}_i$ for $i = 1, \cdots, k$. Since R/dR is finite, it is easy to verify that d is atomic. Thus let $d = d_0 \prod_{i=1}^k p_i^{m_i}$ where $m_i \geqq 0$ and $GCD(d_0, p_i) = 1$. By the Chinese remainder theorem there exists X in R^n such that $X \equiv 0(\operatorname{mod} d_0)$ and $X \equiv X_i(\operatorname{mod} p_i^{m_i+1})$ for $i = 1, \cdots, k$. However, $AX_i \equiv 0(\operatorname{mod} p_i^{m_i+1})$ for $i = 1, \cdots, k$, so that $AX \equiv 0 \operatorname{mod} (d_0 \prod_{i=1}^k p_i^{m_i+1})$, that is $AX \equiv 0(\operatorname{mod} de_0)$. Thus $X + e_0R^n \in H$ and $\gamma(X + e_0R^n) = (X_1 + p_1R^n, \cdots, X_k + p_kR^n)$. Hence γ is an isomorphism and $|H| = \prod_{i=1}^k |H_i|$.

(2) Let $L_i' = \{X + dp_iR^n \mid X \in \mathscr{L}_i\}$ for $i = 1, \cdots, k$. Let i be fixed. Let $\nu: L_i' \to L_i$ denote the R-homomorphism defined by $\nu(X + dp_iR^n) = X + dR^n$ for all X in $\mathscr{L}_i$. Then clearly $\operatorname{Ker}\nu = \{dY + dp_iR^n \mid AY \equiv 0(\operatorname{mod} p_i)\}$ and it follows that

$$|\operatorname{Ker}\nu| = |R/p_iR|^{n-r_i} \equiv |p_i|^{n-r_i}$$

where $r_i = \operatorname{rank} A(\operatorname{mod} p_i)$. Thus $|L_i'| = |p_i|^{n-r_i}|L_i|$ since $\operatorname{Im}\nu = L_i$. However by (1), $|L_i'| = |L||H_i|$. Also since L_i is an R-submodule of L, the quotient module L/L_i is defined and $|L| = |L_i||L/L_i|$. Thus we obtain that $|H_i||L/L_i| = |p_i|^{n-r_i}$. We now show that L/L_i is an R/p_iR vector space. Let $\langle X\rangle = X + dR^n$ for X in R^n. Then $L/L_i = \{\langle X\rangle + L_i \mid X \in \mathscr{L}\}$. For r in R, let $\bar{r} = r + p_iR$ in R/p_iR. We define $\bar{r}(\langle X\rangle + L_i) = \langle rX\rangle + L_i$ for all r in R and X in $\mathscr{L}$. We claim that this is a well-defined R/p_iR multiplication on L/L_i. For let $\bar{r} = \bar{r}'$ and $\langle X\rangle + L_i = \langle X'\rangle + L_i$, where $r, r' \in R$ and $X, X' \in \mathscr{L}$. Then $r - r' \equiv o(\operatorname{mod} p_i)$ and $\langle X\rangle - \langle X'\rangle \in L_i$, that is $\langle X - X'\rangle \in L_i$. Thus there exists Y in $\mathscr{L}_i$ such that $\langle X - X'\rangle = \langle Y\rangle$. We must show that $\langle rX\rangle + L_i = \langle r'X'\rangle + L_i$, that is $\langle rX - r'X'\rangle \in L_i$. We write $rX - r'X' = (r - r')X + r'(X - X')$. However, $X - X' \equiv Y(\operatorname{mod} d)$ and thus $r(X - X') \equiv rY(\operatorname{mod} d)$. So $rX - r'X' \equiv (r - r')X + rY(\operatorname{mod} d)$ and $(r - r')X + rY \in \mathscr{L}_i$. Hence $\langle rX - r'X'\rangle \in L_i$, which establishes the claim. It follows immediately that L/L_i is an R/p_iR vector space since L/L_i is an R-module.

Let s_i denote the dimension of the R/p_iR vector space L/L_i.

Then $|L/L_i| = |p_i|^{s_i}$ and as $|H_i||L/L_i| = |p_i|^{n-r_i}$, we obtain that $|H_i||p_i|^{s_i} = |p_i|^{n-r_i}$. Thus $o \leqq s_i \leqq n - r_i$ and $|H_i| = |p_i|^{n-(r_i+s_i)}$, which completes the proof of (2).

(3) As $|L| = |L_i||p_i|^{s_i}$, it is immediate that $s_i = 0$ if and only if $L = L_i$, that is if and only if for each X in $\mathscr{L}$ there exists X' in $\mathscr{L}_i$ such that $X' \equiv X (\text{mod } d)$.

(4) Suppose that $GCD(d, p_i) = 1$. Let $X \in \mathscr{L}$. By the Chinese remainder theorem there exists X' in R^n such that $X' \equiv X(\text{mod } d)$ and $X' \equiv 0(\text{mod } p_i)$. Thus $AX' \equiv 0(\text{mod } dp_i)$, so that $s_i = o$ by (3).

(5) Let p be a prime dividing d and let $d = d_1 p$. Then $L = \{X + d_1 pR^n \mid X \in \mathscr{L}\}$. However as shown in the proof of (2), $|L| = |p|^{n-r_0}|L_0|$ where $r_0 = \text{rank } A(\text{mod } p)$ and $L_0 = \{X + d_1 R^n \mid X \in \mathscr{L}\}$. Thus if $|L| = 1$, then $n = \text{rank } A(\text{mod } p)$ for any prime $p \mid d$. The converse is trivial.

THEOREM 3. *Let R be a GCD domain. Let $GCD(AX + B, c) = d$ be solvable and suppose that $e = c/d$ is atomic. Let $A' = A/g$ and $d' = d/g$ where $g = GCD(A, d)$. Let $t_0 = d' \prod_{i=1}^k p_i$ denote the minimum modulus of $GCD(AX + B, c) = d$ where $\{p_1, \cdots, p_k\}$ is defined in (*) of Lemma 2. Suppose that R/t_0R is a finite ring. Let $L = \{X + d'R^n \mid A'X \equiv 0(\text{mod } d')\}$ and $L_i = \{X + d'R^n \mid A'X \equiv 0(\text{mod } d'p_i)\}$ for $i = 1, \cdots, k$. Then*

$$N_{t_0} = |L| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-(r_i+s_i)}) \tag{4.1}$$

where r_i denotes rank $A'(\text{mod } p_i)$ *and s_i denotes the dimension of the R/p_iR vector space L/L_i.*

Proof. Let S denote the solution set of $GCD(AX + B, c) = d$. As $g = GCD(A, B, c)$, let $B' = B/g$. Then by Lemma 2, S is also the solution set of $GCD(A'X + B, d'e_0) = d'$ where $e_0 = \prod_{i=1}^k p_i$. Let $\mathscr{S}$ denote the set of X in R^n such that $A'X + B' \equiv 0(\text{mod } d')$. Let $\mathscr{S}_i$ denote the set of X in R^n such that $A'X + B' \equiv 0(\text{mod } d'p_i)$ for $i = 1, \cdots, k$. It is clear that $S = \mathscr{S} \backslash \bigcup_{i=1}^k \mathscr{S}_i$. Let $T_0 = \{X + t_0R^n \mid X \in S\}$. Then $|T_0|$ is what we have denoted by N_{t_0}. Also let $T = \{X + t_0R^n \mid X \in \mathscr{S}\}$ and $T_i = \{X + t_0R^n \mid X \in \mathscr{S}_i\}$ for $i = 1, \cdots, k$. Hence $T_0 = T \backslash \bigcup_{i=1}^k T_i$ and by the method of inclusion and exclusion

$$N_{t_0} = |T_0| = \sum_I (-1)^{|I|} |T_I| \tag{4.2}$$

where the summation is over all subsets I of

$$I_k = \{1, \cdots, k\} \text{ and } T_I = \bigcap_{i=1} T_i \,.$$

Now let $\mathscr{S}_I = \bigcap_{i \in I} \mathscr{S}_i$ and $d'_I = d' \prod_{i \in I} p_i$ for each subset I of

I_k. Then it is easy to see that $\mathscr{S}_I$ is the set of X in R^n such that $A'X + B' \equiv 0(\text{mod } d'_I)$ and $T_I = \{X + t_0R^n \mid X \in \mathscr{S}_I\}$. Let $T'_I = \{X + d'_IR^n \mid X \in \mathscr{S}_I\}$ and let $I' = I_k \backslash I$. Then $|T_I| = |T'_I| \prod_{i \in I'} |p_i|^n$, since $X + d'_IR^n$ consists of $|t_0/d'_I|^n = \prod_{i \in I'} |p_i|^n$ distinct classes of $R^n(\text{mod } t_0)$.

Let $\mathscr{L}_I$ denote the set of X in R^n such that $A'X \equiv 0(\text{mod } d'_I)$. Let $L'_I = \{X + d'_IR^n \mid X \in \mathscr{L}_I\}$. As $\mathscr{S}_i$ is nonempty for $i = 1, \cdots, k$, an argument involving the Chinese remainder theorem shows that each $\mathscr{S}_I$ is nonempty. Hence it follows that $|T'_I| = |L'_I|$. Let $L = \{X + d'R^n \mid X \in \mathscr{L}_\phi\}$ and $L_i = \{X + d'R^n \mid X \in \mathscr{L}_{\{i\}}\}$ for $i = 1, \cdots, k$. Then (1) and (2) of Lemma 3 yield that $|L'_I| = |L| \prod_{i \in I} |p_i|^{n-(r_i+s_i)}$ where $r_i = \text{rank } A'(\text{mod } p_i)$ and $s_i =$ dimension of the R/p_iR vector space L/L_i.

Hence by (4.2),

$$N_{t_0} = |L| \sum_I (-1)^{|I|} \prod_{i \in I} |p_i|^{n-(r_i+s_i)} \prod_{i \in I'} |p_i|^n$$

where the summation is over all subsets I of I_k and $I' = I_k \backslash I$. Thus we may write

$$N_{t_0} = |L| \prod_{i=1}^{k} |p_i|^n \sum_I (-1)^{|I|} \prod_{i \in I} |p_i|^{-(r_i+s_i)}$$

where the summation is over all subsets I of I_k. However,

$$\prod_{i=1}^{k} (1 - |p_i|^{-(r_i+s_i)}) = \sum_I (-1)^{|I|} \prod_{i \in I} |p_i|^{-(r_i+s_i)} ,$$

which yields the formula (4.1) for N_{t_0}. This completes the proof of the theorem.

We remark that if $p_i^{m_i}$ is the highest power of p_i dividing d', then s_i is also the dimension of the R/p_iR vector space K_i^0/K_i where $K_i^0 = \{X + p_i^{m_i}R^n \mid A'X \equiv 0(\text{mod } p_i^{m_i})\}$ and

$$K_i = \{X + p_i^{m_i}R^n \mid A'X \equiv 0(\text{mod } p_i^{m_i+1})\} .$$

Also note that $r_i \geqq 1$ for $i = 1, \cdots, k$.

In Corollaries 1 and 2, the notation is the same as in Theorem 3.

COROLLARY 1. *Let $GCD(AX + B, c) = d$ be solvable and suppose that $e = c/d$ is atomic. Let R/t_0R be finite where $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(AX + B, c) = d$.*

(i) *If $GCD(d', e) = 1$, then*

$$N_{t_0} = |L| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-r_i}) . \tag{4.3}$$

(ii) *If $|L| = 1$, then*

$$N_{t_0} = \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-r_i}) , \tag{4.4}$$

where $r_i = n$ *if* $p_i \mid d'$.

(iii) *If* $n' = \text{rank } A' (\text{mod } p_i)$ *for* $i = 1, \cdots, k$, *where* n' *denotes the smaller of* m *and* n, *then*

$$N_{t_0} = |L| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-n'}) . \tag{4.5}$$

(iv) $N_{t_0} = 1$ *if and only if* (a) $|L| = 1$ *and there exists no prime* $p \mid e$ *such that* $AX + B \equiv 0 (\text{mod } dp)$ *is solvable, or* (b) $n = 1$ *and* $|p| = 2$ *for any prime* $p \mid e$ *such that* $AX + B \equiv 0 (\text{mod } dp)$ *is solvable.*

Proof.

(i) If $GCD(d', p_i) = 1$, then (4) of Lemma 3 shows that $s_i = 0$ in (4.1). Hence if $GCD(d', e) = 1$, then $s_i = 0$ for $i = 1, \cdots, k$, which yields (4.3).

(ii) Suppose that $|L| = 1$. If $p_i \mid d'$, then $n = r_i$ by (5) of Lemma 3 and thus $s_i = 0$ since $s_i \leqq n - r_i$. However if $GCD(d', p_i) = 1$, then $s_i = 0$, so that (4.4) is immediate from (4.1).

In particular if $d = 1$, then N_{t_0} is given by (4.4). If A' is invertible (mod d'), then (4.4) also applies.

(iii) If $n = r_i$, then $s_i = 0$. If $m = r_i$, then the criterion in (3) shows that $s_i = 0$. Thus (4.5) follows from (4.1).

(iv) Suppose that $N_{t_0} = 1$. Then by (4.1), $|L| = 1$ and thus $s_i = 0$ for $i = 1, \cdots, k$. If p_i is a prime dividing e such that $AX + B \equiv 0 (\text{mod } dp_i)$ is solvable, then $|p_i|^n - |p_i|^{n-r_i} = 1$, so that $n = r_i = 1$ and $|p_i| = 2$. Thus either (a) or (b) holds. Conversely if (a) holds, then $N_{t_0} = 1$. If $n = 1$, then clearly $|L| = 1$ and hence (b) implies that $N_{t_0} = 1$.

COROLLARY 2. *Let* $GCD(AX + B, c) = d$ *be solvable and let* $e = c/d$. *Suppose that* R/cR *is a finite ring. Then*

$$N_c = |L| \, |ge|^n \prod_{i=1}^{k} (1 - |p_i|^{-(r_i + s_i)}) . \tag{4.6}$$

Proof. Since R/cR is finite, e is atomic. Thus $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(AX + B, c) = d$. Also R/t_0R is finite since $t_0 \mid c$, so that N_{t_0} is given by (4.1). However $N_c = |c/t_0|^n N_{t_0}$, which yields the result (4.6).

COROLLARY 3. *Suppose that* R/cR *is a finite ring. Then* $GCD(a_1x_1 + \cdots + a_nx_n + b, c) = d$ *is solvable if and only if* $d \mid c$ *and* $GCD(a_1, \cdots, a_n, d) = GCD(a_1, \cdots, a_n, b, c)$. *Let* $a'_j = a_j/g$ *for* $j = 1, \cdots, n$

where $g = GCD(a_1, \cdots, a_n, d)$. *Let* $\{p_1, \cdots, p_k\}$ *be a maximal set of nonassociated prime divisors of* $e = c/d$ *such that* $GCD(a'_1, \cdots, a'_n, p_i) = 1$ *for* $i = 1, \cdots, k$. *Then*

$$N_c = |c|^{n-1}|ge|\prod_{i=1}^{k}(1 - |p_i|^{-1}) . \tag{4.7}$$

Proof. Suppose that $c = de$ and $g = GCD(a_1, \cdots, a_n, b, c)$. Since R/cR is finite, d is atomic and R/pR is a finite field for any prime $p \mid d$. Hence as $g \mid b$, a standard argument shows that $a_1x_1 + \cdots + a_nx_n + b \equiv o(\mathrm{mod}\, d)$ is solvable and has $|g||d|^{n-1}$ distinct solutions $(\mathrm{mod}\, d)$. Thus $GCD(a_1x_1 + \cdots + a_nx_n + b, c) = d$ is solvable since e is atomic. Let $d' = d/g$ and $b' = b/g$. Since $GCD(a'_1, \cdots, a'_n, d'p_i) = 1$ and $R/d'p_iR$ is finite, $a'_1x_1 + \cdots + a'_nx_n + b' \equiv 0(\mathrm{mod}\, d'p_i)$ is solvable for $i = 1, \cdots, k$. It follows that $t_0 = d'\prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(a_1x_1 + \cdots + a_nx_n + b, c) = d$. Let A' denote the $1 \times n$ matrix $(a'_1, \cdots, a'_n)$. Then rank $A'(\mathrm{mod}\, p_i) = 1$ for $i = 1, \cdots, k$. Also $a'_1x_1 + \cdots + a'_nx_n \equiv o(\mathrm{mod}\, d')$ has $|d'|^{n-1}$ distinct solutions $(\mathrm{mod}\, d')$. Thus by (iii) of Corollary 1,

$$N_{t_0} = |d'|^{n-1}\prod_{i=1}^{k}(|p_i|^n - |p_i|^{n-1}) ,$$

which yields (4.7).

COROLLARY 4. *Suppose that* R/cR *is a finite ring where* $c = de$. *Let* $g = GCD(a_1, \cdots, a_m, d)$ *and* $a'_i = a_i/g$ *for* $i = 1, \cdots, m$. *Then* $GCD(a_1x + b_1, \cdots, a_mx + b_m, c) = d$ *is solvable if and only if*

(1) $GCD(a_i, d) \mid b_i$ *for* $i = 1, \cdots, m$,

(2) $a'_ib_j \equiv a'_jb_i(\mathrm{mod}\, d)$ *for* $1 \leqq i < j \leqq m$,

(3) $g = GCD(a_1, \cdots, a_m, b_1, \cdots, b_m, c)$.

Let $\{p_1, \cdots, p_k\}$ *be a maximal set of nonassociated prime divisors of* e *such that for each* p_h, $GCD(a_i, dp_h) \mid b_i$ *for* $i = 1, \cdots, m$ *and* $a'_i \equiv a'_jb_i(\mathrm{mod}\, dp_h)$ *for* $1 \leqq i < j \leqq m$. *Then*

$$N_c = |ge|\prod_{h=1}^{k}(1 - |p_h|^{-1}) .$$

Proof. Let A and B denote the $m \times 1$ matrices with entries $a_1, \cdots, a_m$ and $b_1, \cdots, b_m$ respectively. Since R/dR is finite, the reader may easily verify that the system $Ax + B \equiv 0(\mathrm{mod}\, d)$ is solvable if and only if (1) and (2) hold. Thus as e is atomic, $GCD(Ax + B, c) = d$ is solvable if and only if (1), (2), and (3) hold. Let $GCD(Ax + B, c) = d$ be solvable and let $d' = d/g$. Then it follows that $t_0 = d'\prod_{h=1}^{k} p_h$ is the minimum modulus of $GCD(Ax + B, c) = d$. Let A' denote the $m \times 1$ matrix with entries $a'_1, \cdots, a'_m$. Then rank $A'(\mathrm{mod}\, p_i) = 1$ for

$i = 1, \cdots, k$. Also the system $A'x \equiv 0 (\mathrm{mod}\, d')$ has only the solution $x \equiv \mathrm{o}(\mathrm{mod}\, d')$. Thus by (iii) of Corollary 1, $N_{t_0} = \prod_{h=1}^{k} (|p_h| - 1)$. Hence $N_c = |ge| \prod_{h=1}^{k} (1 - |p_h|^{-1})$.

COROLLARY 5. *Let $c = de$ where e is atomic. Let $g = GCD(a_1, \cdots, a_n, d)$ and $d' = d/g$. Suppose that $R/d'R$ is a finite ring. Then $GCD(a_1x_1 + b_1, \cdots, a_nx_n + b_n, c) = d$ is solvable if and only if $GCD(a_j, d) \mid b_j$ for $j = 1, \cdots, n$ and $g = GCD(a_1, \cdots, a_n, b_1, \cdots, b_n, c)$. Suppose that $R/(\prod_{i=1}^{k} p_i)R$ is finite where $\{p_1, \cdots, p_k\}$ is a maximal set of nonassociated prime divisors of e such that for each p_i, $GCD(a_j, dp_i) \mid b_j$ for $j = 1, \cdots, n$. Then $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(a_1x_1 + b_1, \cdots, a_nx_n + b_n, c) = d$. Let $d_j = GCD(a_j, d)$ and $d'_j = d_j/g$ for $j = 1, \cdots, n$. Then*

$$N_{t_0} = \prod_{j=1}^{n} |d'_j| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-t_i}) \tag{4.8}$$

where t_i denotes the number of j in $\{1, \cdots, n\}$ for which

$$GCD\left(\frac{a_j}{d_j}, p_i\right) = 1 .$$

Proof. Suppose that $d_j \mid b_j$ for $j = 1, \cdots, n$. Let $a'_j = a_j/g$ and $b'_j = b_j/g$ for $j = 1, \cdots, n$. Let A and A' denote the $n \times n$ diagonal matrices with entries $a_1, \cdots, a_n$ and $a'_1, \cdots, a'_n$ respectively. Let B and B' denote the $n \times 1$ matrices with entries $b_1, \cdots, b_n$ and $b'_1, \cdots, b'_n$ respectively. Then the system $A'X + B' \equiv 0(\mathrm{mod}\, d')$ is solvable since $GCD(a'_j, d') \mid b'_j$ for $j = 1, \cdots, n$ and $R/d'R$ is finite. Thus the system $AX + B \equiv 0(\mathrm{mod}\, d)$ is solvable. Hence if $g = GCD(a_1, \cdots, a_n, b_1, \cdots, b_n, c)$, then $GCD(AX + B, c) = d$ is solvable.

Assume that $GCD(AX + B, c) = d$ is solvable. It follows that $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(AX + B, c) = d$. Let $L = \{X + d'R^n \mid A'X \equiv 0(\mathrm{mod}\, d')\}$. Let

$$\mathscr{L}_i = \{X \in R^n \mid A'X \equiv 0(\mathrm{mod}\, d'p_i)\}$$

and $L_i = \{X + d'R^n \mid X \in \mathscr{L}_i\}$ for $i = 1, \cdots, k$. Then by (4.1),

$$N_{t_0} = |L| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-(r_i+s_i)})$$

where $r_i = \mathrm{rank}\, A'(\mathrm{mod}\, p_i)$ and s_i is the dimension of the R/p_iR vector space L/L_i. Clearly $|L| = \prod_{j=1}^{n} |d'_j|$ since $d'_j = GCD(a'_j, d')$ for $j = 1, \cdots, n$. Let $L'_i = \{X + d'p_iR^n \mid X \in \mathscr{L}_i\}$ and $H_i = \{X + p_iR^n \mid X \in \mathscr{L}_i\}$ for $i = 1, \cdots, k$. Then (1) and (2) of Lemma 3 show that $|L'_i| = |L| |H_i|$ where $|H_i| = |p_i|^{n-(r_i+s_i)}$ for $i = 1, \cdots, k$. However, $GCD(a'_j, d'p_i) = d'_j\, GCD(a_j/d_j, p_i)$ and thus

$$|L_i'| = |L| \prod_{j=1}^{n} \left| GCD\left(\frac{a_j}{d_j}, p_i\right)\right|$$

for $i = 1, \cdots, k$. Hence $|p_i|^{n-(r_i+s_i)} = \prod_{j=1}^{n} |GCD(a_j/d_j, p_i)|$ and thus $|p_i|^{n-(r_i+s_i)} = |p_i|^{n-t_i}$, since t_i is the number of j in $\{1, \cdots, n\}$ for which $GCD(a_j/d_j, p_i) = 1$. So $t_i = r_i + s_i$ for $i = 1, \cdots, k$, which yields (4.8).

Note that if R/cR is finite, then

$$N_c = \prod_{j=1}^{n} |d_j e| \prod_{i=1}^{k} (1 - |p_i|^{-t_i}) .$$

COROLLARY 6. *Let R be a principal ideal domain. Let A be an $m \times n$ matrix of* rank *r and let $\alpha_1, \cdots, \alpha_r$ be the invariant factors of A. Let B be an $m \times 1$ matrix and let $(A\colon B)$ have* rank *r' and invariant factors $\beta_1, \cdots, \beta_{r'}$. Then $GCD(AX + B, c) = d$ is solvable if and only if* (1) $d \mid c$, (2) $GCD(\alpha_1, d) = GCD(\beta_1, c)$, (3) $GCD(\alpha_j, d) = GCD(\beta_j, d)$ *for* $j = 1, \cdots, r$ *and* $\beta_{r'} \equiv 0 (\bmod\, d)$ *if* $r' = r + 1$.

Let $\{p_1, \cdots, p_k\}$ be a maximal set of nonassociated prime divisors of $e = c/d$ such that each p_i satisfies (3′) $GCD(\alpha_j, dp_i) = GCD(\beta_j, dp_i)$ *for $j = 1, \cdots, r$ and $\beta_{r'} \equiv 0 (\bmod\, dp_i)$ if $r' = r + 1$. Let $d_j = GCD(\alpha_j, d)$ for $j = 1, \cdots, r$ and $d' = d/d_1$. Then $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(AX + B, c) = d$. Suppose that R/t_0R is finite. Then*

$$N_{t_0} = |d'|^{n-r} \prod_{j=1}^{r} |d_j'| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-t_i}) \tag{4.9}$$

where $d_j' = d_j/d_1$ and t_i denotes the largest j in $\{1, \cdots, r\}$ for which $GCD(\alpha_j/d_j, p_i) = 1$.

Proof. Since R is a principal ideal domain, it is well-known that there exist invertible matrices P and Q such that $PAQ = A_0$ where A_0 is an $m \times n$ matrix in "diagonal form", with nonzero entries $\alpha_1, \cdots, \alpha_r$ and $\alpha_j \mid \alpha_{j'}$ if $j < j'$. The elements $\alpha_1, \cdots, \alpha_r$ are called the invariant factors of A and $\alpha_j = D_j/D_{j-1}$ where D_j denotes the GCD of the determinants of all the $j \times j$ submatrices of A. Clearly $GCD(A, d) = GCD(\alpha_1, \cdots, \alpha_r, d)$, that is $GCD(A, d) = GCD(\alpha_1, d)$ since $\alpha_1 \mid \alpha_j$ for $j = 1, \cdots, r$. Similarly $GCD(A, B, c) = GCD(\beta_1, c)$. However, it is also well-known that the system $AX + B \equiv 0 (\bmod\, d)$ is solvable if and only if condition (3) holds (see [4]). Thus $GCD(AX + B, c) = d$ is solvable if and only if (1), (2), and (3) hold.

Let $GCD(AX + B, c) = d$ be solvable and let $c = de$. Then $t_0 = d' \prod_{i=1}^{k} p_i$ is the minimum modulus of $GCD(AX + B, c) = d$. Suppose that R/t_0R is finite. Let S denote the set of X in R^n such that $GCD(AX + B, c) = d$. Let $PB = B_0$ and let S' denote the set of Y in R^n such that $GCD(A_0Y + B_0, c) = d$. Then clearly $X \in S$ if and only if $Y = Q^{-1}X \in S'$. Thus $GCD(AX + B, c) = d$ and $GCD(A_0Y + B_0, c) = d$ have the same ideal of solution moduli. Let $T_0 = \{X + t_0R^n \mid X \in S\}$ and $T_0' = \{Y + t_0R^n \mid Y \in S'\}$. Then the mapping $f\colon T_0 \to T_0'$ is a bijection, where $f(X + t_0R^n) = Q^{-1}X + t_0R^n$ for all X in S. Hence $|T_0| = |T_0'|$, that is $N_{t_0} = |T_0'|$. Let B_0 have entries $b_1^0, \cdots, b_m^0$ and let $c_0 = GCD(b_{r+1}^0, \cdots, b_m^0, c)$. Then S' is the set of solutions of the linear GCD equation

$$\begin{aligned} GCD(\alpha_1 y_1 + b_1^0, \cdots, \alpha_r y_r + b_r^0, 0\cdot y_{r+1} + 0,& \\ \cdots, 0\cdot y_n + 0, c_0) = d&\ . \end{aligned} \tag{4.10}$$

Thus $t_0 = d' \prod_{i=1}^k p_i$ is also the minimum modulus of (4.10) and hence by (4.8) of Corollary 5,

$$N_{t_0} = |d'|^{n-r} \prod_{j=1}^{r} |d_j'| \prod_{i=1}^{k} (|p_i|^n - |p_i|^{n-t_i})$$

where $d_j' = d_j/d_1$ and t_i is the largest j in $\{1, \cdots, r\}$ for which $GCD(\alpha_j/d_j, p_i) = 1$ since $\alpha_j/d_j \mid \alpha_{j'}/d_{j'}$ if $j < j'$.

If R/cR is finite, then

$$N_c = |c|^{n-r} \prod_{j=1}^{r} |d_j e| \prod_{i=1}^{k} (1 - |p_i|^{-t_i})\ .$$

Finally we remark that the formula for N_{t_0} in (4.1) applies to the class $\mathscr{D}$ of GCD domains R which contain at least one element p such that R/pR is a finite field. Some immediate examples are the integers $\boldsymbol{Z}$, the localizations $\boldsymbol{Z}_{(p)}$ at primes p in $\boldsymbol{Z}$ and $F[X]$ where F is a finite field.

However, an example of such a ring R in $\mathscr{D}$ which is not a PID is the subring R of $\boldsymbol{Q}[X]$ consisting of all polynomials whose constant term is in $\boldsymbol{Z}$. Indeed R is a Bezout domain which cannot be expressed as an ascending union of PID's [1]. Clearly if p is a prime in $\boldsymbol{Z}$, then R/pR is isomorphic to the finite field $\boldsymbol{Z}/p\boldsymbol{Z}$.

We are also indebted to Professor W. Heinzer for the following construction of a ring R in $\mathscr{D}$ which is a UFD but not a PID. Let F be a finite field. Let Y be an element of the formal power series ring $F[[X]]$ such that X and Y are algebraically independent over F. Let V denote the rank one discrete valuation ring $F[[X]] \cap F(X, Y)$ and let $R = F[X, Y][1/X] \cap V$. Then R/XR is isomorphic to F and R is a UFD.

References

1. P. M. Cohn, *Bezout rings and their subrings*, Proc. Camb. Phil. Soc., **64** (1968), 251–264.

2. D. Jacobson and K. S. Williams, *On the solution of linear GCD equations*, Pacific J. Math., **39** (1971), 187-206.
3. I. Kaplansky, *Elementary divisors and modules*, Trans. Amer. Math. Soc., **66** (1949), 464-491.
4. H. J. S. Smith, *On systems of linear indeterminate equations and congruences*, Phil. Trans. London, **151** (1861), 293-326. (Collected Mathematical Papers Vol. **1**, Chelsea, N.Y.), (1965), 367-409.
5. R. Spira, *Elementary problem no. E*1730, Amer. Math. Monthly, **72** (1965), 907.

Received February 22, 1974 and in revised form September 3, 1974. This research was supported in part by NRC Grant-A8749.

UNIVERSITY OF NEW BRUNSWICK

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

A NOTE ON COMPACT SEMIRINGS WHICH ARE MULTIPLICATIVE SEMILATTICES

P. H. KARVELLAS

The topic of this note is the structure of a topological semiring in which a semilattice (commutative, idempotent and associative) multiplication, with identity and connected upper sets, has been postulated. Assuming the topology to be compact, additions compatible with the multiplication can be characterized for certain canonical subsets of the semiring. In particular instances the characterization of addition can be extended to the entire semiring itself.

Certain subintervals, arising naturally from the analysis when the underlying space is the interval [0, 1], are generalized to continuum subsemirings of an arbitrary semiring possessing a semilattice multiplication with identity. The addition in the minimal additive ideal can be specified precisely and each additive subgroup is a single element. If the minimal additive ideal and the set of additive idempotents coincide, a complete description of the semiring addition is possible in terms of homomorphisms of the multiplicative semigroup. The same procedure can be employed when the space is an interval on the real line.

A *topological semiring* $(S, +, \cdot)$ is a Hausdorff space S on which are defined topological semigroups $(S, +)$ and $(S, \cdot)$, for addition and multiplication, such that $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for all x, y, and z in S. This structure will be investigated under the restrictions that $(S, \cdot)$ is a topological semilattice, with identity 1 and multiplicative zero element 0, the set S is compact and upper sets $M(x) = \{y : xy = x\}$ are connected for each x in S. Such a semiring will be called a *semilattice semiring* or *SL-semiring*. Multiplication is therefore commutative and idempotent in a semilattice semiring and an induced partial order, with closed graph, results from defining $x \leqq y$ if $x = xy$.

Unless specifically altered, both $(S, +, \cdot)$ and S shall refer to semilattice semirings in the analysis which follows.

Particular examples of SL-semirings appear in [5], where S is the real number interval [0, 1]. The characterization of such interval SL-semirings is given in Example 1 and employs two continuous functions satisfying certain required conditions on subsets of [0, 1]. A more general space and analysis will, of course, be subject to rather more exaggerated ambiguities.

Ideals will be semigroup ideals in the sense of [1] and kernels (minimal ideals) with be written as $K[+]$ and $K[\cdot]$. In the compact case kernels are nonvoid and closed [7], as are the *idempotent sets* $E[+] = \{x: x = x + x\}$ and $E[\cdot] = \{x: x = x^2\}$. The *union of all additive subgroups* will be written as $H[+]$ and for t in $E[+]$ the *maximal additive subgroup* with identity element t is $H[+](t)$. For a positive integer n and element x, nx denotes the n-fold sum of x. Equivalently nx is the product of two elements of the semiring. The element $(1 + 1)$ will be written as p.

For an element x let $L(x) = \{y: xy = y\}$ and $M(x) = \{y: xy = x\}$. If $x \leqq y$, that is if $x = xy$, then define $C(x, y) = \{z: x \leqq z \leqq y\} = M(x) \cap L(y) = y \cdot M(x)$. In any SL-semiring, $M(x)$ is connected, implying the connectivity of $C(x, y)$ for $x \leqq y$. It is trivial to verify that $C(x, y)$ is a subsemiring if and only if $x \in E[+]$. Lastly, from $S = E[\cdot]$, $x + y = (x + y)^2 = x + p(xy) + y$ for all $x, y \in S$.

2. Connected subsemirings of a semilattice semiring.

In Example 1 is given the characterization, obtained in [5], of all SL-semirings on the interval $[0, 1]$. The resulting subintervals $[0, e]$, $[e, f]$, $[f, p]$, and $[p, 1]$ have obvious generalizations to an arbitrary SL-semiring defined on a general topological space.

EXAMPLE 1. Let $S = [0, 1]$ with multiplication $xy = \min(x, y)$. Any compatible semiring addition, with $x + y = y$ in $K[+]$, can be characterized as follows. Pick arbitrary elements e, f, and p in $[0, 1]$, where $0 \leqq e \leqq f \leqq p \leqq 1$. Let $F: [0, p] \to [e, 1]$ and $G: [0, p] \to [f, 1]$ be continuous functions such that

(1) F is the identity on $[e, p]$;

(2) F decreases on $[0, e]$ and G decreases on $[0, f]$;

(3) for $x \in [0, p]$, $pG(x) = \max(f, pF(x))$.

The addition on S is defined by

$$\begin{aligned} x + y &= p && x, y \geqq p \\ &= xF(y) && y \leqq x,\ y < p \\ &= yG(x) && x < y,\ x < p\,. \end{aligned}$$

The subintervals $[0, e]$, $[e, f]$, $[f, p]$, and $[p, 1]$ are connected subsemirings with the additions below.

$$\begin{array}{llll} x + y = \max(x, y) & x, y \in [0, e] & s + k = k & k \in [e, f],\ s \in S \\ x + y = xy & x, y \in [f, p] & x + y = p & x, y \in [p, 1]\,. \end{array}$$

The additive kernel $K[+]$ is the subinterval $[e, f]$, while $E[+] = [0, p]$.

In any SL-semiring $(S, \cdot)$ is commutative and the kernel $K[\cdot]$ must reduce to a singleton, denoted hereafter by 0 [4]. It is easy to verify that $2x = 4x$ for each x in S and from [3] both $E[+]$ and

$H[+]$ are multiplicative ideals, requiring $0 \in E[+]$. Because $(S, \cdot)$ has an identity, $E[+]$ is closed under addition [3], and both $E[+]$ and $H[+]$ are connected [8]. Alternatively $p = 1 + 1 = p^2 = p + p$ and $px = x + x$ for each x in S. The map $x \to px$ is continuous and $M(0) = S$ is connected. Hence $E[+] = pS$ is connected. As will be proven subsequently, $E[+] = H[+]$. Noting that $S = M(0)$ and is connected, we have the result below.

THEOREM 1. *Let $(S, +, \cdot)$ be a semilattice semiring.*

(1) *$K[\cdot] = \{0\} \subseteqq E[+]$ and S is a connected set.*

(2) *$E[+] = \{x + x : x \in S\}$ and is an additive subsemigroup.*

(3) *$E[+]$ and $H[+]$ are connected multiplicative ideals.*

The next result characterizes the operations in the minimal additive ideal $K[+]$.

THEOREM 2. *Let $(S, +, \cdot)$ be a semilattice semiring. Then:*

(1) *$K[+]$ is a subsemiring of S contained in $E[+]$.*

(2) *There exist elements e and f in S such that $K[+] = C(e, f)$ and $f = 1 + k + 1$ for each element $k \in K[+]$.*

(3) *$K[+] = (S + e) + (e + S)$, with each element z in $K[+]$ uniquely of the form $z_1 + z_2$, where $z_1 \in S + e$ and $z_2 \in e + S$. Moreover, for elements x_1, x_2 in $S + e$ and y_1, y_2 in $e + S$, the kernel operations are given by*

$$(x_1 + y_1) + (x_2 + y_2) = x_1 + y_2$$
$$(x_1 + y_1) \cdot (x_2 + y_2) = x_1x_2 + y_1y_2$$
$$(e + S) \cap (S + e) = \{e\} .$$

Proof. Because $S^2 \cap K[+]$ is nonvoid, the additive kernel is a subsemiring using a result from [6]. From $S = E[\cdot]$ and Theorem 1 of [7] each additive subgroup is totally disconnected. However, $K[+]$ is the union of the connected maximal subgroups $H[+](t) = t + S + t$ for t in $K[+] \cap E[+]$ [8]: hence $H[+](t) = \{t\}$ for each $t \in K[+] \cap E[+]$ and thus $K[+] \subseteqq E[+]$. The compact, commutative subsemigroup $(K[+], \cdot)$ has a multiplicative kernel which is a single point. Let $\{e\}$ denote this kernel. Then $f = 1 + e + 1$ is in $K[+]$ and, for each element k in $K[+]$, $e \leqq k$ while

$$fk = k + e + k \in k + S + k = H[+](k) = \{k\}$$

$$1 + k + 1 = f(1 + k + 1) = f + k + f \in H[+](f) = \{f\}$$

proving that $1 + K[+] + 1 = \{f\}$ and $K[+] \subseteqq C(e, f)$. For any element $x \in C(e, f)$, $x = xf = x(1 + e + 1) = x + e + x \in C(e, f) \cap K[+]$

and hence $K[+] = C(e, f)$. The characterization of addition in $K[+]$ follows directly from Theorem 1.3.10 of [4] and the triviality of maximal additive subgroups in $K[+]$. For x_1, x_2 in $S + e$ and y_1, y_2 in $e + S$ we have that $x_1 = x_1 + e$, $y_1 = e + y_1$ and $H[+](e) = e + S + e = \{e\}$, implying therefore that $(e + S) \cap (S + e) \subseteqq H[+](e)$ and that

$$\begin{aligned}(x_1 + y_1) \cdot (x_2 + y_2) &= x_1x_2 + y_1x_2 + x_1y_2 + y_1y_2 \\ &= x_1x_2 + (e + y_1)x_2 + (x_1 + e)y_2 + y_1y_2 \\ &= x_1x_2 + e + y_1x_2 + x_1y_2 + e + y_1y_2 \\ &= x_1x_2 + e + y_1y_2 \\ &= x_1x_2 + y_1y_2 \,.\end{aligned}$$

The subsets of interest are the following: $E[+] = pS$, $K[+] = C(e, f)$, $M(p) = \{x: px = p\}$, $L(e) = eS$ and $1 + S + 1$. Both $E[+]$ and $K[+]$ have been shown to be connected subsemirings from the preceding arguments. As proven in Theorem 4, the requirement that $M(x)$ be connected for each x in S results in $p = p + 1$ and implies triviality of addition in $M(p)$. If the restriction on upper sets is removed, partial results can still be obtained.

THEOREM 3. *Let $(T, +, \cdot)$ be a compact semiring, with $E[+] = \{q\}$, such that $(T, \cdot)$ is a semilattice with identity* 1. *Then*:

(1) $1 + x = x + 1$ *and* $q = 1 + 1 = x + q + x$ *for all* x *in* T.

(2) $(T, +)$ *is commutative.*

(3) $T + T$ *is the additive kernel.*

Proof. Since $T = E[\cdot]$, $1 + 1 = (1 + 1)^2 = (1 + 1) + (1 + 1) \in E[+]$ and thus $q = 1 + 1$. Moreover, $K[\cdot] \subseteqq E[+] = \{q\}$. Hence $q = qx$ for each x in T. It is easily shown that $1 + x = (1 + x)^2 = 1 + 3x = q + 1 + x$ for each element x of T. Analogously $x + 1 = x + 1 + q$. As a result one obtains the equations

$$\begin{aligned}(x + 1) \cdot (1 + x) &= x(1 + x) + (1 + x) = q + 1 + x = 1 + x \\ &= (x + 1) + (x + 1)x = x + 1 + q = x + 1 \,.\end{aligned}$$

Moreover, $x + q + x = x + qx + x = x(2q) = q$. In a similar manner it can be proven that $x + y = (x + y) \cdot (y + x) = y + x$ for all x and y in T. Addition in T is therefore commutative.

Lastly, because $(T, +)$ is a compact semigroup with a single idempotent element, $K[+] = H[+](q) = T + q + T$ [8]. Thus, for x and y in T, $x + y = (x + y)^2 = x + q(xy) + y = x + q + y \in K[+]$. Therefore $T + T \subseteqq T + q + T = K[+]$, implying that $K[+] = T + T$.

THEOREM 4. *Let $(S, +, \cdot)$ be a semilattice semiring, $p = 1 + 1$. Then*:

(1) *$(M(p), +, \cdot)$ is a subsemiring with trivial addition.*
(2) *$L(e) = eS$ is a distributive topological lattice.*

Proof. From $M(p) = \{x: px = p\}$ it is clear that $M(p)$ is a continuum subsemiring with a single additive idempotent. Theorem 3 applies and it is now only necessary to note that the additive kernel of the subsemiring $M(p)$ is the connected additive group $M(p) + p + M(p)$. However, $M(p) \subseteqq E[\cdot]$ and from [7] this group must also be totally disconnected. Consequently $M(p) + M(p) = \{p\}$ and $p = 1 + 1 = p + 1 = 1 + p$.

Recall that $K[+] = C(e, f)$ where $\{e\}$ is the multiplicative kernel of the subsemiring $K[+]$. The subcontinuum $eS = L(e)$ is a subsemiring with identity e and $e = e + x = x + e$ for each $x = ex$ in eS. Thus for elements x and y of eS we obtain

$$(x + y)x = xe + xy = x(e + y) = xe = x$$
$$(x + y)y = xy + ey = (x + e)y = ey = y\ .$$

Therefore, $x + y \in M(x) \cap M(y)$ and for any $t \in M(x) \cap M(y)$ it follows that $t(x + y) = tx + ty = x + y$. That is, $x + y$ is the least upper bound of x and y in the partial order defined by the semilattice multiplication and consequently $(eS, +, \cdot)$ is a lattice. Since multiplication distributes over addition, both lattice distributive laws hold.

COROLLARY 5. *Let $(S, +, \cdot)$ be a semilattice semiring. If $E[+] = \{0\}$ then $S + S = \{0\}$.*

THEOREM 6. *Let $(S, +, \cdot)$ be a semilattice semiring and let f denote the maximal element of the additive kernel, while $p = 1 + 1$. Then:*

(1) *These are equivalent statements.*
(a) *$(E[+], +)$ is commutative.*
(b) *$x + p = p + x$ for all x in $E[+]$.*
(c) *$x + p = p + x$ for all x in S.*

(2) *If $(E[+], +)$ is commutative, then $(E[+], +, \cdot)$ is a topological lattice if and only if $f = p$.*

Proof. Recall that $E[+]$ is a connected subsemiring. For any x in S we have that $x + p = (x + p)^2 = (p + 1)x + p$ and $(p + 1)x \in E[+]$. Thus if $x + p = p + x$ for x in $E[+]$, the same result holds in S, and vice versa.

Clearly, (a) $\rightarrow$ (b). Assume that elements of S commute with p under addition. For $x, y \in E[+]$, $x = x + x = px$, $xy = pxy$, $y = y + y = py$ and thus the equations below are obtained.

$$\begin{aligned} x + y = (x + y)^2 = x + xy + y = x(p + y) + y &= xy + (x + y) \\ &= x + (x + p)y = (x + y) + xy \end{aligned}$$

$$\begin{aligned} y + x = (y + x)^2 = y + xy + x &= xy + (y + x) \\ &= (y + x) + xy \,. \end{aligned}$$

It follows that

$$\begin{aligned} (x + y) \cdot (y + x) = x(y + x) + y(y + x) &= xy + (x + y) + xy \\ = (x + y)y + (x + y)x &= xy + (y + x) + xy \end{aligned}$$

which implies that $(E[+], +)$ is commutative.

Assume now that addition in $E[+]$ is commutative. Because distinct idempotents in $K[+]$ do not commute in the compact case [4], we obtain $K[+] = \{f\}$. If $f = p$ then, from Theorem 4, $E[+]$ is a distributive topological lattice. Conversely, if $E[+]$ is a lattice then, since one distributive law holds, $E[+]$ is a distributive lattice. Therefore, because $a = a(a + b) = a + (ab)$ for all a and b in the lattice $E[+]$, we obtain

$$p = p + pf = p^2 + pf = p(p + f) = pf = f \,.$$

The following example illustrates the general idempotent semilattice semiring with commutative addition which can be constructed on an interval.

EXAMPLE 2. Let $S = [z, p]$ be an interval of real numbers with min multiplication. Fix an element f in S and denote the subintervals $[z, f]$ by A and $[f, p]$ by B respectively. If $\{f\}$ is the additive kernel of an idempotent and commutative addition semiring on $[z, p]$, then $B = p + B$ and $x + y = \min(x, y)$ in B, while $x + y = \max(x, y)$ in A. The map $f\colon S \to B$ defined by $f(x) = 1 + x$ is continuous and is the identity on B. Moreover, f reverses order on A ($xy = x$ in A implies $f(x) \cdot f(y) = f(y)$ in B). Any such addition on S is therefore given by the characterization

$$\begin{aligned} x + y &= xF(y) \qquad y \leqq x \\ &= yF(x) \qquad x < y \end{aligned}$$

where $F\colon S \to B$ is continuous, the identity on B and order-reversing on A.

The existence of the three elements $p(= 1 + 1)$, e and f, where $K[+] = C(e, f)$, has allowed the characterization of addition in $M(p)$, $K[+]$ and $L(e)$. The next result completes the description of connected subsemirings which are analogues of the subintervals appearing in Example 1.

THEOREM 7. *Let* $(S, +, \cdot)$ *be a semilattice semiring,* $p = 1 + 1$ *and* $K[+] = C(e, f)$ *for elements* $e \leqq f$ *in* $E[+]$. *Then:*

(1) $H[+] = E[+]$ *and each additive subgroup is a single point.*

(2) $1 + S + 1 = 1 + E[+] + 1 \subseteqq M(f) \cap E[+]$ *with addition given by* $x + y = xy = y + x$.

(3) *For* $x \in 1 + S + 1$, $y \in M(p)$, $x + y = x = y + x$.

(4) $M(f) + K[+] + M(f) = \{f\}$.

(5) $e + 1 \geqq e + s$ *and* $1 + e \geqq s + e$ *for all* s *in* S.

(6) $S + p + S \subseteqq E[+]$.

(7) *The boundary* B *of* $E[+]$ *is connected.*

Proof. For $t \in E[+]$ the maximal additive subgroup $H[+](t)$ is a subsemiring since $t = t^2$ [2]. Moreover, $H[+](t) \subseteqq M(t)$ since for each $x \in H[+](t)$, $tx \in E[+] \cap H[+](t) = \{t\}$. Consequently $x + x = px = t$ and therefore $x = x + t = (1 + p)x = t$ for each x in $H[+](t)$. Hence $H[+] \subseteqq E[+]$ and each additive subgroup is a single element.

Clearly $1 + E[+] \subseteqq 1 + S$ and, because $1 + x = (1 + x)^2 = 1 + px$ for each element x, the reverse inclusion also holds. Similarly $S + 1 = E[+] + 1$ and for each element x of S we have that

$$\begin{aligned} 1 + x + 1 = (1 + x + 1)^2 &= (1 + x + 1) + 3x + (1 + x + 1) \\ &= p(1 + x) + p(x + 1) \\ &= p(1 + x + 1) \in E[+] \,. \end{aligned}$$

In addition it follows that $f = f + x + f = f(1 + x + 1)$, implying that $1 + S + 1 \subseteqq M(f) \cap E[+]$. For any two elements x and y of $1 + S + 1$, $px = px + 1$ and $py = 1 + py$ and hence

$$\begin{aligned} x + y = (x + y)^2 = x + p(xy) + y &= x(1 + py) + y \\ &= p(xy) + y = p(xy) = xy \end{aligned}$$

and in a similar manner $y + x = xy$. Moreover, for $x \in 1 + S + 1$ and $y \in M(p)$ we obtain

$$x + y = xy + y = (x + 1)y = xy = x \,.$$

For elements $k \in K[+]$, and $m, n \in M(f)$, we have that

$$k + n = f(k + n) = k + fn = k + f$$

$$m + k + n = (f + k) + n = f + k + f = f \,.$$

Consequently $M(f) + K[+] + M(f) = \{f\}$.

For any element $s \in S$ it follows that $(e + s) \leqq (e + 1)$ since

$$(e + 1)(e + s) = e + es + e + s = e + s$$

and similarly $(s + e) \leqq (1 + e)$. In addition, for elements x and y of

S, $px + 1 = x + 1$, $1 + y = 1 + py$ and therefore $x + p + y = p(x + 1 + y) \in E[+]$, implying $S + p + S \subseteqq E[+]$.

Lastly, consider the set $T = S \backslash E[+]$, which is connected since for each t in T the interval $C(t, 1) \subseteqq T$. Consequently pT is also connected and $pT \subseteqq E[+]$. For x in T let $R(x) = \{y: px = py\}$. Then $R(x) \cap E[+] = \{px\}$, $x \in R(x)$ and it is easily verified that $R(x)$ is a compact subsemiring of S. Moreover, $C(px, y) \subseteqq R(x)$ for each y in $R(x)$, implying that $R(x)$ is connected. Suppose now that px is contained in the interior of $E[+]$. There then exists an open set U, containing px, and contained in $E[+]$. However, $U \cap R(x) = \{px\}$ is an open and closed subset of the connected set $R(x)$. Consequently pT is contained in the boundary B of $E[+]$. It is now only necessary to note that if $r \in B$, then for any open set W containing r there exists an open set V, containing r, such that $pV \subseteqq W$. Thus, since $V \cap T$ is nonempty, r is a limit point of the connected set pT and B is connected.

Identification of the various connected subsemirings of a general semilattice semiring with the subintervals obtained in Example 1 yields the correspondences: $L(e)$ with $[0, e]$; $M(p)$ with $[p, 1]$; and, $1 + S + 1$ with $[f, p]$. The addition in the additive kernel $K[+]$ of a general SL-semiring is that of a rectangular band [1], while the existence of a cutpoint in the Example 1 case produces either a left- or right-trivial addition [4].

The construction of "characterizing functions", as given in Example 1, is apparently futile for a general semilattice semiring. However, as demonstrated below, the situation $K[+] = E[+]$ is amenable to this approach.

3. Semilattice semirings with $K[+] = E[+]$.

In the case of SL-semiring with $K[+] = E[+]$ it is possible to obtain a complete characterization of the addition in terms of semilattice homomorphisms on the multiplicative semigroup. The following lemma establishes some preliminary results.

LEMMA 8. *Let S be a semilattice semiring with $K[+] = E[+]$. Then*:

(1) $S + S \subseteqq E[+]$.

(2) *For* $x, y \in S$, $\{x + y\} = x + S + y$, $0 + x \leqq 0 + 1 \leqq x + 1$ *and* $x + 0 \leqq 1 + 0 \leqq 1 + x$.

(3) *For* $k \in K[+]$, $k + M(f) = \{k + 1\}$, $\{f\} = M(f) + k + M(f)$.

(4) *The maps* $x \xrightarrow{F} (1 + x)$ *and* $x \xrightarrow{G} (x + 1)$ *are semiring homomorphisms with* $F(x + y) = F(y)$ *and* $G(x + y) = G(x)$. *Addition in S is given by*

$$x + y = G(x) \cdot F(y) .$$

(5) *For* $x, y \in S$, $M(x + 0) \cap M(0 + x) = M(px)$ *and* $M(f) = M(1 + x) \cap M(y + 1) = M(x + 1) \cap M(1 + y)$.

Proof. Noting that $p = f$ and that $E[+](= K[+])$ is both an additive and multiplicative ideal, we have the result

$$x + y = (x + y)^2 = x + p(xy) + y \in K[+]$$

for each x and y in S. Recall that $H[+](px) = x + S + x = \{px\}$ and therefore, using both distributive laws, we obtain

$$\begin{aligned}(x + 1) \cdot (x + 0 + 1) &= (x + x) + 0 + (x + 1) = x + 1 \\ &= px + (x + 0 + 1) = x + 0 + 1 .\end{aligned}$$

Analogously $1 + x = 1 + 0 + x$. Using $\{p(xy)\} = xy + S + xy$ the following equations hold.

$$\begin{aligned}x + f + y &= p(x + f + y) = p(x + 1 + y) = x + 1 + y \\ &= (x + fx + xy) + (xf + f + yf) + (xy + fy + y) \\ &= fx + f(xy) + fy \\ &= x + p(xy) + y = x + y .\end{aligned}$$

Therefore, for any x and y in S, it follows that

$$\begin{aligned}x + S + y = f(x + S + y) &= (x + S) + (S + y) \\ &= (x + 1 + S) + (S + 1 + y) \\ &= x + f + y = x + y .\end{aligned}$$

For each x in S, $0 + x = x(0 + 1) \leqq 0 + 1$. Similarly we have that $(x + 1) \cdot (0 + 1) = 0 + x + 0 + 1 = 0 + 1 \leqq x + 1$. For k in $K[+]$ and m in $M(f)(= M(p))$, $k + m = p(k + m) = k + 1$. Analogously $M(f) + k = \{1 + k\}$, thereby establishing (3) as a special case of Theorem 7 (4).

Consider the maps $F, G: S \rightarrow K[+]$ defined by $F(x) = 1 + x$, $G(x) = x + 1$. Both are semiring homomorphisms and addition in S is given by

$$x + y = x + 1 + xy + y = (x + 1) \cdot (1 + y) = G(x) \cdot F(y) .$$

Lastly, $x + 0$, $0 + x \leqq px$. And, if $t \in M(x + 0) \cap M(0 + x)$, then $tx + 0 = x + 0$, $0 + x = 0 + tx$, implying the result

$$t(px) = tx + 0 + tx = x + 0 + x \in x + S + x = \{px\} .$$

Similarly, $M(1 + x) \cap M(y + 1) = M(x + 1) \cap M(1 + y) = M(f)$.

The next example describes a general semilattice semiring under the restriction that the additive kernel $K[+]$ is the set $E[+]$ of additive idempotents.

EXAMPLE 3. Let $(S, \cdot)$ be a compact topological semilattice, with identity element 1 and connected upper sets. Let p be any fixed element of S. If F and G are continuous semilattice homomorphisms from S into pS such that

(a) $(F \circ F)(x) = F(x)$, $(G \circ G)(x) = G(x)$ for all x in S;

(b) $F(x)G(x) = px$ for all x in S;

(c) $(F \circ G)(x) = (G \circ F)(x) = p$ for all x in S:

(where "$\circ$" denotes composition) and an addition is defined on S by

$$x + y = G(x)F(y)$$

for all x and y in S, then $(S, +, \cdot)$ is a semilattice semiring with additive kernel $K[+] = E[+] = pS$.

THEOREM 9. *Let $(S, \cdot)$ be a compact topological semilattice, with identity element* 1 *and connected upper sets.*

(a) *For any fixed element p of S, and homomorphisms F and G into pS defining an addition $(+)$ as in Example* 3, *$(S, +, \cdot)$ is a semilattice semiring with $K[+] = E[+] = pS$.*

(b) *Conversely, if $(+)$ is the addition of a semilattice semiring on the set S, with $K[+] = E[+]$ and addition compatible with the given semilattice multiplication, then the maps $F, G: S \rightarrow E[+]$ defined by $F(x) = 1 + x$, $G(x) = x + 1$ satisfy the properties of Example* 3 *when p is taken to be the element* $(1 + 1)$ *of S.*

Proof. The verification of part (a) is trivial, albeit tedious. If, on the other hand, $(S, +, \cdot)$ is a semilattice semiring with $E[+] = K[+]$, and the maps F and G are as defined, then both are continuous multiplicative homomorphisms, as proven in Lemma 8. Clearly $F(F(x)) = 1 + F(x) = p + x = 1 + x = F(x)$ and $G(G(x)) = G(x)$ for all x in S. Analogously $F(x) \cdot G(x) = (1 + x) \cdot (x + 1) = x + 1 + x = px$. Moreover, $(F \circ G)(x) = 1 + G(x) = 1 + x + 1 = p$. Lastly, as shown in Lemma 8, addition satisfies the definition given in Example 3.

The final two results, presented without proof, describe a SL-semiring in which $E[+] = K[+]$ and $S \backslash E[+] \subseteqq M(1 + 0) \cup M(0 + 1)$. Note that the latter condition is not sufficient to describe the characterization on the interval given in Example 1.

LEMMA 10. *Let S be a semilattice semiring with $E[+] = K[+]$. Then these are equivalent statements for an element x of S.*

(1) $1 + x = f$ $[x + 1 = f]$:

(2) $x \in M(0 + 1)$ $[x \in M(1 + 0)]$:

(3) $px = x + 1$ $[px = 1 + x]$.

THEOREM 11. *Let* $(S, +, \cdot)$ *be a semilattice semiring, with* $E[+] = K[+]$, *in which* $S \backslash E[+] \subsetneqq M(1+0) \cup M(0+1)$. *Then addition in* S *is given by*:

$$\begin{aligned} x + y &= py && x, y \in M(1+0) \\ &= px && x, y \in M(0+1) \\ &= f && x \in M(1+0),\ y \in M(0+1) \\ &= p(xy) && x \in M(0+1),\ y \in M(1+0) \\ &= G(x) \cdot y && x \in E[+],\ y \in M(1+0) \\ &= F(y) && x \in M(1+0),\ y \in E[+] \\ &= G(x) && x \in E[+],\ y \in M(0+1) \\ &= x \cdot F(y) && x \in M(0+1),\ y \in E[+] \end{aligned}$$

where $F, G: S \to E[+]$ *are defined by* $F(x) = 1 + x$, $G(x) = x + 1$.

The author would like to express his appreciation to Professor Michael Friedberg for his suggestions and criticism.

REFERENCES

1. A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, Volumes I and II, The American Mathematical Society, Mathematical Surveys No. 7, 1961.
2. M. O. Poinsignon Grillet, *Subdivision rings of a semiring*, Fund. Math., **67** (1970), 67-74.
3. P. H. Karvellas, *Inversive semirings*, J. Australian Math. Soc., (to appear).
4. A. B. Paalman de Miranda, *Topological Semigroups*, Mathematisch Centrum, Amsterdam, 1964.
5. K. R. Pearson, *Certain topological semirings in R_1*, J. Australian Math. Soc., **8** (1968), 171-182.
6. ———, *The three kernels of a compact semiring*, J. Australian Math. Soc., **10** (1969), 299-319.
7. J. Selden, *A note on compact semirings*, Proc. Amer. Math. Soc., **15** (1964), 882-886.
8. A. D. Wallace, *The structure of topological semigroups*, Bull. Amer. Math. Soc., **61** (1955), 95-122.

Received March 28, 1973 and in revised form July 24, 1973.

UNIVERSITY OF ALBERTA, CANADA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

STIELTJES DIFFERENTIAL-BOUNDARY OPERATORS, II

ALLAN M. KRALL

The differential boundary system

$$Ly = (y + H[Cy(0) + Dy(1)] + H_1\Psi)' + Py ,$$

$$Ay(0) + By(1) + \int_0^1 dK(t)y(t) = 0 ,$$

$$\int_0^1 dK_1(t)y(t) = 0 ,$$

and its adjoint system are written as Stieltjes integral equation systems with end point boundary conditions. Fundamental matrices are exhibited and, from these, a spectral analysis and a Green's matrix are produced. These are used to achieve spectral resolutions in both self-adjoint and nonself-adjoint situations.

1. Introduction. This article is a continuation of [2] and [6] which showed the density of the domain of L in $\mathscr{L}_n^p[0, 1]$, $1 \leqq p < \infty$, when the boundary functionals satisfied certain conditions, and which derived the dual operator in $\mathscr{L}_n^q[0, 1]$, $1/p + 1/q = 1$, in those circumstances. Rather than repeat those results, we prefer to refer the reader to the articles mentioned. For our purposes here it is sufficient to state that y is an n dimensional vector in $\mathscr{L}_n^p[0, 1]$; A and B are $m \times n$ matrices, $m \leqq 2n$, such that rank $(A\colon B) = m$; C and D are $(2n - m) \times n$ matrices such that $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is nonsingular; K is an $m \times n$ matrix valued function of bounded variation such that the measure it generates satisfies $dK(0) = A$, $dK(1) = B$; K_1 is an $r \times n$ matrix valued function of bounded variation which is not absolutely continuous, satisfying $dK_1(0) = 0$, $dK_1(1) = 0$; H and H_1 are, respectively, $n \times (2n - m)$ and $n \times s$ matrix valued functions of bounded variation, H_1 not absolutely continuous; P is a continuous $n \times n$ matrix; and, finally, Ψ is an s dimensional constant vector.

Because we wish to exhibit the contributions of K, K_1, H, H_1 at 0 and 1 separately, integrals involving their resulting measures will not include contributions at 0 or 1. At all other points, however, we do assume that these functions are regular as defined by Hildebrandt [4]. This results in considerable simplification throughout. Of course, all integrals are Lebesgue or Lebesgue-Stieltjes integrals.

It is convenient to note that the adjoint system has the form

$$L^*z = -(z + K^*[\tilde{A}z(0) + \tilde{B}z(1)] + K_1^*\phi)' + P^*z ,$$

$$\tilde{C}z(0) + \tilde{D}z(1) + \int_0^1 dH^*(t)z(t) = 0 ,$$

$$\int_0^1 dH_1^*(t)z(t) = 0 ,$$

where ϕ is an r dimensional constant vector, and $\tilde{A}$, $\tilde{B}$, $\tilde{C}$, $\tilde{D}$ satisfy

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} -\tilde{A}^* & -\tilde{C}^* \\ \tilde{B}^* & \tilde{D}^* \end{pmatrix} = \begin{pmatrix} -\tilde{A}^* & -\tilde{C}^* \\ \tilde{B}^* & \tilde{D}^* \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = I_{2n} .$$

2. Integral equation representation. Let us make the following definitions. Let

$$\begin{aligned}
\xi_1 &= y , \\
\xi_2 &= Ay(0) + \int_0^t dK(x)y(x) , \\
\xi_3 &= Cy(0) + Dy(1) , \\
\xi_4 &= \int_0^t dK_1(x)y(x) , \\
\xi_5 &= \Psi .
\end{aligned}$$

Then the equation $Ly = 0$, together with the boundary conditions is equivalent to the system

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix}(t) = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix}(0) + \int_0^t d\begin{pmatrix} -Q & 0 & -H & 0 & -H_1 \\ K & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ K_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}(x)\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix}(x) ,$$

where $Q(t) = \int_0^t P(x)dx$,

$$\begin{pmatrix} A & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ C & 0 & -\frac{1}{2}I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix}(0) + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ B & I & 0 & 0 & 0 \\ D & 0 & -\frac{1}{2}I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \end{pmatrix}(1) = 0 .$$

If $M(t)$ represents the Stieltjes measure in the integral equation, then Hildebrandt's $\Delta M^{\pm}(t)$ has zero entries along the diagonal. Hence $I \pm \Delta M^{\pm}$ is always nonsingular.

The adjoint system $L^*z = 0$, together with the boundary conditions is

$$\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{pmatrix}(t) = \begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{pmatrix}(0) - \int_0^t d\begin{pmatrix}-Q^* & K^* & 0 & K_1^* & 0\\ 0 & 0 & 0 & 0 & 0\\ -H^* & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ -H^* & 0 & 0 & 0 & 0\end{pmatrix}(x)\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{pmatrix}(x)\,,$$

$$\begin{pmatrix}I & A^* & C^* & 0 & 0\\ 0 & 0 & -D^* & 0 & 0\\ 0 & 0 & I & 0 & 0\\ 0 & 0 & 0 & 0 & I\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{pmatrix}(0) + \begin{pmatrix}0 & 0 & -C^* & 0 & 0\\ I & -B^* & D^* & 0 & 0\\ 0 & 0 & I & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & I\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{pmatrix}(1) = 0\,.$$

These representations should be compared to those found in [5] which they generalize under certain conditions.

In addition we note that the problem $Ly = \lambda y$ has a similar representation. The only change necessary is to replace $Q(t) = \int_0^t P(x)dx$ by $Q(t) - \lambda t$. The nonhomogeneous problem $Ly = f$ has a representation as a nonhomogeneous integral equation with an additional term

$$F(t) = \int_0^t \begin{pmatrix} f\\ 0\\ 0\\ 0\\ 0\end{pmatrix}(x)dx$$

on the right side.

3. Fundamental matrices. We can express the homogeneous integral problem generated by $(L - \lambda I)y = 0$ together with the boundary conditions in a more compact way by the expressions

$$\xi(t) = \xi(0) + \int_0^t dM_\lambda(x)\xi x\,,$$

$$R\xi(0) + S\xi(1) = 0\,;$$

likewise the adjoint system by

$$\eta(t) = \eta(0) - \int_0^t dM_\lambda^*(x)\eta(x)\,,$$

$$\tilde{R}\eta(0) + \tilde{S}\eta(1) = 0\,.$$

We shall assume in addition that $M_\lambda(t)$ is regular:

$$M_\lambda(t) = 1/2[M_\lambda(t+) + M_\lambda(t-)] ,$$
$$M(0) = M(0+) , \qquad M(1) = M(1-) .$$

Hildebrandt [4] and Vejvoda and Tvrdy [8] have shown that under these conditions the first integral equation has a solution given by $\xi(t) = U_\lambda(0, t)\xi(0)$, where $U_\lambda(s, t)$ is the uniform limit of Picard-like approximations beginning with I (hence U_λ is analytic in λ) satisfying

$$U_\lambda(s, t) = I + \int_s^t dM_\lambda(x) U_\lambda(s, x) .$$

U_λ has the additional properties $U_\lambda(t, t) = I$, and $U_\lambda(r, t) U_\lambda(s, r) = U_\lambda(s, t)$. U_λ is therefore a fundamental matrix when M_λ is absolutely continuous.

Similarly the adjoint equation has a solution given by $\eta(t) = V_{\lambda^*}(0, t)\eta(0)$, where $V_{\lambda^*}(s, t)$ satisfies

$$V_{\lambda^*}(s, t) = I - \int_s^t dM_\lambda^*(x) V_{\lambda^*}(s, x) ,$$

$V_{\lambda^*}(t, t) = I$, $V_{\lambda^*}(r, t) V_{\lambda^*}(s, r) = V_{\lambda^*}(s, t)$.

Since M_λ is regular, it is possible to show that U_λ and V_{λ^*} are related through the formula

$$U_\lambda(s, t) = V_\lambda(t, s) .$$

Hence $U_\lambda(s, t)^{-1} = V_\lambda(s, t)$. Regularity, however, is not inherited from M_λ unless $(\Delta^+ M_\lambda)^2 \equiv 0$. This occurs only when $\Delta^+ K \Delta^+ H \equiv 0$, $\Delta^+ K_1 \Delta^+ H \equiv 0$, $\Delta^+ K \Delta^+ H_1 \equiv 0$, $\Delta^+ K_1 \Delta^+ H_1 \equiv 0$, and will not be necessary.

The fundamental matrices U_λ and V_λ may be easily calculated in the same way as was done in [5]. If $Y(t)$ is a fundamental matrix for $Y' + PY = 0$ satisfying $Y(0) = I$, and

$$\mathscr{H}(t) = \int_0^t e^{-\lambda x} Y(t) Y(x)^{-1} dH(x) ,$$

$$\mathscr{H}_1(t) = \int_0^t e^{-\lambda x} Y(t) Y(t)^{-1} dH_1(x) ,$$

$$\mathscr{K}(t) = \int_0^t dK(x) e^{\lambda x} Y(x) ,$$

$$\mathscr{K}_1(t) = \int_0^t dK_1(x) e^{\lambda x} Y(x) ,$$

$$\mathscr{L}(t) = \int_0^t dK(z) \int_0^z e^{\lambda(z-x)} Y(z) Y(x)^{-1} dH(x) ,$$

$$\mathscr{L}_{01}(t) = \int_0^t dK(z) \int_0^z e^{\lambda(z-x)} Y(z) Y(x)^{-1} dH_1(x) ,$$

$$\mathscr{L}_{10}(t) = \int_0^t dK_1(z) \int_0^z e^{\lambda(z-x)} Y(z) Y(x)^{-1} dH(x) ,$$

$$\mathscr{L}_{11}(t) = \int_0^t dK_1(z) \int_0^z e^{\lambda(z-x)} Y(z) Y(x)^{-1} dH_1(x) ,$$

and $\mathscr{M}(t)$, $\mathscr{M}_{01}(t)$, $\mathscr{M}_{10}(t)$, $\mathscr{M}_{11}(t)$ are defined by the same formulae as $\mathscr{L}(t)$, $\mathscr{L}_{01}(t)$, $\mathscr{L}_{10}(t)$, $\mathscr{L}_{11}(t)$ with only the limits of integration with respect to x changed to from z to t, then

$$U_\lambda(0, t) = \begin{pmatrix} e^{\lambda t} Y(t) & 0 & -e^{\lambda t}\mathscr{H}(t) & 0 & -e^{\lambda t}\mathscr{H}_1(t) \\ \mathscr{K}(t) & I & -\mathscr{L}(t) & 0 & -\mathscr{L}_{01}(t) \\ 0 & 0 & I & 0 & 0 \\ \mathscr{K}_1(t) & 0 & -\mathscr{L}_{10}(t) & I & -\mathscr{L}_{11}(t) \\ 0 & 0 & 0 & 0 & I \end{pmatrix},$$

and

$$V_\lambda(0, t) = \begin{pmatrix} e^{-\lambda t} Y(t)^{-1} & 0 & Y(t)^{-1}\mathscr{H}(t) & 0 & Y(t)\mathscr{H}_1(t) \\ -\mathscr{K}(t)e^{-\lambda t}Y(t)^{-1} & I & -\mathscr{M}(t) & 0 & -\mathscr{M}_{01}(t) \\ 0 & 0 & I & 0 & 0 \\ -\mathscr{K}_1(t)e^{-\lambda t}Y(t)^{-1} & 0 & -\mathscr{M}_{10}(t) & I & -\mathscr{M}_{11}(t) \\ 0 & 0 & 0 & 0 & I \end{pmatrix}.$$

By applying the boundary condition of U_λ the following theorem immediately follows.

THEOREM 3.1. *If $Y(t)$ is a fundamental matrix for $Y' + PY = 0$ satisfying $Y(0) = I$, then the system*

$$Ly = \lambda y ,$$

$$Ay(0) + By(1) + \int_0^1 dK(t)y(t) = 0 ,$$

$$\int_0^1 dK_1(t)y(t) = 0$$

is compatible if and only if the rank of

$$\begin{pmatrix} A & -I & 0 & 0 & 0 \\ Be^\lambda Y(1) + \mathscr{K}(1) & I & -Be^\lambda\mathscr{H}(1) - \mathscr{L}(1) & 0 & -Be^\lambda\mathscr{H}_1(1) - \mathscr{L}_{01}(1) \\ De^\lambda Y(1) + C & 0 & -De^\lambda\mathscr{H}(1) - I & 0 & -De^\lambda\mathscr{H}_1(1) \\ 0 & 0 & 0 & I & 0 \\ \mathscr{K}_1(1) & 0 & -\mathscr{L}_{10}(1) & I & -\mathscr{L}_{11}(1) \end{pmatrix}$$

is less than $3n + r + s$. If $m = n$, the system is compatible if and only if the determinant of the matrix above is zero.

We shall assume throughout the remainder of this article that $m = n$ in order to derive eigenfunction expansions under various conditions.

4. The Green's matrix. Whenever the homogeneous problem is not comparable, the nonhomogeneous problem possesses a unique solution generated by a Green's matrix, just as is the case for the regular Sturm-Liouville problem. Hildebrandt [4] shows that the solution to

$$\xi(t) = \int_0^t dM_\lambda(s)\xi(s) + \mathscr{F}(t) ,$$

$$\xi(0) = \mathscr{F}(0)$$

is given by

$$\xi(t) = U_\lambda(0, t)\mathscr{F}(0) + \int_0^t U_\lambda(s, t)d\mathscr{F}(s)$$

whenever $\Delta^{\pm}\mathscr{F} \equiv 0$. Since in our situation $\mathscr{F}(t) = F(t) + \xi(0)$, where $F(t)$ is absolutely continuous, $F'(t) = f_0(t) = (f(t), 0 \cdots 0)^T$, we find that the solution can be expressed by

$$\xi(t) = U_\lambda(t, 0)y(0) + \int_0^t U_\lambda(s, t)f_0(s)ds .$$

If $\xi(1)$ is calculated and $R\xi(0) + S\xi(1)$ is set equal to 0, $\xi(0)$ is determined, and the solution takes the form

$$\xi(t) = \int_0^1 \mathscr{G}_\lambda(s, t)f_0(s)ds ,$$

where the Green's function $\mathscr{G}$ is given by

$$\begin{aligned} \mathscr{G}_\lambda(s, t) &= U(0, t)[R + SU_\lambda(0, 1)]^{-1}RU_\lambda(0, s)^{-1},\ s < t ,\\ &= -U(0, t)[R + SU_\lambda(0, 1)]^{-1}SU_\lambda(0, 1)U_\lambda(0, s)^{-1},\ s > t . \end{aligned}$$

This is the same formula as that encountered in the regular Sturm-Liouville problem. The Green's function $\mathscr{G}$ possesses the properties, including the adjoint properties, usually attributed to Green's functions.

We note in particular that λ is in the spectrum of the operator L if and only if

$$\det [R + SU_\lambda(0, 1)] = 0 .$$

Since $[R + SU_\lambda(0, 1)]$ is analytic in λ, this implies that either the entire complex plane is in the point spectrum of L, or else the spectrum of L consists only of isolated eigenvalues, accumulating only at ∞.

5. Self-adjoint Stieltjes differential-boundary expansions. It was shown earlier in [6] that the operator $T = iL$ is self-adjoint in

$\mathscr{L}_n^2[0, 1]$ if and only if

1. $P^* = -P$
2. $m = n,\ r = s.$
3. $K = [BD^* - AC^*]H^*$ a.e.
4. $AA^* = BB^*$
5. $H[CC^* - DD^*] = 0$ a.e.
6. $K_1 = MH_1^*$, where M is a nonsingular $r \times r$ matrix.

This being the case, then the spectrum of T is contained in the real axis. Every point with nonzero imaginary part lies in the resolvent. This implies that $\det[R + U_\lambda(0, 1)S] = 0$ only at isolated real points with ∞ their only limit. An application of the spectral resolution theorem for self-adjoint operators on a Hilbert space results in the following.

THEOREM 5.1. *If T is self-adjoint, then*

1. *The spectrum of T consists of a denumerable set of real eigenvalues, accumulating only at ∞.*

2. *Each eigenvalue corresponds to at most n eigenfunctions. Eigenfunctions corresponding to different eigenvalues are orthogonal.*

3. *For each complex number λ, not an eigenvalue, $(T - \lambda I)^{-1}$ exists and can be represented by a unique linear integral operator*

$$(T - \lambda I)^{-1}f(t) = \int_0^1 G_\lambda(s, t)f(s)ds .$$

4. *The Green's function $G_\lambda(s, t)$ satisfies*

a. *As a function of t, $s \neq t$,*

$$(T - \lambda I)G_\lambda(s, t) = 0 .$$

b. $AG_\lambda(s, 0) + BG_\lambda(s, 1) + \int_0^1 dK(t)G_\lambda(s, t) = 0$ *a.e. in s.*

c. $\int_0^1 dK_1(t)G_\lambda(s, t) = 0$ *a.e. in s.*

d. $G_\lambda(t, s) = G_\lambda^*(s, t)$ *a.e. in s and t.*

e. *The eigenfunctions of T are complete in $\mathscr{L}_n^2[0, 1]$.*

If those corresponding to the same eigenvalue have been made orthonormal (denote them by $\{y_i\}_1^\infty$), then for all f in $\mathscr{L}_n^2[0, 1]$

$$f = \sum_1^\infty (f, y_i)y_i .$$

Operators self-adjoint under a transformation are substantially more complex and will be discussed in a subsequent paper. At this point the existence of such a transformation except in trivial cases is doubtful.

6. Nonself-adjoint Stieltjes differential-boundary expansions. Expansions for nonself-adjoint systems have been derived in certain earlier circumstances. First, for the case where $H = 0$, $H_1 = 0$, $K_1 = 0$ or when $H = 0$, $H_1 = 0$, $K = 0$ (the adjoint of the former), an expansion was derived in [2] using familiar techniques. Second, when $H_1 = 0$, $K_1 = 0$ (so $r = 0$, $s = 0$) and H and K are absolutely continuous, an expansion was derived in [5].

In the present situation troubles arise. The bottom terms in the matrix of Theorem 3.1 do not all asymptotically have nice limits as $\operatorname{Re}(\lambda) \to \infty$, a necessary sort of condition previously. For example, when

$$\begin{aligned} K_{j/6}(t) &= 0,\ 0 \leqq t < \frac{j}{6}\,, \\ &= 1,\ \frac{j}{6} < t \leqq 1\,, \end{aligned}$$

the system

$$Ly = (y + K_{1/6}[y(0) - y(1)] + K_{2/6}\Psi)'$$

$$y(0) + y(1) + \int_0^1 dK_{3/6}y = 0\,,$$

$$\int_0^1 d[K_{4/6} + K_{5/6}]y = 0\,,$$

has eigenvalues which are zeros of the determinant of

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ e^{\lambda} + e^{3\lambda/6} & 1 & -e^{5\lambda/6} - e^{2\lambda/6} & 0 & -e^{4\lambda/6} - e^{\lambda/6} \\ e^{\lambda} + 1 & 0 & -e^{5\lambda/6} - e^{2\lambda/6} & 0 & -e^{4\lambda/6} \\ 0 & 0 & 0 & 1 & 0 \\ e^{4\lambda/6} + e^{5\lambda/6} & 0 & -e^{3\lambda/6} - e^{4\lambda/6} & 1 & -e^{2\lambda/6} - e^{3\lambda/6} \end{bmatrix}.$$

These are $\lambda = (2k + 1)6\pi i$; $k = 0, \pm 1, \cdots$. As $\operatorname{Re}\lambda \to -\infty$, however, the matrix has a singular limit.

However, the system

$$Ly = (y + K_{3/6}\Psi)'$$

$$y(0) + y(1) = 0\,,$$

$$\int_0^1 dK_{3/6}y = 0\,,$$

has as its eigenvalue determining matrix

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -e^{\lambda} & 1 & 0 & 0 & -e^{\lambda/2} \\ 1+e^{\lambda} & 0 & -1 & 0 & -e^{\lambda/2} \\ 0 & 0 & 0 & 1 & 0 \\ -2e^{\lambda/2} & 0 & 0 & 1 & -1 \end{bmatrix} .$$

The eigenvalues are easily seen to be $\lambda = 2k\pi i$, $k = 0, \pm 1, \cdots$. The limit of the matrix above as $\operatorname{Re}\lambda \to -\infty$ is nonsingular. Frankly, the author does not entirely understand what is going on.

It is possible to extend the results of [5] under some rather severe restrictions. Let us assume that $H_1 = 0$ and $K_1 = 0$ so that a 3 dimensional vector representation (with $\xi_4 = 0$ and $\xi_5 = 0$) is possible. In addition assume that H is continuous (or by considering the adjoint problem that K is continuous). One system has the form

$$Ly = (y + H[Cy(0) + Dy(1)])' + Py$$

$$Ay(0) + By(1) + \int_0^1 dKy = 0 .$$

If y is replaced by $\tilde{y}$ under the invertable transformation $y = Y\tilde{y}$ $(Y' + PY = 0)$, then we find the equations $Ly = f$, $Ly = \lambda y$ are equivalent to

$$\left(\tilde{y} + \left[Y^{-1}H - \int_0^t Y^{-1}Pdx\right][CY(0)\tilde{y}(0) + DY(1)\tilde{y}(1)]\right)' = Y^{-1}f \text{ or } = \lambda\tilde{y} .$$

The new equations are of the same form as the old, with the same assumptions, with the absence in the second set of the term Py. This results in an equivalent system in which the terms Y and Y^{-1} are missing, a considerable simplification in calculation. We shall henceforth assume that $P = 0$.

The following lemma is the analog of Lemmas 6.4–6.8 of [5].

LEMMA 6.1. (a) $\lim_{\operatorname{Re}(\lambda)\to\infty} \mathscr{H}(t) = 0$ a.e. *In particular* $\lim_{\operatorname{Re}(\lambda)\to\infty} \mathscr{H}(1) = 0$.

(b) $\lim_{\operatorname{Re}(\lambda)\to\infty} e^{\lambda t}[\mathscr{H}(1) - \mathscr{H}(t)] = 0$ a.e.

(c) $\lim_{\operatorname{Re}(\lambda)\to\infty} e^{-\lambda t}\mathscr{K}(t) = 0$ a.e. *In particular* $\lim_{\operatorname{Re}(\lambda)\to\infty} e^{-\lambda}\mathscr{K}(1) = 0$.

(d) $\lim_{\operatorname{Re}(\lambda)\to\infty} [\mathscr{K}(t) \cdot \mathscr{H}(1) - \mathscr{L}(t)] = 0$ a.e.

(e) $\lim_{\operatorname{Re}(\lambda)\to\infty} \mathscr{M}(t) = 0$ a.e. *In particular* $\lim_{\operatorname{Re}(\lambda)\to\infty} \mathscr{M}(1) = 0$.

Proof. Let V_α^β stand for the total variation from α to β.

(a) If $0 < a < t$, then for an appropriate norm

$$\begin{aligned}\|\mathscr{H}(t)\| &= \left\|\int_0^t e^{-\lambda x}d\mathscr{H}(x)\right\| \\ &\leqq \left\|\int_0^a e^{-\lambda x}d\mathscr{H}(x)\right\| + \left\|\int_a^t e^{-\lambda x}d\mathscr{H}(x)\right\| \\ &\leqq V_0^a\|\mathscr{H}\| + e^{-a\lambda}V_a^t\|\mathscr{H}\| .\end{aligned}$$

The first can be made less than half of any preassingned ε if a is sufficiently close to 0. The second is less than $\varepsilon/2$ if $\mathrm{Re}(\lambda)$ is sufficiently large.

(b) $$\begin{aligned}\|e^{\lambda t}[\mathscr{H}(1) - \mathscr{H}(t)]\| &= \left\|e^{\lambda t}\int_t^1 e^{-\lambda x}d\mathscr{H}(x)\right\| \\ &\leqq \left\|e^{\lambda t}\int_{t+\delta}^1 e^{\lambda x}d\mathscr{H}(x)\right\| + \left\|e^{\lambda t}\int_t^{t+\delta} e^{\lambda x}d\mathscr{H}(x)\right\|\end{aligned}$$

when $t \leqq t + \delta \leqq 1$. The second term is less than $V_t^{t+\delta}\|\mathscr{H}\|$. This can be made less than any $\varepsilon/2$ by choosing δ small. The first is bounded by $e^{-\lambda\delta}V_0^1\|\mathscr{H}\|$ which becomes small as $\mathrm{Re}(\lambda) \rightarrow \infty$.

(c) This is shown by the same technique as was used in (a).

(d) $$\begin{aligned}\|\mathscr{K}(t)\mathscr{H}(1) - \mathscr{L}(t)\| &= \left\|\int_0^t d\mathscr{K}(z)\int_z^1 e^{\lambda(z-x)}d\mathscr{H}(x)\right\| \\ &\leqq \left\|\int_0^t d\mathscr{K}(z)\int_{z+\delta}^1 e^{\lambda(z-x)}d\mathscr{H}(x)\right\| \\ &\quad + \left\|\int_0^t d\mathscr{K}(z)\int_z^{z+\delta} e^{\lambda(z-x)}d\mathscr{H}(x)\right\| .\end{aligned}$$

The second term is bounded by $V_0^1\|\mathscr{K}\| \cdot \sup_z V_z^{z+\delta}\|\mathscr{H}\|$. Since $\mathscr{H}$ is continuous on $[0, 1]$ this can be made uniformly small if δ is sufficiently close to 0. The first term is then bounded by $e^{-\lambda\delta}V_0^1\|\mathscr{K}\|$ $V_0^1\|\mathscr{H}\|$ which has zero limit as $\mathrm{Re}(\lambda) \rightarrow \infty$.

(e) This is shown by the same technique as was used in (d).

It is now possible to determine the location of the eigenvalues of L.

THEOREM 6.2. *The eigenvalues of L are the zeros of the determinant of*

$$\varDelta_1 = \begin{pmatrix} A & -I & 0 \\ Be^{\lambda} + \mathscr{K}(1) & I & -Be^{\lambda}\mathscr{H}(1) - \mathscr{L}(1) \\ De^{\lambda} + C & 0 & -De^{\lambda}\mathscr{H}(1) - I \end{pmatrix}.$$

If A is nonsingular, they are bounded on the left in the complex plane. If B is nonsingular, they are bounded on the right in the complex plane. Hence when both A and B are nonsingular, the eigenvalues of L be in a vertical strip.

Since $\det\varDelta_1$ *is almost periodic in* $\mathrm{Im}(\lambda)$, *when A and B are nonsingular, the number of zeros lying in a vertical strip* $|\mathrm{Re}(\lambda)| < h$ *also satisfying* $k < \mathrm{Im}(\lambda) < k + 1$ *is bounded by some number* *independent of* $\measuredangle$. *For any* $\delta > 0$ *there is a corresponding* $m(\delta) \gg 0$ *such that*

$$|\det \Delta_1| > m(\delta)$$

for λ *lying in the strip* $|\mathrm{Re}(\lambda)| < h$ *and outside circles of radius* δ *with centers at the zeros of* $\det \Delta_1$.

Proof. An elementary calculation shows, when A is nonsingular, that as $\mathrm{Re}(\lambda) \to -\infty$, $\det \Delta_1 = (\det A + o(1))$, which ultimately cannot be zero. Similarly, using Lemma 6.1, when B is nonsingular, as $\mathrm{Re}(\lambda) \to \infty$, $\det \Delta_1 = -e^{\lambda}(\det B + o(1))$, which is also ultimately nonzero. The final statements follow from [7, pp. 264–269].

We are now in a position to quote directly the results in §6 of [5]. Please note that the phrases "uniformly in $\cdots$" appearing there should be replaced by "for all x, ξ in $(0, 1)$". Actually a.e. will do fine. Such is our present situation. Assuming A and B are nonsingular, we quote:

THEOREM 6.3. *Let* λ_0 *be in the resolvent set for* L. *Let* $\{\lambda_i\}_1^\infty$ *be the eigenvalues of* L (*which for convenience we assume to be simple*). *Let* $\{Y_i\}_1^\infty$ *and* $\{Z_i\}_1^\infty$ *be the associated eigenfunctions and adjoint eigenfunctions, assuming that* $\int_0^1 Z_i^* Y_i dx = 1$. *Then the Green's function for* L, $G_{\lambda_0}(s, t) = \mathscr{G}_{11}(s, t)$ *satisfies*

$$G_{\lambda_0}(s, t) = \sum_{i=1}^{\infty} \frac{Y_i(t) Z_i^*(s)}{\lambda_i - \lambda_0} \quad \text{a.e.}$$

The proof is by contour integration using the asymptotic estimates established in this section as well as that in [5, §6], suitably avoiding the zeros of $\det \Delta_1$ as we know is possible.

By integrating $G_{\lambda_0}(s, t) \cdot f(s)$ with respect to s before the contour approaches ∞ and appealing to the Lebesgue dominated convergence theorem, we find:

THEOREM 6.4. *Let* f *in* $\mathscr{L}_n^p[0, 1]$ *be in the domain of* L, *then*

$$f(t) = \sum_{i=1}^{\infty} Y_i(t) \int_0^1 Z_i^*(s) f(s) ds .$$

COROLLARY 6.5. *If* f *in* $\mathscr{L}_n^p[0, 1]$ *is in the domain if* L *and* g *in* $\mathscr{L}_n^q[0, 1]$ *is in the domain of* L^*, *then* (*Parseval's Equality*)

$$\int_0^1 g^*(t) f(t) dt = \sum_{i=1}^{\infty} \int_0^1 g^*(t) Y_i(t) dt \int_0^1 Z_i^*(s) f(s) ds .$$

The problem of expansions in the general case remains open.

References

1. R. C. Brown and A. M. Krall, *Adjoints of multipoint-integral boundary value problems*, Proc. Amer. Math. Soc., **37** (1973), 213-216.
2. ———, *Ordinary differential operators under Stieltjes boundary conditions*, submitted for publication.
3. R. C. Brown, G. B. Green, and A. M. Krall, *Eigenfunction expansions under multipoint-integral boundary conditions*, Ann. Mat. Pura. Appl., to appear.
4. T. H. Hildebrandt, *On systems of linear differentio-Stieltjes-integral equations*, Illinois J. Math., **3** (1959), 352-373.
5. A. M. Krall, *Differential-boundary operators*, Trans. Amer. Math. Soc., **154** (1971), 429-458.
6. ———, *Stieltjes differential-boundary operators*, Proc. Amer. Math. Soc., to appear.
7. B. Ja. Levin, *Distribution of Zeros of Entire Functions*, Amer. Math. Soc., Providence, R.I., 1964.
8. O. Vejvoda and M. Tvrdy, *Existence of solutions to a linear integro-boundary-differential equation with additional conditions*, Ann. Mat. Pura Appl., **89** (1971), 169-216.

Received March 7, 1974.

THE PENNSYLVANIA STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

ON THE INNER APERTURE AND INTERSECTIONS OF CONVEX SETS

D. G. LARMAN

If $C_1, \cdots, C_n$ are n convex surfaces or sets in d-dimensional Euclidean space E^d, then it is of some interest to study the invariance properties of $\bigcap_{i=1}^n (C_i + \boldsymbol{a}_i)$ for all choices of vectors $\boldsymbol{a}_i$ in E^d. Such considerations occur naturally in identifying an object irrespective of the direction in which it approaches the observer.

For example, Melzak [2] and Lewis [1] have investigated the conditions under which the intersection $\bigcap_{i=1}^d (C_i + \boldsymbol{a}_i)$ of certain convex surfaces always is a single point. These surfaces arise from the work of Ratcliff and Hartline [3] concerning varying light intensities upon different visual elements of the eye.

In this article we study such intersections and in Theorem 1, we show that the result of Melzak [1] has an associated Helly number in E^2 but not in E^3. In Theorem 2 we give a necessary and sufficient condition for $\bigcap_{i=1}^n C_i + \boldsymbol{a}_i$ to be nonempty, whenever $C_1, \cdots, C_n$ are convex sets, in terms of the outward normals. This condition is not easy to apply in that it involves the outward normals to intersections of d-membered subsets. So in Theorem 3 we give a sufficient condition in terms of inner and outer apertures which is widely applicable. Finally, in Theorem 4, we give a characterization of the sets which can arise as inner apertures. I am indebted to Z. A. Melzak for suggesting these problems to me.

To define the inner and outer aperture, let D be a convex subset of E^d. If $l \equiv l(\boldsymbol{u}, \boldsymbol{v})$,

$$l = \{\boldsymbol{u} + \lambda \boldsymbol{v}, \lambda \geqq 0\}$$

is a typical ray in E^d, $\boldsymbol{u}, \boldsymbol{v} \in E^d$, $\boldsymbol{v} \neq \boldsymbol{o}$, define

$$\theta(\lambda, D) = \text{dist.}\,\{\boldsymbol{u} + \lambda \boldsymbol{v}, E^d \backslash D\}$$

and

$$\theta(D) = \sup_{\lambda \geqq 0} \theta(\lambda)$$

where

$$\text{dist.}\,\{A, B\} = \inf_{\substack{\boldsymbol{a} \in A \\ \boldsymbol{b} \in B}} ||\boldsymbol{a} - \boldsymbol{b}||$$

when A, B are nonempty subsets of E^d. The inner aperture $\mathscr{I}(D)$ of D is the union of those rays $l(\boldsymbol{u}, \boldsymbol{v}) - \boldsymbol{u}$ emanating from the origin

$\boldsymbol{o}$ such that $\theta(l(\boldsymbol{u}, \boldsymbol{v}), D) = +\infty$. So, if D contains $\boldsymbol{o}$, $\mathscr{I}(D)$ is the union of those rays $l \equiv l(\boldsymbol{o}, \boldsymbol{u})$ in D such that $\lambda\boldsymbol{u}$ can be made an arbitrarily large distance from the boundary of D for λ sufficiently large. The outer cone $O(D)$ of D is what is usually known as the characteristic cone namely the set of all rays $l(\boldsymbol{u}, \boldsymbol{v}) - \boldsymbol{u}$ emanating from $\boldsymbol{o}$ with $l(\boldsymbol{u}, \boldsymbol{v})$ contained in D. Both $O(D)$ and $\mathscr{I}(D)$ are convex cones and $O(D)$ is closed whenever D is closed. In general, of course, $O(D)$ can be any convex cone in E^d but this is not the case for $\mathscr{I}(D)$. It will follow from Theorem 4 that $\mathscr{I}(D)$ is a G_δ-convex cone with the property that whenever a ray $l \in \mathrm{cl.}\,\{\mathscr{I}(D)\}\backslash\mathscr{I}(D)$ then the smallest exposed face $F(l)$ of $\mathrm{cl.}\,\{\mathscr{I}(D)\}$ containing l also is contained in $\{\mathrm{cl.}\,\mathscr{I}(D)\}\backslash\mathscr{I}(D)$.

THEOREM 1. *Let $C_1^*, \cdots, C_n^*$ be n convex sets in E^d whose d-dimensional interiors are nonempty and do not contain a line. Let $C_1, \cdots, C_n$ be the convex surfaces bounding $C_1^*, \cdots, C_n^*$ respectively. Then $\bigcap_{j=1}^n (C_j + \boldsymbol{a}_j)$ is at most a single point for all choices $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ of points in E^d if and only if there does not exist n parallel lines of support $l_1, \cdots, l_n$ to $C_1^*, \cdots, C_n^*$ respectively. In E^2 this is true if and only if some four membered subset $C_{j_1}^*, \cdots, C_{j_4}^*$ do not have parallel lines of support. However, in E^3 and for every $n \geqq 3$ there exist convex sets $C_1^*, \cdots, C_n^*$, whose relative interiors do not contain a line, such that every $n-1$ membered subset have parallel lines of support but this is not so for $C_1^*, \cdots, C_n^*$.*

LEMMA 1. *Let $A_1, \cdots, A_n$ be spherically convex subsets (possibly open, half-open or closed semicircles) of the unit circle S^1 such that*

$$\bigcap_{\nu=1}^{4} (A_{i_\nu} \cup - A_{i_\nu}) \neq \varnothing,\ 1 \leqq i_\nu \leqq n,\ \nu = 1, \cdots, 4\,.$$

Then

$$\bigcap_{i=1}^{n} (A_i \cup - A_i) \neq \varnothing\,.$$

Proof. We parametrise S^1 in terms of the angle θ made with some fixed line through the origin and consider the semicircular interval $[0, \pi]$. The intersection $A_i \cup - A_i$ with $[0, \pi]$ is either

(i) an interval $\langle c_i, d_i \rangle$ not containing either 0 or π,

or (ii) $[0, \pi]$,

or (iii) two intervals $[0, a_i >, < b_i, \pi]$, the first containing 0 and the second containing π.

The classification yields a corresponding subdivision I_1, I_2, I_3 of $\{1, \cdots, n\}$. Let

$$[0, a_{i_1}\rangle = \bigcap_{i \in I_3} [0, a_i\rangle$$
$$\langle b_{i_2}, \pi] = \bigcap_{i \in I_3} \langle b_i, \pi] \ .$$

If $\langle c_i, d_i\rangle$ and $\langle c_j, d_j\rangle$, $i, j \in I_1$ both meet $[0, a_{i_1}\rangle$ and

$$(1) \qquad \langle c_i, d_i\rangle \cap \langle c_j, d_j\rangle \cap [0, a_{i_1}\rangle = \emptyset$$

then at least one of these intervals is contained in $[0, a_{i_1}\rangle$. But then

$$(A_i \cup - A_i) \cap (A_j \cup - A_j) \cap (A_{i_1} \cup - A_{i_1}) \cap (A_{i_2} \cup - A_{i_2})$$

is contained in $[0, a_{i_1}\rangle \cup - [0, a_{i_1}\rangle$ and consequently, by (1), is empty, which is contradiction. So, if

$$I_1^1 = \{i \in I_1 : \langle c_i, d_i\rangle \cap [0, a_{i_1}\rangle \neq \emptyset\}$$

we have, from Helly's theorem, that

$$(2) \qquad [0, a_{i_1}\rangle \cap \bigcap_{i \in I_1^1} \langle c_i, d_i\rangle \neq \emptyset \ .$$

Similarly, if

$$(3) \qquad \begin{gathered} I_1^2 = \{i \in I_1 : \langle c_i, d_i\rangle \cap \langle b_{i_2}, \pi] \neq \emptyset\} \\ \langle b_{i_2}, \pi] \cap \bigcap_{i \in I_1^2} \langle c_i, d_i\rangle \neq \emptyset \ . \end{gathered}$$

If there exists $i_3 \in I_1 \backslash I_1^1$ and $i_4 \in I_1 \backslash I_1^2$ then

$$\bigcap_{\nu=1}^{4} A_{i_\nu} \cup - A_{i_\nu} = \emptyset \ ,$$

so either $I_1^1 = I_1$ or $I_1^2 = I_1$ and, using (2) and (3),

$$\bigcap_{i=1}^{n} A_i \cup - A_i \neq \emptyset \ .$$

REMARK. This is the best possible result for if $A_1 = [0, \pi/2]$, $A_2 = [\pi/4, 3\pi/4]$, $A_3 = [\pi/2, \pi]$, $A_4 = [3\pi 4, 5\pi/4]$ then

$$\bigcap_{\nu=1}^{3} A_{i_\nu} \cup - A_{i_\nu} \neq \emptyset, 1 \leqq i_1 < i_2 < i_3 \leqq 4$$

but

$$\bigcap_{i=1}^{4} A_i \cup - A_i = \emptyset \ .$$

LEMMA 2. *There exist n closed spherically convex two dimensional subsets $D_1, \cdots, D_n$ on S^2, none of which contain antipodal points, such that for every $n-1$ membered subset $D_{i_1}, \cdots, D_{i_{n-1}}$ there exists* *a great circle of* S^2 *which meets each* D_{i_ν}, *but there does not exist a great circle meeting each of* $D_1, \cdots, D_n$.

Proof. In [4], Santalo constructs, for each $n \geqq 3$, a family of n compact convex two dimensional sets $F_1, \cdots, F_n$ in E^2 so that each $n-1$ members of the family admit a common transversal but the entire family does not have a common transversal. We mention that such an example is the family of n circular discs whose centers have polar coordinates $\rho = 1$ and $\theta = 2k\pi/n$, $k = 1, \cdots, n$ and whose radii are all equal to $\cos^2 \pi/n$ or $\cos^2 \pi/n + \cos^2 \pi/2n - 1$ according as whether n is even or odd.

Now, if we place the configuration $F_1, \cdots, F_n$ into a plane tangent to S^2, let $D_1, \cdots, D_n$ be the corresponding closed spherically convex subsets of S^2 obtained by the projection of $F_1, \cdots, F_n$ into S^2 from the origin. Clearly $D_1, \cdots, D_n$ satisfy the requirements of the lemma.

Proof of Theorem 1. The proof of the first part is essentially due to Melzak [1] but as he makes the restriction that $d = n$ we repeat the details.

If there exist n parallel lines of support $l_1, \cdots, l_n$ to $C_1^*, \cdots, C_n^*$ respectively then by translating the line l_j into the relative interior of C_j if necessary, $j = 1, \cdots, n$ we obtain n nondegenerate similarly orientated chords $[\boldsymbol{p}_j, \boldsymbol{q}_j]$ of C_j^* parallel to l_j such that

$$||\boldsymbol{p}_1 - \boldsymbol{q}_1|| = \cdots = ||\boldsymbol{p}_n = \boldsymbol{q}_n|| .$$

Hence, if $\boldsymbol{a}_j = \boldsymbol{p}_i - \boldsymbol{p}_j$, $j = 1, \cdots, n$

$$\bigcap_{j=1}^{n} C_j^* + \boldsymbol{a}_j \supset \{\boldsymbol{p}_1, \boldsymbol{q}_1\}$$

and so contains at least two points.

On the other hand, if there exist vectors $\boldsymbol{a}_j$, $j = 1, \cdots, n$ such that $\bigcap_{j=1}^n C_j^* + \boldsymbol{a}_j$ contains at least two points say $\boldsymbol{p}, \boldsymbol{q}$ then, by considering two dimensional sections of C_j, C_j has a line of support l_j parallel to $[\boldsymbol{p}, \boldsymbol{q}]$ and hence $l_1, \cdots, l_n$ are parallel lines of support to $C_1, \cdots, C_n$ respectively which completes the proof of the first part.

In E^2 we may select a set A_i of unit tangent vectors $\boldsymbol{u}$ to C_i^* by ensuring that the outward normal lies on the left hand side of $\boldsymbol{u}$ when viewed from the point of contact on C_i in a clockwise direction. Then A_i is a spherically convex subset of S^1 which is either S^1 or is contained in semicircle according to whether or not C_i is bounded. Now $C_1^*, \cdots, C_n^*$ do not have parallel lines of support if and only if

$$\bigcap_{i=1}^{n} (A_i \cup - A_i) = \varnothing .$$

This, by Lemma 1, is true if and only if there exists some four membered subset of $C_1^*, \cdots, C_n^*$ which do not possess parallel lines of support which completes the proof of the second part of the theorem.

In E^3 and for each $n \geqq 2$ consider the n closed spherically convex subsets $D_1, \cdots, D_n$ of S^2 afforded by Lemma 2. If $\langle , \rangle$ denotes scalar product consider the set of closed half-spaces $\mathscr{H}_i$ such that $H^- \in \mathscr{H}_i$ if

$$H^- = \{\boldsymbol{x}\colon \langle \boldsymbol{x}, \boldsymbol{u} \rangle \leqq 1\} \quad \text{for some} \quad \boldsymbol{u} \in D_i \, .$$

Let

$$C_i^* = \bigcap_{\mathscr{H}_i} H^- \, , \qquad\qquad i = 1, \cdots, n \, .$$

Then D_i is the set of outward normals to C_i^* and so as D_i is two dimensional, C_i^* does not contain a line, $i = 1, \cdots, n$. Also for every $n-1$ membered subset $C_{i_1}^*, \cdots, C_{i_{n-1}}^*$ of $C_1, \cdots, C_n$ the corresponding set of outward normals $D_{i_1}, \cdots, D_{i_{n-1}}$ all meet some great sphere $S \equiv S(i_1, \cdots, i_{n-1})$. Consequently, if l is a line perpendicular to aff. S, $C_{i_1}, \cdots, C_{i_{n-1}}$ each possess lines of support parallel to l.

On the other hand, if $C_1, \cdots, C_n$ possess parallel lines of support then there would exist a great sphers S^1 of S^2 which meets each of $D_1, \cdots, D_n$ which, by Lemma 2, is not so. Hence $C_1, \cdots, C_n$ do not possess parallel lines of support, which completes the proof of Theorem 1.

We observe the following lemma which is easily established by separating two disjoint convex sets by a hyperplane.

LEMMA 3. *Two convex sets C_1, C_2 in E^d cannot be separated by translation if and only if $N(C_1) \cap (-N(C_2)) = \boldsymbol{o}$, where $N(C_i)$ is the convex cone of outward normals to C_i, $i = 1, 2$.*

Using Helly's theorem we readily verify the following lemma.

LEMMA 4. *If $C_1, \cdots, C_n$ are convex sets in E^d, then $\bigcap_{i=1}^n (C_i + \boldsymbol{a}_i) \neq \emptyset$ for all points $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ in E^d if and only if $\bigcap_{\nu=1}^{d+1} (C_{i_\nu} + \boldsymbol{a}_{i_\nu}) \neq \emptyset$ for all points $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ in E^d and for every $d+1$ membered subset $\{C_{i_\nu}\}_{\nu=1}^{d+1}$ of $\{C_i\}_{i=1}^n$.*

Using Lemmas 3 and 4 we obtain

THEOREM 2. *If $C_1, \cdots, C_n$ are convex sets in E^d then $\bigcap_{i=1}^n (C_i + \boldsymbol{a}_i) \neq \emptyset$ for all points $\boldsymbol{a}_1, \cdots \boldsymbol{a}_n$ in E^d if and only if*

$$\{-N(C_{i_1})\} \cap N\Big(\bigcup_{\nu=2}^{d+1} C_{i_\nu}\Big) = \emptyset$$

for all $d+1$ *membered subcollections* $\{C_{i_\upsilon}\}_{\upsilon=1}^{d+1}$ *of* $\{C_i\}_{i=1}^n$.

However, this condition is not completely satisfactory in that $N(\bigcup_{\upsilon=2}^{d+1} C_{i_\upsilon})$ is a function of $\bigcup_{\upsilon=2}^{d+1} C_{i_\upsilon}$ rather than a combination of functions of each C_{i_υ}. We shall resolve this problem to a certain extent in Theorem 3 by giving a widely applicable sufficient condition.

THEOREM 3. *Let* $C_1, \cdots, C_n$ *be* n *convex sets in* E^d. *Then*

$$\bigcap_{i=1}^n (C_i + \boldsymbol{a}_i) \neq \emptyset \tag{4}$$

for all choices of $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ *if there exists* j *such that*

$$O(\text{cl. } C_j) \cap \bigcap_{\upsilon=1}^{d+1} \mathscr{I}(C_{i_\upsilon}) \neq \emptyset$$

for all $d+1$ *membered subcollections* $\{C_{i_\upsilon}\}_{\upsilon=1}^{d+1}$ *of* $\{C_i\}_{\substack{i=1 \\ i \neq j}}^n$. *Further, if at least of* cl. $C_1, \cdots,$ cl. C_n *does not contain a line, each is unbounded and* $C_1, \cdots, C_n$ *cannot be separated by translation,* i.e., (4) *holds for all* $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ *then*

$$\bigcap_{j=1}^n O(\text{cl. } C_j) \neq \emptyset\ .$$

Proof. Let l be a ray of $O(\text{cl. } C_j) \cap \bigcap_{i=1}^n \mathscr{I}(C_i)$ which, by Helly's theorem, is nonempty. We may suppose, without loss of generality, that $\boldsymbol{o} \in C_1 \cap \cdots \cap C_n$. Then, if $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ are points of E^d,

$$l + \boldsymbol{a}_i \subset C_i + \boldsymbol{a}_i\ , \qquad i = 1, \cdots, n\ .$$

If $l = \{\lambda \boldsymbol{u}, \lambda \geqq 0\}$, then, as $l \subset \mathscr{I}(C_i)$, $i \neq j$, there exists λ_i such that $\lambda \boldsymbol{u} + \boldsymbol{a}_j$ is in C_i, $\lambda \geqq \lambda_i$.

So, if $\lambda^* = \max_{1 \leqq i \leqq n}. \lambda_i$,

$$\lambda^* u + a_j \in \bigcap_{i=1}^n C_i \quad \text{as required}\ .$$

To prove the second part, let C_i^* denote the closure of C_i, $i = 1, \cdots, n$. We may assume that C_1 and C_1^* do not contain a line and that for some n, $\bigcap_{i=1}^{n-1} C_i^*$ is unbounded, which is certainly true for $n = 2$. As $\bigcap_{i=1}^{n-1} C_i^*$ is convex closed and unbounded it follows that $O(\bigcap_{i=1}^{n-1} C_i^*)$ is nonempty. Further, as $\bigcap_{i=1}^{n-1} C_i^*$ is contained in C_1^*, $\bigcap_{i=1}^{n-1} C_i^*$ and $O(\bigcap_{i=1}^{n-1} C_i^*)$ do not contain a line. Let l be a ray of $O(\bigcap_{i=1}^{n-1} C_i^*)$, say $l = \{\lambda \boldsymbol{u}, \lambda \geqq 0\}$. If $O(\bigcap_{i=1}^n C_i^*)$ is empty then, in particular, $\bigcap_{i=1}^n C_i^*$ must be a compact convex set.

If $\lambda \geqq 0$,

$$\lambda \boldsymbol{u} + \bigcap_{i=1}^{m-1} C_i \subset \bigcap_{i=1}^{m-1} C_i\ ,$$

and consequently,

$$(5) \qquad \left(\lambda u + \bigcap_{i=1}^{m-1} C_i\right) \cap C_m = \left(\lambda u + \bigcap_{i=1}^{m-1} C_i\right) \cap \left(\bigcap_{i=1}^{m} C_i\right).$$

If no matter how large λ is taken, $(\lambda u + \bigcap_{i=1}^{m-1} C_i) \cap C_m$ contains a point $z(\lambda)$ say then, by (5), $z(\lambda)$ is confined to a compact set $\bigcap_{i=1}^{m} C_i$ and $z(\lambda) - \lambda u \in \bigcap_{i=1}^{m-1} C_i$, $\lambda \geqq 0$. It follows that $-l$ is a ray of $O(\bigcap_{i=1}^{m-1} C_i^*)$ which is a contradiction to C_1^* not containing a line. So $\bigcap_{i=1}^{m} C_i^*$ is an unbounded closed convex set and hence $O(\bigcap_{i=1}^{m} C_i^*)$ is nonempty. So repeating this process for $m = 1, 2, \cdots, n$ we conclude that $O(\bigcap_{i=1}^{n} C_i^*)$ is nonempty as required.

DEFINITION. We say that a collection $\mathscr{H}$ of closed half-spaces in E^d is *closed* if whenever $\{H_i^-\}_{i=1}^{\infty}$ is a sequence of closed half-spaces in $\mathscr{H}$, where

$$H_i^- = \{x: \langle x, u_i \rangle \leqq \alpha_i\},\ u_i \text{ a unit vector},$$

and $u_i \rightarrow u$, $\alpha_i \rightarrow \alpha$ as $i \rightarrow \infty$ then the closed half-space

$$H^- = \{x: \langle x, u \rangle \leqq \alpha\}$$

is in $\mathscr{H}$. We say that a collection $\mathscr{H}$ of closed half-spaces is F_σ if it is the countable union of closed collections.

If $\mathscr{H}$ is a closed collection of closed half-spaces notice that the set $\bigcup_{H^- \in \mathscr{H}} H$, where H is the bounding hyperplane of H^-, is a closed set and consequently $\bigcap_{H^- \in \mathscr{H}} \operatorname{int} H^-$ is a relatively open subset of $\bigcap_{H^- \in \mathscr{H}} H^-$.

THEOREM 4. *A set C in E^d is the inner aperture of some convex subset of E^d if and only if*

$$C = o \cup \bigcap_{\mathscr{H}} \operatorname{int.} H^-$$

where $\mathscr{H}$ is an F_σ-collection of closed half-spaces and $o \in H$, the bounding hyperplane of H^-, for all $H^- \in \mathscr{H}$.

REMARK. So, in particular, C has to be a G_δ-convex cone with apex the origin such that if $x \in \{\text{cl.}\, C\}\backslash C$ then the smallest exposed face $F(x)$ of cl. C that contains x is also contained in $\{\text{cl.}\, C\}\backslash C$. In E^3 the converse is also true.

Proof. We shall assume that the theorem is true in $d - 1$ dimensions, the theorem being trivial for $d = 1$.

(i) *Necessity.* Let C be the inner aperture of some convex set D in E^d where, since $\mathscr{I}(D) = \mathscr{I}(\text{cl.}\, D)$ we may suppose that D is closed. If $D = E^d$ then $C = E^d$ and, by convention,

$$C = \bigcap_{\mathscr{H}} \text{int. } H^- = E^d$$

where $\mathscr{H}$ is the empty set of closed half-spaces.

Otherwise $D \neq E^d$ and so possesses at least one hyperplane of support M say with D contained in the closed half-space M^-. We may suppose, without loss of generality, that $\boldsymbol{o} \in M$. If D contains a (maximal) linear subspace L of dimension at least one then $L \subset M$ and

$$D = F + L$$

where F is a closed convex subset of $L^\perp$. By the inductive assumption the inner aperture $\mathscr{I}(F)$ of F can be written

$$\mathscr{I}(F) = \boldsymbol{o} \cup \bigcap_{\mathscr{H}^*} \text{int. } H^{*-}$$

where $\mathscr{H}^*$ is a closed subset of the closed half-spaces in $L^\perp$. Then

$$C = \boldsymbol{o} \cup \bigcap_{\mathscr{H}} \text{int. } H^-$$

where $\mathscr{H}$ is the closed collection of closed half-spaces in E^d formed by taking H^- in $\mathscr{H}$ if

$$H^- = L + H^{*-}$$

where $H^{*-} \in \mathscr{H}^*$.

If D does not contain a line then the set of rays in D is a closed convex cone K which has a hyperplane of support say $\{x_d = 0\}$ with

$$K \cap \{x_d = 0\} = \boldsymbol{o} .$$

Let π_ν denote the hyperplane $x_d = \nu$, $\nu \geqq 0$. Let l be a typical ray of K,

$$\alpha_\nu(l) = \text{dist. } \{(l\pi_\nu), \pi_\nu(E^d \backslash D)\} ,$$

and

$$\alpha(l) = \sup_{\nu \geqq 0} \alpha_\nu(l) .$$

By considering two dimensional sections through l it is easily verified that $\alpha_\nu(l)$ increases with ν. Also

$$l \subset C \quad \text{if and only if} \quad \alpha(l) = +\infty .$$

So, if

$$C_i = \{l: l \text{ is a ray in } K, \alpha(l) > i\} ,$$

then

$$C = \bigcap_{i=1}^{\infty} C_i \,. \tag{6}$$

Now $C_i K$, $i = 1, 2, \cdots$ and

$$K = \boldsymbol{o} \cup \bigcap_{\mathscr{H}} \operatorname{int.} H^- \tag{7}$$

where $\mathscr{H}$ is the collection of closed half-spaces, whose bounding hyperplanes contain $\boldsymbol{o}$, such that $K \backslash \boldsymbol{o} \subset \operatorname{int.} H^-$. If $\hat{K} = K \cap S^{d-1}$, let $\mathscr{H}_j^*$ denote the closed set of the closed half-spaces H^-,

$$H^- = \{\boldsymbol{x}\colon \langle \boldsymbol{x}, \boldsymbol{u}\rangle \leqq 0\}$$

where

$$\langle -\boldsymbol{u}, \boldsymbol{k}\rangle \leqq -2^{-j} \,, \qquad \text{for all} \quad k \in \hat{K} \,.$$

Then $\mathscr{H} = \bigcup_{j=1}^{\infty} \mathscr{H}_j^*$ and so, using (6), (7) it is enough to show that

$$C_i = K \cap \bigcap_{\mathscr{H}_i} \operatorname{int.} H^-$$

where $\mathscr{H}_i$ is a closed collection of closed half-spaces of E^d whose bounding hyperplanes goes through $\boldsymbol{o}$.

Suppose now that l is a ray of $K \backslash C_i$. Then

$$\alpha(l) \leqq i \,.$$

For $j = 1, 2, \cdots$, there exist points $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots$, with $\boldsymbol{a}_j \in \pi_j \cap \operatorname{bdy.} D$ such that

$$\|\boldsymbol{a}_j - \{\pi_j \cap l\}\| \leqq i \,. \tag{8}$$

Let H_j denote a hyperplane of support to D at $\boldsymbol{a}_j$, with $D \subset H_j^-$. As we may suppose that $K \neq \boldsymbol{o}$, H_j is not parallel to the hyperplane π_1. So $H_j \cap \pi_1$ is a line in π_1. If we consider the two plane σ_j through l and $\boldsymbol{a}_j$ then H_j meets σ_j in a line l_j. As l_j supports $\sigma_j \cap D$, it follows, using (8), that

$$\|l_j \cap \pi_1 - l \cap \pi_1\| \leqq i \,. \tag{9}$$

Consequently the $(d-2)$ affine space $\pi_1 \cap H_j$ lies within a distance i of $l \cap \pi_1$. So we may suppose, by picking subsequences if necessary, that $\pi_1 \cap H_j \to \pi_1 \cap H_0$ as $j \to \infty$ and $l_j \cap \pi_1$ tends to a point which, with a view to later developments, we denote by $l_0 \cap \pi_1$. Let the line through the points $\boldsymbol{a}_j$ and $l_j \cap \pi_1$ be l_j^*, $j = 1, 2, \cdots$. As (8), (9) hold, l_j^* converges to a line l_0 through $l_0 \cap \pi_1$ and parallel to l. Consequently $H_j \to H_0$ as $j \to \infty$. So $D \subset H_0^-$ and

(10) $$||\pi_\nu \cap l_0 - \pi_\nu \cap l|| = \beta \leqq i \, , \qquad \text{if} \quad \nu \geqq 0 \, ,$$

β a constant. We claim that

$$H_0^- + \{\pi_1 l - \pi_1 l_0\} = H_0'^- \text{ say ,}$$

contains K and H_0' supports K and passes through $\boldsymbol{o}$. Certainly

(11) $$l \subset H_0'$$

and so H_0' passes through $\boldsymbol{o}$. If there exists a ray l^* in $K \backslash H_0'^-$, then l^* meets H_0 which contradicts $D \subset H_0^-$.

Now let $\mathscr{H}_i$ denote those closed half-spaces H^- such that the bounding hyperplane H supports K and there exists a closed half-space H^{*-} containing H^- such that H^* supports D; H^* is parallel to H and a distance, in the hyperplane π_1, at most i from H.

By (11),

(12) $$C_i \supset K \cap \bigcap_{\mathscr{H}_i} \text{int. } H^- \, ,$$

where $\mathscr{H}_i$ is a closed set of closed half-spaces.

Conversely, if l is a ray of

$$K \backslash \{K \cap \bigcap_{\mathscr{H}_i} \text{int. } H^-\}$$

then there exists H^- in $\mathscr{H}_i$ such that $l \subset H$. Then there exists a closed half-space H^{*-} which contains D such that H^* is parallel to H and the distance between H and H^* is at most i. Consequently

$$\alpha_\nu(l) \leqq i, \nu \geqq 0$$

and so $l \not\subset C_i$. Hence

(13) $$C_i \subset K \cap \bigcap_{\mathscr{H}_i} \text{int. } H^- \, .$$

Combining (12) and (3),

$$C_i = K \cap \bigcap_{\mathscr{H}_i} \text{int. } H^-$$

which completes the proof of the necessity of the conditions.

(ii) *Sufficiency*. Suppose now that

$$C = \boldsymbol{o} \cup \bigcap_{\mathscr{H}} \text{int. } H^-$$

where $\mathscr{H}$ is an F_σ-collection of closed half-spaces and $\boldsymbol{o} \in H$ for all $H^- \in \mathscr{H}$. So we may write $\mathscr{H} = \bigcup_{i=1}^\infty \mathscr{H}_i$ where the $\mathscr{H}_i$ form an increasing sequence of closed collections.

Consider the closed convex cone

$$C_0 = \text{cl.}\, C = \bigcap_{\mathscr{H}} H^- .$$

If $C_0 = E^d$ then $C = E^d$ and C is its own inner aperture. Otherwise C_0 possesses one hyperplane of support M through $\boldsymbol{o}$ with C_0 contained in the closed half-space M^-. If $M \cap C_0$ contains a maximal linear subspace L of dimension at least 1 then we may write $C_0 = F + L$ where F is a proper closed convex cone in L. Notice that $L \subset H$ for each $H^- \in \mathscr{H}$ and consequently we may write

$$H^- = L + H^{*-} \quad \text{for each} \quad H^- \in \mathscr{H} ,$$

where H^{*-} is a closed half-space in L whose bounding hyperplane H^* passes through $\boldsymbol{o}$. Consequently

$$C = \boldsymbol{o} \cup \{\{\bigcap_{\mathscr{H}} \text{int.}\, H^{*-}\} + L\} .$$

By the inductive assumption, there exists a closed convex set D^* in L such that

$$\boldsymbol{o} \cup \bigcap_{\mathscr{H}} \text{int.}\, H^{*-}$$

is the inner aperture of D^* in L. Let

$$D = D^* + L$$

and then C is the inner aperture of D.

Henceforth therefore we may suppose that C_0 is a proper closed convex cone in E^d i.e., C_0 does not contain a line and we can also suppose that the ray

$$X_d^+ = \{(0, \cdots, 0, x_d),\, x_d \geqq 0\}$$

is in C_0 and that the hyperplane $\pi_0 = \{x_d = 0\}$ supports C_0 with $\pi_0 \cap C_0 = \boldsymbol{o}$. Then, as for K in the proof of necessity,

$$C_0 = \boldsymbol{o} \cup \bigcap_{\mathscr{H}_0} \text{int.}\, H^-$$

where $\mathscr{H}_0$ is a closed set of closed half-spaces whose bounding hyperplanes pass through $\boldsymbol{o}$. We may suppose that

$$\mathscr{H}_0 \subset \mathscr{H}_1 \subset \mathscr{H}_2 \subset \cdots$$

and let

$$C_i = \boldsymbol{o} \cup \bigcap_{\mathscr{H}_i} \text{int.}\, H^- , \qquad i = 0, 1, 2, \cdots .$$

We shall produce inductively a nested sequence of closed convex sets $\{C_i^*\}_{i=0}^{*}$ such that C_i is the inner aperture of C_i^* and indeed

$$C_{i+1}^* = C_i^* \cap \bigcap_{\mathscr{H}_i} H^{*-},\ i \geqq 0 \tag{14}$$

where, if $H^- \in \mathscr{H}_i$ then H^{*-} is that closed half-space containing H^- such that H^* and H are parallel and at a distance i apart in the hyperplane π_1.

We begin the induction by taking

$$C_0^* = \{\boldsymbol{x} = (x_1, \cdots, x_d),\ x_d \geqq 0 \quad \text{and} \quad \text{dist.}\,(\boldsymbol{x}, C_0 \cap \pi_{x_d}) \leqq x_d^{1/2}\}\ .$$

Clearly C_0^* is closed and it is convex since, from above, $C_0^* \cap \pi_\nu$ is convex, $\nu \geqq 0$ and so C_0^* cannot possess a point of concavity. We shall show that

$$\mathscr{I}(C_0^*) = C_0\ . \tag{15}$$

First notice that if $\boldsymbol{u} = (u_1, \cdots, u_d)$ is a unit vector in C_0 then $u_d > 0$. So, if $l = \{\lambda u\colon \lambda \geqq 0\}$ is the corresponding ray in C_0

$$\theta_\lambda = \alpha_{\lambda u_d}(l) \geqq \sqrt{\lambda u_d} > 0\ .$$

So, if m is a positive number

$$\theta_\lambda \geqq m \tag{16}$$

provided $m^2/u_d \leqq \lambda$. It is an almost immediate consequence of (16) that $l \subset \mathscr{I}(C_0^*)$ and hence $C_0 \subset \mathscr{I}(C_0^*)$.

Suppose next that the ray

$$l' = \{\lambda \boldsymbol{v},\ \lambda \geqq 0\}$$

is not in C_0. If $v_d \leqq 0$ then $\lambda \boldsymbol{v} \notin C_0^*$ for all $\lambda > 0$ and then certainly $l' \not\subset \mathscr{I}(C_0^*)$. If $v_d > 0$ then $l' \cap \pi_\nu$ is a single point for each $\nu \geqq 0$ and there exists $\eta > 0$ such that

$$\text{dist.}\,(\boldsymbol{v}, C_0 \cap \pi_{v_d}) > \eta\ .$$

So

$$\text{dist.}\,(\lambda \boldsymbol{v}, C_0 \pi_{\lambda v_d}) > \lambda \eta\ . \tag{17}$$

But, if $l' \subset \mathscr{I}(C_0^*)$ then, in particular, $\lambda \boldsymbol{v} \in C_0^*$ for each $\lambda \geqq 0$. So

$$\text{dist.}\,(\lambda \boldsymbol{v}, C_0 \pi_{\lambda v_d}) \leqq (\lambda v_d)^{1/2},\ \lambda \geqq 0\ . \tag{18}$$

However, provided $\lambda > v_d/\eta^2$ it follows from (17) that (18) is false. Consequently $l' \not\subset \mathscr{I}(C_0^*)$ which establishes (15).

Suppose inductively that for some $m \geqq 1$ we have constructed m closed convex sets $C_0^*, \cdots, C_{m-1}^*$ in E^d with C_i being the inner aperture of C_i^*, $i = 0, \cdots, m - 1$. Indeed,

(19) $$C^*_{i+1} = C^*_i \cap \bigcap_{\mathscr{H}_{i+1}} H^{*-}\ , \qquad i = 0, 1, \cdots, m-2\ ,$$

where, if $H^- \in \mathscr{H}_{i+1}$ then H^{*-} is that closed half-space containing H^- such that H^* and H are parallel and at a distance $i+1$ apart in the plane π_1.

For each $H^- \in \mathscr{H}_m$, let H^{*-} be that closed half-space containing H^- such that H^* and H are parallel and at a distance m apart in the plane π_1. Define

(20) $$C^*_m = C^*_{m-1} \cap \bigcap_{\mathscr{H}_m} H^{*-}\ .$$

We claim that the inner aperture of C^*_m is C_m i.e.,

(21) $$\mathscr{I}(C^*_m) = C_m\ .$$

If l is a ray of C_0 not in C_m then l is in some hyperplane H where $H^- \in \mathscr{H}_m$. Consequently, by considering the corresponding closed half-space H^{*-}, we deduce that $\alpha(l) \leqq m$, and so $l \not\subset \mathscr{I}(C^*_m)$. Hence $\mathscr{I}(C^*_m) \subset C_m$.

On the other hand, suppose that $l \in C_m$. That the set

$$\bigcup_{\mathscr{H}_m} H^* = H_m \text{ say}$$

is a closed set and does not meet the ray $l \backslash \boldsymbol{o}$. As each hyperplane H, with $H^- \in \mathscr{H}_m$, passes through $\boldsymbol{o}$, it follows that

(22) $$\text{dist.}\,(l \cap \pi_\nu, H_m) \longrightarrow +\infty \quad \text{as} \quad \nu \longrightarrow +\infty\ .$$

Also $l \in \mathscr{I}(C^*_{m-1})$ and so

(23) $$\text{dist.}\,(l \cap \pi_\nu, E^d \backslash C^*_{m-1}) \longrightarrow +\infty \quad \text{as} \quad \nu \longrightarrow +\infty\ .$$

Consequently using (20), (22), (23),

$$\text{dist.}\,(l \cap \pi_\nu, E^d \backslash C^*_m) \longrightarrow +\infty \quad \text{as} \quad \nu \longrightarrow +\infty\ .$$

Therefore, $l \subset \mathscr{I}(C^*_m)$ and so $C_m \subset \mathscr{I}(C^*_m)$ which completes the verification of (21).

The results (20), (21) verify (19) for m and we can now suppose that the C^*_m have been defined so that (20), (21) hold for $m = 0, 1, 2, \cdots$. Define

$$C^* = \bigcap_{m=0}^{\infty} C^*_m$$

and we shall show that $\mathscr{I}(C^*) = C$.

Suppose that l is a ray of C_0 not in $\mathscr{I}(C^*)$. Then there exists m such that $\alpha_\nu(l) \leqq m$, $\nu \geqq 0$. So l is not in $\mathscr{I}(C^*_{m+1}) = C_{m+1}$. Consequently l is not in C. So $C \subset \mathscr{I}(C^*)$.

On the other hand, suppose that l is a ray of C_0 which is not in C. Then l is not in C_m for some $m \geqq 0$. So

$$l \not\subset \mathscr{I}(C_m^*) \supset \mathscr{I}(C^*) .$$

Hence $\mathscr{I}(C^*) \subset C$ and this finally establishes that

$$\mathscr{I}(C^*) = C$$

which completes the proof of Theorem 4.

References

1. J. E. Lewis, *On a problem of uniqueness arising in connection with a neurophysical control mechanism*, Submitted to Information and Control.
2. Z. A. Melzak, *On a uniqueness theorem and its application to a neurophysical control mechanism*, Information and Control, **5** (1962), 163-172.
3. F. Ratcliff and H. K. Hartline, *The response of Limus optic nerve fibers to patterns of illumination on the receptor mosaic*, J. Genl. Physiol., **42** (1959), 1241-1255.
4. L. A. Santaló, *Un teorema sobre conjuntos de paralelepipedos de aristas paralelas*, Publ. Inst. Mat. Univ. Nac. Litoral, **2** (1940), 49-60 and **3** (1942), 202-210.

Received April 30, 1973 and in revised form August 21, 1973. Research supported by the National Research Council of Canada, while visiting the University of British Columbia.

UNIVERSITY COLLEGE LONDON

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

ON THE REGULARITY OF THE P^n-INTEGRAL AND ITS APPLICATION TO SUMMABLE TRIGONOMETRIC SERIES

S. N. MUKHOPADHYAY

The symmetric P^{2m}-integral (and P^{2m+1}-integral) as defined by R. D. James in "Generalized nth primitives", Trans. Amer. Math. Soc., 76 (1954), is useful to solve problems relating to trigonometric series (see R. D. James: Summable trigonometric series, Pacific J. Math., 6 (1956)). But the definition of the integral is not valid, since Lemma 5.1 of the former paper of James, which is the basis of the whole theory, is incomplete due to the fact that the difference of two functions having property B_{2m-2} may not have this property. Therefore, all the subsequent results of James also remain incomplete and a complete systematic definition of the integral is needed.

In the present paper a definition of the P^{2m}-integral (and P^{2m+1}-integral) is given and it is shown that all the results of the later paper of James remain valid with this integral.

1. Definitions and Notations. Most of the definitions and notations of [8] will be used with essential modifications. The generalized symmetric derivative [8] (also called symmetric de La Vallée Poussin derivative [18]) of even and odd orders and the generalized unsymmetric derivative [8] (also called Peano derivative [13] or unsymmetric de La Vallée Poussin derivative [11]) of a function f at x_0 will be denoted by $D^r f(x_0)$ and $f_{(r)}(x_0)$ respectively, where r denotes the order of the respective derivatives. If $D^{2k}f(x_0)$ exists, $0 \leqq k \leqq m-1$, define $\theta_{2m}(f; x_0, h)$ by

$$\frac{h^{2m}}{(2m)!}\theta_{2m}(f; x_0, h) = \frac{1}{2}\{f(x_0+h) + f(x_0-h\} - \sum_{k=0}^{m-1}\frac{h^{2k}}{(2k)!}D^{2k}f(x_0) .$$

The upper generalized symmetric derivate of f at x_0 of order $2m$ is defined as

$$\bar{D}^{2m}f(x_0) = \limsup_{h\to 0}\theta_{2m}(f; x_0, h) .$$

Replacing 'lim sup' by 'lim inf' one gets the definition of $\underline{D}^{2m}f(x_0)$.

The function f is said to satisfy the property $\bar{\mathscr{S}}_{2m}$ at x_0, written as $f \in \bar{\mathscr{S}}_{2m}(x_0)$, if

$$\limsup_{h\to 0} h\theta_{2m}(f; x_0, h) \geqq 0 ,$$

and $f \in \underline{\mathscr{S}}_{2m}(x_0)$ if $-f \in \bar{\mathscr{S}}_{2m}(x_0)$. The function f is said to be smooth at x_0 of order $2m$ if

$$\lim_{h\to 0} h\theta_{2m}(f; x_0, h) = 0 .$$

Clearly smoothness of order $2m$ implies smoothness of order $2m-2$ and if f is smooth at x_0 of order $2m$ then $f \in \overline{\mathscr{S}}_{2m}(x_0) \cap \underline{\mathscr{S}}_{2m}(x_0)$. For symmetric derivatives of odd order, $\theta_{2m+1}(f; x_0, h)$, $\overline{D}^{2m+1}f(x_0)$, $\underline{D}^{2m+1}f(x_0)$, $\overline{\mathscr{S}}_{2m+1}(x_0)$, $\underline{\mathscr{S}}_{2m+1}(x_0)$ are defined analogously.

If $f_{(r)}(x_0)$ exists, $0 \leqq r \leqq n-1$, $\gamma_n(f; x_0, h)$ is defined as

$$\frac{h^n}{n!}\gamma_n(f; x_0, h) = f(x_0 + h) - \sum_{r=0}^{n-1} \frac{h^r}{r!} f_{(r)}(x_0) .$$

The upper generalized unsymmetric derivate of f at x_0 of order n is defined as

$$\overline{f}_{(n)}(x_0) = \limsup_{h\to 0} \gamma_n(f; x_0, h)$$

with a similar definition for $\underline{f}_{(n)}(x_0)$. By restricting h suitably one can define one-sided derivates which are denoted by $\overline{f}^{+}_{(n)}(x_0)$, etc. For convenience, the first order derivates $\overline{f}_{(1)}(x_0)$, $\overline{f}^{+}_{(1)}(x_0)$, etc., will be denoted simply by $\overline{f}(x_0)$, $\overline{f}^{+}(x_0)$, etc. The ordinary nth derivative of f at x_0 will be denoted by $f^{(n)}(x_0)$.

A function f is said to satisfy the property $\mathscr{R}$ in an interval I, written $f \in \mathscr{R}$ in I, if for every perfect set $P \subset I$, there is a portion of P in which f restricted to P is continuous (see [17]). A function f is said to satisfy the property $\mathscr{T}$ in (a, b), written $f \in \mathscr{T}$ in (a, b), if there exists a function F continuous in $[a, b]$ such that $F_{(n)} = f$ in (a, b) for some n. The class of all Darboux functions will be denoted by $\mathscr{D}$. From the properties of Darboux functions it follows that if $D^{2k}f \in \mathscr{D}$ and if g is continuous then $D^{2k}f + g \in \mathscr{D}$. This fact will be used in the sequel. For the definition of n-convex functions we refer to [8, 1].

We now come to the definition of major and minor functions. Let f be defined in (a, b) and let $a = a_1 < a_2 < \cdots < a_{2m} = b$. A function Q is said to be a P^{2m}-major function or simply a major function of f over $(a_i; 1 \leqq i \leqq 2m)$ if

(i) Q is continuous in $[a, b]$,

(ii) $D^{2m-2}Q$ exists and $D^{2k} \in \mathscr{R} \cap \mathscr{T}$ in (a, b), $0 \leqq k \leqq m-1$,

(iii) $Q(a_i) = 0$, $1 \leqq i \leqq 2m$,

(iv) $\underline{D}^{2m}Q \geqq f$ a.e. in (a, b),

(v) $\underline{D}^{2m}Q > -\infty$, except on an enumerable set $E \subset (a, b)$,

(vi) Q is smooth of order $2m$ on E.

The function q is a minor function of f if $-q$ is a major function of $-f$. The P^{2m+1}-major functions and P^{2m+1}-minor functions are defined similarly.

This definition of major and minor functions differs from that of James [8] in allowing certain exceptional sets in (iv) and (v). But this is standard and is also noted by James in his modified definition of the P^{2m}-integral [9]. Another difference is in condition (ii) where we are assuming $D^{2k}Q \in \mathscr{R} \cap \mathscr{T}$ instead of James' [8] requirement that Q has properties A_{2m} and B_{2m-2}. (The property $\mathscr{R}$ is weaker than A_{2m} by Lemma 3.2 of [8] and the property $\mathscr{T}$ is stronger than B_{2m-2} by Lemma 8.1 of [8] or by Theorem 2 of [13].) But this is necessary since the difference of two functions in $\mathscr{R} \cap \mathscr{T}$ is in $\mathscr{R} \cap \mathscr{T}$ which is not true with the property B_{2m-2}. We shall prove in the sequel that this is a proper definition of major and minor functions and the P^{2m}-integral defined by these major and minor functions is capable of handling trigonometric series.

2. Preliminary lemmas.

LEMMA 2.1. *If f is smooth of order $2m+1$, as well as of order $2m+2$, at x_0 then $f_{(2m)}(x_0)$ exists. If $f_{(n)}(x_0)$ exists then f is smooth of order $n+1$. More generally, if $\underline{f}^{+}_{(n+1)}(x_0)$, $\underline{f}^{-}_{(n+1)}(x_0)$, $\bar{f}^{+}_{(n+1)}(x_0)$, $\bar{f}^{-}_{(n+1)}(x_0)$ are all finite, then*

$$\limsup_{h\to 0} h\theta_{n+2}(f; x_0, h) \leqq \frac{n+2}{2}\{\bar{f}^{+}_{(n+1)}(x_0) - \underline{f}^{-}_{(n+1)}(x_0)\}$$

$$\liminf_{h\to 0} h\theta_{n+2}(f; x_0, h) \geqq \frac{n+2}{2}\{\underline{f}^{+}_{(n+1)}(x_0) - \bar{f}^{-}_{(n+1)}(x_0)\} .$$

Proof. The first part is clear. For the last part, since $f_{(n)}(x_0)$ exists, $D^r f(x_0)$ exists, $0 \leqq r \leqq n$, and

$$\frac{1}{2}\{\gamma_{n+1}(f; x_0, h) + \gamma_{n+1}(f; x_0, -h)\} = \theta_{n+1}(f; x_0, h) \tag{2.1}$$

$$\frac{1}{2}\{\gamma_{n+1}(f; x_0, h) - \gamma_{n+1}(f; x_0, -h)\} = \frac{h}{n+2}\theta_{n+2}(f; x_0, h) . \tag{2.2}$$

From (2.1)

$$\lim_{h\to 0} h\theta_{n+1}(f; x_0, h) = 0 ,$$

and from (2.2)

$$\frac{n+2}{2}\{\underline{f}^{+}_{(n+1)}(x_0) - \bar{f}^{-}_{(n+1)}(x_0)\} \leqq \liminf_{h\to 0} h\theta_{n+2}(f; x_0, h) .$$

The other relation follows similarly.

LEMMA 2.2. *If $G_{(n-1)}(x_0)$ and $D^nG(x_0)$ exist and if $G \in \bar{\mathscr{S}}_{n+2}(x_0)$ then the function $\omega_{n+1}(G; x_0, h)$ defined by*

$$(2.3)\quad \frac{h^{n+1}}{(n+1)!}\omega_{n+1}(G; x_0, h) = G(x_0 + h) - \sum_{r=0}^{n-1}\frac{h^r}{r!}G_{(r)}(x_0) - \frac{h^n}{n!}D^nG(x_0)$$

satisfies the relation

$$\limsup_{h\to 0+}\omega_{n+1}(G; x_0, h) \geqq \liminf_{h\to 0-}\omega_{n+1}(G; x_0, h) .$$

Proof. Since

$$\omega_{n+1}(G; x_0, h) - \omega_{n+1}(G; x_0, -h) = \frac{2h}{n+2}\theta_{n+2}(G; x_0, h) ,$$

and since $G \in \bar{\mathscr{S}}_{n+2}(x_0)$, the proof is immediate.

LEMMA 2.3. *If $f_{(n)}$ exists in (a, b) and $x_0 \in (a, b)$ then*

$$(2.4)\qquad \underline{(f_{(n)})}^+(x_0) \leqq \underline{f}^+_{(n+1)}(x_0),\ \bar{f}^+_{(n+1)}(x_0) \leqq \overline{(f_{(n)})}^+(x_0)\ , \quad \text{etc.}$$

$$(2.5)\qquad \underline{(f_{(n)})}(x_0) \leqq \underline{D}^{n+1}f(x_0),\ \bar{D}^{n+1}f(x_0) \leqq \overline{(f_{(n)})}(x_0) .$$

Proof. If $n = 0$ this is immediate. Suppose $n \geqq 1$. Then f is continuous in (a, b). Let $x_0 \in [\alpha, \beta] \subset (a, b)$. Then each $f_{(k)}$ is C_k-continuous in $[\alpha, \beta]$, $0 \leqq k \leqq n$, by Lemma 11.1 of [8]. From the definition of Cesaro derivative (see [4]) we have $C_nD^+f_{(n)}(x_0) = \bar{f}^+_{(n+1)}(x_0)$, where $C_nD^+f_{(n)}(x_0)$ is the right hand upper nth Cesaro derivate of $f_{(n)}$ at x_0. Since $C_0D^+f_{(n)}(x_0)$ is the first order derivate $\overline{(f_{(n)})}^+(x_0)$, (2.4) follows from Theorem 2.1 of [4]. Lastly, from (2.1), $\underline{D}^{n+1}f(x_0) \geqq f_{(n+1)}(x_0)$ and hence (2.5) follows from (2.4).

LEMMA 2.4. *Let g be continuous in $[a, b]$ and $\bar{D}^2g \geqq 0$ in (a, b), except on an enumerable set $E \subset (a, b)$ and let $g \in \bar{\mathscr{S}}_2(x)$ for $x \in E$. Then g is convex in $[a, b]$.*

This is proved in [19, I, p. 328], which sharpens a result of de La Vallée Poussin (see [16, Lemma 3]).

3. $2m$-convex functions. In this section and in § 4, the results are stated in a more general form than is necessary for P^{2m}-major and P^{2m+1}-major functions. Since every member in $\mathscr{T}$ possesses Darboux property [13], we have $\mathscr{T} \cap \mathscr{R} \subset \mathscr{D} \cap \mathscr{R}$ and hence these results are applicable in §§ 5 and 6.

THEOREM 3.1, 2m. *Suppose that*

(i) *f is continuous in $[a, b]$,*
(ii) *$D^{2m-2}f$ exists and $D^{2k}f \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m-1$,*
(iii) *$\bar{D}^{2m}f \geqq 0$ in (a, b), except on an enumerable set $E \subset (a, b)$,*
(iv) *$f \in \bar{\mathscr{S}}_{2m}(x)$ for $x \in E$.*
Then $D^{2m-2}f$ is convex in (a, b) and it is the continuous derivative $f^{(2m-2)}$ in (a, b).

The above theorem is true for $m = 1$ by Lemma 2.4. So, we assume that the theorem is true for $m = m_0$ i.e., Theorem 3.1, $2m_0$ is true and we prove that Theorem 3.1, 2 $(m_0 + 1)$ is also true and so the theorem will be proved to be true for all m by induction on m. We require the following auxiliary lemmas:

LEMMA 3.1, $2m_0$. *Suppose that*
(i) *G is continuous in $[a, b]$,*
(ii) *$D^{2m_0}G$ exists in (a, b) and is $\mathscr{L}$-integrable in $[a, b]$,*
(iii) *$D^{2k}G \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m_0 - 1$.*
Then $\Psi - G$ is a polynomial of degree at most $2m_0 - 1$ in $[a, b]$, where

$$\Psi(x) = \frac{1}{(2m_0-2)!}\int_a^x (x-t)^{2m_0-2}g(t)dt$$

$$g(x) = \int_a^x D^{2m_0}G(t)dt$$

and $G^{(2m_0-1)}$ exists and is continuous in (a, b).

Proof. As in [10, Theorem 18], one can construct a sequence of continuous functions $\{A_i\}$ which converges uniformly to g in $[a, b]$ as $i \to \infty$ and for all i

$$\underline{(A_i)}(x) > D^{2m_0}G(x)\,, \qquad x \in (a, b)\,.$$

For each i, define

$$U_i(x) = \frac{1}{(2m_0-2)!}\int_a^x (x-t)^{2m_0-2}A_i(t)dt\,, \qquad x \in [a, b]\,.$$

Then $\{U_i\}$ converges uniformly to Ψ in $[a, b]$ as $i \to \infty$. Since A_i is continuous, taking $(2m_0 - 1)$th derivative

$$U_i^{(2m_0-1)}(x) = A_i(x)\,, \qquad x \in (a, b)\,.$$

So, by (2.5) we have

$$\text{(3.1}\qquad \underline{(A_i)}(x) = \underline{(U_i^{(2m_0-1)})}(x) \leqq \underline{D}^{2m_0}U_i(x)\,, \qquad x \in (a, b)\,.$$

Since by construction $\underline{(A_i)}(x) > D^{2m_0}G(x)$ for $x \in (a, b)$, (3.2) $\quad \bar{D}^{2m_0}[U_i - G](x) > \bar{D}^{2m_0}U_i(x) - \underline{(A_i)}(x)\,, \qquad x \in (a, b)\,.$

Hence from (3.1) and (3.2)

$$\bar{D}^{2m_0}[U_i - G](x) > 0\,, \qquad x \in (a, b)\,.$$

Since $D^{2k}G \in \mathscr{D} \cap \mathscr{R}$ and $D^{2k}U_i$ is continuous in (a, b) for $0 \leqq k \leqq m_0 - 1$, $D^{2k}[U_i - G] \in \mathscr{D} \cap \mathscr{R}$ in (a, b) for $0 \leqq k \leqq m_0 - 1$. Hence by Theorem 3.1, $2\mathrm{m}_0$, $D^{2m_0-2}[U_i - G]$ is convex in (a, b) and so $U_i - G$ is $2m_0$-convex in (a, b) and by the continuity, $U_i - G$ is $2m_0$-convex in $[a, b]$. Since $U_i - G$ converges uniformly to $\Psi - G$ in $[a, b]$, $\Psi - G$ is $2m_0$-convex in $[a, b]$. It can be similarly shown that $\Psi - G$ is $2m_0$-concave in $[a, b]$. Hence $\Psi - G$ is a polynomial of degree at most $2m_0 - 1$. Since $\Psi^{(2m_0-1)}$ exists and is continuous, $G^{(2m_0-1)}$ also exists and is continuous in (a, b).

LEMMA 3.2, $2\mathrm{m}_0$. *Let G be continuous in $[a, b]$ and let $D^{2m_0}G$ exist in (a, b) and be $\mathscr{L}$-integrable in $[a, b]$. Let $G^{(2m_0-1)}$ exist and be continuous in (a, b). If $D^{2m_0}G$ attains a maximum at $x_0 \in (a, b)$ then*

$$\limsup_{h\to 0+} \omega_{2m_0+1}(G; x_0, h) \leqq 0 \leqq \liminf_{h\to 0-} \omega_{2m_0+1}(G; x_0, h)\,,$$

where ω is the function defined in (2.3) *with $n = 2m_0$.*

Proof. Let

$$J(x) = \int_a^x D^{2m_0}G(t)dt\,, \qquad x \in [a, b]\,.$$

Then by Lemma 3.1, $2\mathrm{m}_0$ $J - G^{(2m_0-1)}$ is constant. Since $G^{(2m_0-1)}$ is continuous in (a, b), by mean value property, for any h, $0 < h < b - x_0$, there is η, $0 < \eta < 1$, such that

$$\begin{aligned}\omega_{2m_0+1}(G; x_0, h) &= \frac{2}{(\eta h)^2}\{G^{(2m_0-1)}(x_0 + \eta h) - G^{(2m_0-1)}(x_0) - \eta h D^{2m_0}G(x_0)\} \\ &= \frac{2}{(\eta h)^2}\int_{x_0}^{x_0+\eta h}\{D^{2m_0}G(t) - D^{2m_0}G(x_0)\}dt\,.\end{aligned}$$

Therefore, since $D^{2m_0}G$ is maximum at x_0,

$$\limsup_{h\to 0+} \omega_{2m_0+1}(G; x_0, h) \leqq 0\,.$$

The other part follows similarly.

LEMMA 3.3, $2\mathrm{m}_0$. *Suppose that*

(i) *F is continuous in $[a, b]$,*

(ii) *$D^{2m_0-2}F$ exists and $D^{2k}F \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m_0 - 1$,*
(iii) *$\bar{D}^{2m_0}F \geqq 0$ in (a, b), except on an enumerable set $E \subset (a, b)$,*
(iv) *$F \in \overline{\mathscr{S}}_{2m_0}(x)$ for $x \in E$.*

Then

$$\theta_{2m_0}(F; x, h) \geqq 0 \ , \quad \textit{for all } x, h, \ a < x - h < x + h < b \ .$$

LEMMA 3.4, $2m_0$. *Suppose that*
(i) *G is continuous in $[a, b]$,*
(ii) *$D^{2m_0}G$ exists and $D^{2k}G \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m_0$,*
(iii) *$D^{2m_0}G$ attains a maximum at $x_0 \in (a, b)$.*

Then

$$\bar{D}^{2m_0+2}G(x_0) \leqq 0 \ .$$

The proof of Lemma 3.3, $2m_0$ is similar to that of Lemma 4.1, $2m_0$ of [8]. Lemma 3.4, $2m_0$ can be proved by using Lemma 3.3, $2m_0$ in the same manner as in Lemma 4.2, $2m_0$ of [8].

LEMMA 3.5, $2m_0$. *Suppose that*
(i) *f is continuous in $[a, b]$,*
(ii) *$D^{2m_0}f$ exists and $D^{2k}f \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m_0$,*
(iii) *$\bar{D}^{2m_0+2}f \geqq 0$ in (a, b), except on an enumerable set $E \subset (a, b)$,*
(iv) *$f \in \overline{\mathscr{S}}_{2m_0+2}(x)$ for $x \in E$,*
(v) *$D^{2m_0}f$ is upper semicontinuous in (a, b) and $\mathscr{L}$-integrable in $[a, b]$.*

Then $D^{2m_0}f$ is convex in (a, b).

Proof. We first consider the special case when the inequality in (iii) is strict inequality. Suppose that $D^{2m_0}f$ is not convex in (a, b). Then there is a subinterval $[\alpha, \beta] \subset (a, b)$ such that

$$\begin{aligned}\rho(x) &= D^{2m_0}f(x) - \frac{1}{\beta - \alpha}\{(\beta - x)D^{2m_0}f(\alpha) + (x - \alpha)D^{2m_0}f(\beta)\} \\ &= D^{2m_0}f(x) - px - q\end{aligned}$$

takes positive values somewhere in (α, β). Since ρ is upper semicontinuous in $[\alpha, \beta]$ and $\rho(\alpha) = \rho(\beta) = 0$, ρ attains maximum in (α, β). So, if μ is sufficiently near to p then the function $D^{2m_0}G$, where

$$G(x) = f(x) - \mu\frac{x^{2m_0+1}}{(2m_0 + 1)!} - q\frac{x^{2m_0}}{(2m_0)!} \ ,$$

also attains its maximum in (α, β), say, at x_μ. Hence by Lemma 3.4, $2m_0$

$$\bar{D}^{2m_0+2}G(x_\mu) = \bar{D}^{2m_0+2}f(x_\mu) \leqq 0 .$$

Hence $x_\mu \in E$. Now G satisfies the hypotheses of Lemma 3.1, $2m_0$ and hence $G^{(2m_0-1)}$ exists and is continuous in (a, b). Also since $f \in \bar{\mathscr{S}}_{2m_0+2}(x)$ for $x \in E$, $G \in \bar{\mathscr{S}}_{2m_0+2}(x_\mu)$. Hence by Lemma 2.2

$$\limsup_{h\to 0+} \omega_{2m_0+1}(G; x_\mu, h) \geqq \liminf_{h\to 0-} \omega_{2m_0+1}(G; x_\mu, h)$$

where ω is the function as defined in (2.3) with $n = 2m_0$, $x_0 = x_\mu$. But by Lemma 3.2, $2m_0$, since $D^{2m_0}G$ is maximum at x_μ,

$$\limsup_{h\to 0+} \omega_{2m_0+1}(G; x_\mu, h) \leqq 0 \leqq \liminf_{h\to 0-} \omega_{2m_0+1}(G; x_\mu, h)$$

and hence

$$\liminf_{h\to 0-} \omega_{2m_0+1}(G; x_\mu, h) = 0$$

i.e.,

$$\liminf_{h\to 0-} \omega_{2m_0+1}(f; x_\mu, h) = \mu .$$

Thus for each μ sufficiently near to p there exists $x_\mu \in E$ and for different μ the points x_μ are also different. This contradicts the fact that E is enumerable.

To complete the proof, consider, for arbitrary $\varepsilon > 0$, the function g_ε where

$$g_\varepsilon(x) = f(x) + \varepsilon \cdot \frac{x^{2m_0+2}}{(2m_0 + 2)!} .$$

Then by the above special case, $D^{2m_0}g_\varepsilon$ is convex in (a, b) and since ε is arbitrary, $D^{2m_0}f$ is convex in (a, b), completing the proof.

Proof of Theorem 3.1, $2(m_0+1)$. To prove the theorem we remark that under the hypotheses, if $D^{2m_0}f$ is continuous in an interval $(\alpha, \beta) \subset (a, b)$, then by Lemma 3.1, $2m_0$, $f^{(2m_0-1)}$ exists and is continuous in (α, β) and so by Lemma 7 of [18], $D^{2m_0}f$ is the continuous ordinary derivative $f^{(2m_0)}$ in (α, β). Hence applying the mean value property it can be shown that $\bar{D}^2(f^{(2m_0)}) \geqq \bar{D}^{2m_0+2}f$ and that $f^{(2m_0)} \in \bar{\mathscr{S}}_2(x)$ if $f \in \bar{\mathscr{S}}_{2m_0+2}(x)$ for points in (α, β) and so by Lemma 2.4, $f^{(2m_0)}$ is convex in (α, β).

Let U be the set of all points x in (a, b) such that there is a neighborhood of x in which $D^{2m_0}f$ is continuous. Then U is open. Let (α, β) be any component interval of U. Then $D^{2m_0}f$ is continuous in (α, β) and so by the above remark $D^{2m_0}f$ is convex in (α, β). Hence $\lim_{x\to\alpha+} D^{2m_0}f(x)$ and $\lim_{x\to\beta-} D^{2m_0}f(x)$ exist and by the property $\mathscr{D}$, $D^{2m_0}f$ is continuous in $[\alpha, \beta] \cap (a, b)$. Let $P = (a, b) - U$. Then P is closed in (a, b). Since $D^{2m_0}f$ is continuous in the closure (relative to (a, b)) of each component interval of U, P is perfect in (a, b). If possible, suppose that $P \neq 0$. Then there is $[c, d] \subset (a, b)$ such that $[c, d] \cap P$ is a nonvoid perfect set. Since $D^{2k}f \in \mathscr{R}$ in (a, b), there is a portion of $[c, d] \cap P$, say, $H = [a_0, b_0] \cap P$ on which $D^{2k}f/H$ is continuous for each k, $0 \leqq k \leqq m_0$. It can be shown, as in Theorem 4.1, $2(m_0 + 1)$ of [8] that $D^{2m_0}f$ is upper semicontinuous in $[a_0, b_0]$. Hence there is M such that $D^{2m_0}f(x) \leqq M$ for $x \in [a_0, b_0]$. Since the theorem is true for $m = m_0$, the function $F(x) = Mx^2/2 - D^{2m_0-2}f(x)$ is convex in (a_0, b_0). Choose a_1, b_1, such that $a_0 < a_1 < b_1 < b_0$ and $P \cap (a_1, b_1) \neq 0$. Then by Lemma 3.16 of [19, I, p. 328], D^2F exists almost everywhere in (a_0, b_0) and is $\mathscr{L}$-integrable in $[a_1, b_1]$. Since F is continuous, $D^2F = M - D^{2m_0}f$ holds whenever D^2F exists and hence $D^{2m_0}f$ is $\mathscr{L}$-integrable in $[a_1, b_1]$. So, by Lemma 3.5, $2m_0$, $D^{2m_0}f$ is convex in (a_1, b_1). Hence $D^{2m_0}f$ is continuous in (a_1, b_1). This contradicts the fact that $(a_1, b_1) \cap P \neq 0$. Hence $P = 0$ and so $D^{2m_0}f$ is continuous in (a, b). Hence by our earlier remark $D^{2m_0}f$ is convex in (a, b). The rest follows from Lemma 3.1, $2m_0$ and Lemma 7 of [18]. This completes the proof of the theorem for $m = m_0 + 1$.

Thus the theorem is true for all m and so henceforth we shall omit $2m$ in refering to this theorem. The usual extension of the above theorem is the following

THEOREM 3.2. *Suppose that*

(i) *f is continuous in* $[a, b]$,

(ii) *$D^{2m-2}f$ exists and $D^{2k}f \in \mathscr{D} \cap \mathscr{R}$ in* (a, b), $0 \leqq k \leqq m - 1$,

(iii) $\bar{D}^{2m}f \geqq 0$ a.e. *in* (a, b),

(iv) *$\bar{D}^{2m}f > -\infty$, except on an enumerable set $E \subset (a, b)$,*

(v) *$f \in \bar{\mathscr{S}}_{2m}(x)$, for $x \in E$.*

Then $D^{2m-2}f$ is convex in (a, b) and $D^{2m-2}f$ is the continuous derivative $f^{(2m-2)}$ in (a, b).

This can be proved from Theorem 3.1 by using standard argument used to prove Theorem 1.1 of [5] or Theorem 16 of [1] and so we omit it.

REMARK 3.1. The property $D^{2k}f \in \mathscr{D}$ for $0 \leqq k \leqq m - 1$, in the above theorem plays an important role. For, consider the function f where

$$f(x) = \begin{cases} x^2, & x \geqq 0 \\ -x^2, & x < 0 . \end{cases}$$

Then D^2f exists everywhere but $D^2f \notin \mathscr{D}$. Also f satisfies all the other conditions of the above theorem and $D^4f = 0$ everywhere; but D^2f is neither convex nor concave in any interval including 0.

REMARK 3.2. The above example shows that if $\underline{D}^{2m}f$ replaces $\bar{D}^{2m}f$ in the hypotheses (iii) and (iv) of the above theorem and if in (v) smoothness of f of order $2m$ is assumed everywhere, then even under this stronger conditions the theorem is false without the property $D^{2k}f \in \mathscr{D}$.

4. **$(2m+1)$-convex functions.** Now it is natural to ask whether the analogous results hold for odd order derivatives. In [8], it is indicated that the proof of Theorem 4.1, 3 of [8] was similar to that of a theorem of Saks [14]. But Saks used the lower derivate $\underline{D}^3f$ and not $\bar{D}^3f$ and so the induction on m in [8] ensures the validity of Theorem 4.1, $2m+1$ of [8], provided $\bar{D}^{2m+1}f$ is replaced by $\underline{D}^{2m+1}f$ in its hypotheses. But if in the hypotheses of Theorem 4.1, $2m+1$ of [8], $\bar{D}^{2m+1}f$ is replaced by $\underline{D}^{2m+1}f$ then this new theorem is only a consequence of Theorem 4.1, $2(m+1)$ of [8] for the integrated function. The proof of Theorem 4.1, $2m+1$ of [8] is thus incomplete. We complete the proof in the following more general theorem.

THEOREM 4.1. *Suppose that*

(i) *f is continuous in $[a, b]$,*

(ii) *$D^{2m-1}f$ exists and $D^{2k+1}f \in \mathscr{D} \cap \mathscr{R}$ in (a, b), $0 \leqq k \leqq m-1$,*

(iii) *$\bar{D}^{2m+1}f \geqq 0$ in (a, b), except on an enumerable set $E \subset (a, b)$,*

(iv) *$f \in \bar{\mathscr{S}}_{2m+1}(x)$ for $x \in E$.*

Then $D^{2m-1}f$ is convex in (a, b) and it is the continuous derivative $f^{(2m-1)}$ in (a, b).

The proof is similar to that of Theorem 3.1. It is necessary to prove this theorem for $m=1$ and to do this, Lemmas 4.1, 1, 4.2, 1, 4.4, 1, 4.5, 1, which are analogous to Lemmas 3.1, $2m_0$, 3.2, $2m_0$, 3.4, $2m_0$, 3.5, $2m_0$, will be needed. The proofs of Lemmas 4.2, 1 and 4.5, 1 are similar to those of Lemmas 3.2, $2m_0$ and 3.5, $2m_0$ respectively. In proving Lemma 4.1, 1 one is to appeal to a result of [12] instead of assuming Theorem 3.1, $2m_0$ as it was done in Lemma 3.1, $2m_0$ and in proving Lemma 4.4, 1 one is to notice that since $D^1G \in \mathscr{D}$, by the same result of [12], D^1G has mean value property and hence for any h there is ξ, $x_0 - h < \xi < x_0 + h$, such that

$$h^2\theta_3(G; x_0, h) = 3!\,\{D^1G(\xi) - D^1G(x_0)\} \leqq 0$$

giving $\bar{D}^3G(x_0) \leqq 0$. The proof of Theorem 4.1 for $m=1$ will now follow the same line of argument as in Theorem 3.1, $2(m_0+1)$. The $\mathscr{L}$-integrability of D^1f will follow from the fact that $F(x) = Mx - f(x)$ is nondecreasing in $[a_0, b_0]$, [12] and $M - D^1f$ is the derivative of

F where it exists. Proving the above theorem for $m = 1$ and supposing it to be true for $m = m_0$, all the lemmas beginning 4.1, $2m_0 + 1$ through 4.5, $2m_0 + 1$ can be proved and the proof of the theorem for $m = m_0 + 1$ can be completed. We remark that an analogue of Theorem 3.2 is also true in this case.

5. **The P^{2m}-integral.** We now come to the definition of the integral. We must show that the definition of major and minor functions, as introduced earlier, actually helps to obtain a proper definition of the integral. For, because of the presence of the exceptional set E in condition (v) and (vi) of the definition of major function we cannot apply directly Theorem 3.2 to prove that $Q - q$ is a $2m$-convex function for arbitrary major and minor functions Q and q respectively. (As the definition of the P^{2m}-integral in [9] and that of the P^2-integral in [7] are also affected by the exceptional sets S and E_0 respectively, (see [9] and [7]) they would also need this clarification; but the definition of the P^2-integral in [6] is not affected since the smoothness of major and minor functions is assumed everywhere). We shall follow the method adopted in [15].

LEMMA 5.1. *Given $\varepsilon_0 > 0$ and $x_0 \in (a, b)$ there is a major function Q for the function $t(x) \equiv 0$ such that*

(i) $Q^{(2m-2)}$ *is continuous in* $[a, b]$,

(ii) $\underline{D}^2 Q^{(2m-2)} \geqq 0$ *in* (a, b),

(iii) $\lim_{h\to 0} h\theta_2(Q^{(2m-2)}; x_0, h) > 0$, $\lim_{h\to 0} h\theta_2(Q^{(2m-2)}; x, h) = 0$, *for* $x \neq x_0$,

(iv) $|Q^{(2m-2)}| \leqq \varepsilon_0$ *in* (a, b),

(v) $|h\theta_2(Q^{(2m-2)}; x, h)| \leqq \varepsilon_0$, *for* $x \neq x_0$, *and* $x, x \pm h \in (a, b)$.

Proof. Let g be the function such that

$$g(x_0) = 0\ , \quad g(a) = \frac{1}{2}\min\cdot\left[\varepsilon_0(x_0 - a), \frac{\varepsilon_0}{(2m)!}\right],$$
$$g(b) = \frac{1}{2}\min\cdot\left[\varepsilon_0(b - x_0), \frac{\varepsilon_0}{(2m)!}\right],$$

and g is linear and continuous in each of the interval $[a, x_0]$ and $[x_0, b]$ and let G be the $(2m-2)$th indefinite integral of g in $[a, b]$. Then the function Q defined by

$$Q(x) = G(x) - \sum_{i=1}^{2m} \lambda(x; a_i) G(a_i)$$

satisfies the requirements, where

$$\lambda(x; a_i) = \prod_{\substack{j=1 \\ j\neq i}}^{2m} \frac{x - a_j}{a_i - a_j}\ , \quad a = a_1 < a_2 < \cdots < a_{2m} = b\ . \tag{5.1}$$

LEMMA 5.2. *If Q is a major function of f and $\varepsilon > 0$, then there is a major function Q_ε such that*

$$|D^{2m-2}Q_\varepsilon - D^{2m-2}Q| \leqq \varepsilon\ , \quad \underline{D}^{2m}Q_\varepsilon > -\infty\ , \quad in\ (a, b)\ .$$

Proof. Let $x_1, x_2, \cdots, x_k, \cdots$ be an enumeration of the exceptional set $E \subset (a, b)$, where $\underline{D}^{2m}Q = -\infty$ holds. For each positive integer k, let F_k be the major function obtained from Lemma 5.1 with ε_0 and x_0 replaced by $\varepsilon/2^k$ and x_k respectively. Set

$$\Psi(x) = \sum_{k=1}^{\infty} F_k^{(2m-2)}(x)\ , \qquad F(x) = \sum_{k=1}^{\infty} F_k(x)\ .$$

The first series being uniformly and absolutely convergent, Ψ is continuous and $\Psi = F^{(2m-2)}$. By the mean value property there is η, $0 < \eta < 1$, such that

$$\theta_{2m}(F; x, h) = \theta_2(\Psi; x, \eta h) = \sum_{k=1}^{\infty} \theta_2(F_k^{(2m-2)}; x, \eta h)$$

and since by (i), (ii) of Lemma 5.1 and by Theorem 3.1, each $F_k^{(2m-2)}$ is convex in (a, b), $\underline{D}^{2m}F \geqq 0$ in (a, b). Also, for $x_i \in E$, the series $\sum_{k=i+1}^{\infty} h\theta_2(F_k^{(2m-2)}; x_i, h)$ is uniformly and absolutely convergent with respect to h and hence

$$\begin{aligned}\lim_{h\to 0} h\theta_{2m}(F; x_i, h) &= \lim_{h\to 0} h\theta_2(\Psi; x_i, h)\\ &= \lim_{h\to 0} \sum_{k=1}^{\infty} h\theta_2(F_k^{(2m-2)}; x_i, h)\\ &= \lim_{h\to 0} h\theta_2(F_i^{(2m-2)}; x_i, h)\\ &> 0\ .\end{aligned}$$

Now set

$$Q_\varepsilon(x) = Q(x) + F(x)\ .$$

Then if $x_i \in E$,

$$\lim_{h\to 0} h\theta_{2m}(Q_\varepsilon; x_i, h) = \lim_{h\to 0} h\theta_{2m}(F; x_i, h) > 0$$

and hence $\underline{D}^{2m}Q_\varepsilon(x_i) = \infty$. Clearly Q_ε is a major function of f and by construction $|D^{2m-2}Q_\varepsilon - D^{2m-2}Q| \leqq \varepsilon$.

LEMMA 5.3. *If Q and q are any major and minor functions then $Q - q$ is $2m$-convex.*

Proof. By Lemma 5.2, for each positive integer n there is a major function Q_n and a minor function q_n such that

(5.2) $$|D^{2m-2}Q_n - D^{2m-2}Q| \leqq \frac{1}{n}, \underline{D}^{2m}Q_n > -\infty\,, \quad \text{in } (a, b)$$

and a similar relation for q_n holds. Hence $\underline{D}^{2m}[Q_n - q_n] \geqq 0$ a.e. in (a, b) and $\underline{D}^{2m}[Q_n - q_n] > -\infty$ in (a, b). Since $D^{2k}Q_n \in \mathscr{T}$, and $D^{2k}q_n \in \mathscr{T}$, we have $D^{2k}[Q_n - q_n] \in \mathscr{T}$ and hence $D^{2k}[Q_n - q_n] \in \mathscr{D}$, for each k, $0 \leqq k \leqq m-1$, [13]. So, by Theorem 3.2 $D^{2m-2}[Q_n - q_n]$ is convex in (a, b) and hence by (5.2) and a relation for q_n, $D^{2m-2}[Q - q]$ is convex in (a, b) and so the result follows.

Lemma 5.3 gives the analogue of Lemma 5.1 of [8]. Once this lemma is proved all the subsequent results of [8] can be deduced with this definition of major and minor functions. The definition of P^{2m}-integral thus obtained remains valid and all the results of [8] except Theorem 5.4 of [8] are true. We state Theorem 5.4 of [8] in our setting whose proof is similar to that in [8].

THEOREM 5.1. *If G is such that*

(i) *G is continuous in $[a, b]$,*

(ii) *$D^{2m-2}G$ exists and $D^{2k}G \in \mathscr{R} \cap \mathscr{T}$ in (a, b), $0 \leqq k \leqq m-1$,*

(iii) *$D^{2m}G$ exists* a.e. *in (a, b),*

(iv) *$-\infty < \underline{D}^{2m}G \leqq \bar{D}^{2m}G < \infty$, except on an enumerable set $E \subset (a, b)$,*

(v) *G is smooth of order $2m$ on E,*

then $D^{2m}G$ is P^{2m}-integrable over $(a_i; x)$, where $a \leqq a_1 < a_2 < \cdots < a_{2m} \leqq b$, and if $a_r \leqq x < a_{r+1}$, then

$$(-1)^r \int_{(a_i)}^{x} f(t)d_{2m}t = G(x) - \sum_{i=1}^{2m} \lambda(x; a_i)G(a_i)$$

where λ is the function defined in (5.1).

6. The P^{2m+1}-integral. The definition of P^{2m+1}-integral can be obtained from the P^{2m+1}-major and minor functions in the same manner as in the case of P^{2m}-integral. The P^1-integral i.e., P^{2m+1}-integral for $m = 0$ is not defined in [8]. Theorem 3 of [12] shows that the definition of P^1-integral is also valid and so the definition of symmetric P^n-integral is valid for all $n \geqq 1$.

7. The unsymmetric P^n-integral. The unsymmetric P^n-integral as defined in [8] is not affected by Lemma 5.1 of [8]. We state here the conditions to be satisfied by an unsymmetric P^n-major function Q of the function f in our improved setting:

(i) Q is continuous in $[a, b]$,

(ii) $Q_{(n-1)}$ exists in (a, b),

(iii) $Q(a_i) = 0$, $1 \leqq i \leqq n$,

(iv) $\underline{Q}_{(n)} \geqq f$ a.e. in (a, b),

(v) $\underline{Q}_{(n)} > -\infty$, except on an enumerable set $E \subset (a, b)$.
It is easy to verify that for any major and minor function, Q and q, the difference $Q - q$ is n-convex. The definition of the unsymmetric P^n-integral now follows that of the symmetric P^n-integral. For different approach we refer to [2, 3].

8. Application to trigonometric series. Now we shall show that the results of [9] remain true with this definition of the P^{2m}-integral. For the notations $A_n^k(x)$, $B_n^k(x)$ and the upper and the lower (C, k) sums $S^k(x)$ and $s^k(x)$, which we shall use here, we refer to [9] (see also [19, I, pp. 74–77]).

THEOREM 8.1. (Cf. Theorem 6.2 of [9].) *Suppose that the series*

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \tag{8.1}$$

is summable (C, k) *almost everywhere to a finite function* f *on* $[0, 2\pi]$ *and let*

$$-\infty < s^k(x) \leqq S^k(x) < \infty\ , \tag{8.2}$$

except on an enumerable set in $[0, 2\pi]$. *If for* $x \in [0, 2\pi]$

$$A_n^{k-1}(x) = o(n^k)\ , \qquad B_n^{k-1}(x) = o(n^k)\ , \tag{8.3}$$

as $n \to \infty$, *then* $f(x)$, $f(x)\cos rx$, $f(x)\sin rx$, *are* P^{k+2}*-integrable over* $(\alpha_i; x)$ *and the coefficients of* (8.1) *are given by*

$$a_r = \frac{\delta_k}{2^{k+1}\pi^{k+2}}\int_{(\alpha_i)}^{0} f(x)\cos rx\, d_{k+2}x \tag{8.4}$$

$$b_r = \frac{\delta_k}{2^{k+1}\pi^{k+2}}\int_{(\alpha_i)}^{0} f(x)\sin rx\, d_{k+2}x \tag{8.5}$$

where

$$\delta_k = \frac{(k+2)!}{\left\{\left(\frac{k+2}{2}\right)!\right\}^2} \qquad \textit{if } k \textit{ is even}\ ,$$

$$= \frac{(k+2)!}{\left(\frac{k+1}{2}\right)!\left(\frac{k+3}{2}\right)!} \qquad \textit{if } k \textit{ is odd}\ .$$

Proof. Since (8.1) is summable (C, k), the series obtained by integrating (8.1) term by term $k+2$ times converges uniformly to a continuous function F and

$$a_n = o(n^k)\ , \qquad b_n = o(n^k)\ ,$$

as $n \to \infty$, (see [18]) and hence F is smooth of order $k+2$ (see [9, Theorem 3.1]). Since the once-integrated series of (8.1) is also summable $(C, k-1)$ a.e. in $[0, 2\pi]$ (see [11]), F is smooth of order $k+1$; hence by Lemma 2.1, $F_{(k)}$ exists and by Lemma 6 of [18], $F_{(i)} \in \mathscr{R}$ in $(0, 2\pi)$ for $0 \leqq i \leqq k$. By [18, Theorem B] we get from (8.2)

$$-\infty < \underline{D}^{k+2}F(x) \leqq \bar{D}^{k+2}F(x) < \infty$$

except on an enumerable set and $D^{k+2}F(x) = f(x)$ a.e. in $(0, 2\pi)$. So, by Theorem 5.1, f is P^{k+2}-integrable over $(\alpha_i; x)$. The proofs that $f(x)\cos rx$ and $f(x)\sin rx$ are also P^{k+2}-integrable and that the coefficients of (8.1) are given by (8.4) and (8.5) are similar to those given in [9, Theorem 4.2 and its corollary].

REFERENCES

1. P. S. Bullen, *A criterion for n-convexity*, Pacific J. Math., **36** (1971), 81-98.
2. ———, *The P^n-integral*, J. Australian Math. Soc., (2), **14** (1972), 219-236.
3. P. S. Bullen and S. N. Mukhopadhyay, *Peano derivatives and general integrals*, Pacific J. Math., **47** (1973), 43-58.
4. J. C. Burkill, *The Cesaro-Perron scale of integration*, Proc. London Math. Soc., (2), **39** (1935), 541-552.
5. ———, *Integrals and trigonometric series*, Proc. London Math. Soc., (3), **1** (1951), 46-57.
6. R. D. James and W. H. Gage, *A generalized integral*, Trans. Royal Soc. Canada, Sec. III, **40** (1946), 25-35.
7. R. D. James, *A generalized integral II*, Canad. J. Math., **2** (1950), 297-306.
8. ———, *Generalized* nth *primitives*, Trans. Amer. Math. Soc., **76** (1954), 149-176.
9. ———, *Summable trigonometric series*, Pacific J. Math., **6** (1956), 99-110.
10. R. L. Jeffery, *Trigonometric Series*, Toronto, 1956.
11. J. Marcinkeiwicz and A. Zygmund, *On the differentiability of functions and summability of trigonometric series*, Fundamenta Mathematicae, **26** (1936), 1-43.
12. S. N. Mukhopadhyay, *On Schwarz differentiability IV*, Acta Mathematica Academiae Scientiarum Hungaricae, **17** (1966), 129-136.
13. H. W. Oliver, *The exact Peano derivative*, Trans. Amer. Math. Soc., **76** (1954), 444-456.
14. S. Saks, *On generalized derivatives*, J. London Math. Soc., **7** (1933), 247-251.
15. S. J. Taylor, *An integral of Perron's type defined with the help of trigonometric series*, Quarterly J. Math., (Oxford), (2), **6** (1955), 255-274.
16. S. Verblunsky, *On the theory of trigonometric series II*, Proc. London Math. Soc., **34** (1932), 457-491.
17. ———, *On the theory of trigonometric series V*, Fundamenta Mathematicae, **21** (1933), 168-210.
18. F. Wolf, *Summable trigonometric series: an extension of uniqueness theorem*, Proc. London Math. Soc., (2), **45** (1939), 328-356.
19. A. Zygmund, *Trigonometric series*, Cambridge, 1968.

Received January 7, 1974 and in revised form September 25, 1974. The ideas of this work were originated when the author was in a visiting position at The University of British Columbia supported by the National Research Council of Canada Grant No. 67-2986. The author wishes to thank Professor P. S. Bullen for various helpful discussions.

THE UNIVERSITY OF BURDWAN, BURDWAN, WEST BENGAL, INDIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

ON (J, M, m)-EXTENSIONS OF BOOLEAN ALGEBRAS

DWIGHT W. READ

The class $\mathscr{K}$ of all (J, M, m)-extensions of a Boolean algebra $\mathscr{A}$ can be partially ordered and always contains a maximum and a minimal element, with respect to this partial ordering. However, it need not contain a smallest element. Should $\mathscr{K}$ contain a smallest element, then $\mathscr{K}$ has the structure of a complete lattice. Necessary and sufficient conditions under which $\mathscr{K}$ does contain a smallest element are derived. A Boolean algebra $\mathscr{A}$ is constructed for each cardinal m such that the class of all m-extensions of $\mathscr{A}$ does not contain a smallest element. One implication of this construction is that if a Boolean algebra $\mathscr{A}$ is the Boolean product of a least countably many Boolean algebras, each of which has more than one m-extension, then the class of all m-extensions of $\mathscr{A}$ does not contain a smallest element. The construction also has as implication that neither the class of all $(m, 0)$-products nor the class of all (m, n)-products of an indexed set $\{\mathscr{A}_t\}_{t \in T}$ of Boolean algebras need contain a smallest element.

1. Sikorski [2] has investigated the question of imbedding a given Boolean algebra $\mathscr{A}$ into a complete or m-complete Boolean algebra $\mathscr{B}$ and has shown that in the case where the imbedding map is not a complete isomorphism, the imbedding need not be unique up to isomorphism. He further has shown that if $\mathscr{K}$ is the class of all (J, M, m)-extensions of a Boolean algebra $\mathscr{A}$, then $\mathscr{K}$ has a naturally defined partial ordering on it and always contains a maximum and a minimal element. He has left as an open question whether it always contains a smallest element. La Grange [1] has given an example which implies that $\mathscr{K}$ need not always contain a smallest element. However, the question of when does $\mathscr{K}$ in fact contain a smallest element is of interest as it turns out that should $\mathscr{K}$ contain a smallest element, it has the structure of a complete lattice.

In § 2, necessary and sufficient conditions are given for $\mathscr{K}$ to contain a smallest element. In addition, the principle behind La Grange's example is generalized in Proposition 2.10 to show that if $\mathscr{A}$ is not m-representable then the class $\mathscr{K}$ of all (J, M, m')-extension of $\mathscr{A}$, where $\bar{\bar{J}}, \bar{\bar{M}} < \sigma$ and $m' > M$, will not contain a smallest element.

Since the proof of this result requires that J and M have cardinality $\leqq \sigma$, it is of interest to ask if the class of all m-extensions contain a smallest element in general, and the answer is no.

In § 3, a Boolean algebra $\mathscr{A}$ is constructed for each cardinal m such that the class $\mathscr{K}$ of all m-extensions of $\mathscr{A}$ does not contain a smallest element. The construction has as implication (Theorems 3.1 and 3.2; Corollary 3.1) that for each algebra in a rather broad group of Boolean algebras, the class of all m-extensions will not contain a smallest element. In particular, this group includes all Boolean algebras which are the Boolean product of at least countably many Boolean algebras each of which has more than one m-extension.

Finally, in the last section, Sikorski's result that there is an equivalence between the class $\mathscr{P}$ of all $(m, 0)$-products of an indexed set $\{\mathscr{A}_t\}_{t\in T}$ of Boolean algebras and the class of all (J, M, m)-extensions of the Boolean product $\mathscr{A}_0$ of $\{\mathscr{A}_t\}_{t\in T}$, for suitably defined J and M, is generalized to show there is an equivalence between the class $\mathscr{P}_n$ of all (m, n)-products of $\{\mathscr{A}_t\}_{t\in T}$ and all (J, M, m)-extensions of $\hat{\mathscr{F}}$, where $\hat{\mathscr{F}}$ is the field of sets generated by a certain set $\mathscr{S}$, for suitably defined J and M. Then the above results imply that neither $\mathscr{P}$ nor $\mathscr{P}_n$ need contain a smallest element.

The notation throughout follows that of Sikorski [2].

2. Let n be the cardinality of a set of generators for the Boolean algebra $\mathscr{A}$, let $\mathscr{A}_{m,n}$ be a free Boolean m-algebra with a set of n free m-generators, let $\mathscr{A}_{0,n}$ be the free Boolean algebra generated by this set of n free m-generators and let g be a homomorphism from $\mathscr{A}_{0,n}$ to $\mathscr{A}$. Let Δ_0 be the kernel of this homomorphism and let I be the set of all m-ideals Δ in $\mathscr{A}_{m,n}$ such that:

a. $\Delta \cap \mathscr{A}_{0,n} = \Delta_0$;

b. Δ contains all the elements

$$A_0 - \bigcup_{A\in S_1} A\,, \qquad \bigcup_{A\in S_1} A - A_0\,,$$
$$A_0 - \bigcap_{A\in S_2} A\,, \qquad \bigcap_{A\in S_2} A - A_0\,,$$

where $A_0 \in \mathscr{A}_{0,n}$ and $\mathscr{S}_1, \mathscr{S}_2$ are any subsets of $\mathscr{A}_{0,n}$ of cardinality $\leqq m$ such that:

$$g(\mathscr{S}_1) \in J\,, \qquad g(A_0) = \bigcup_{A\in\mathscr{S}_1} g(A)$$
$$g(\mathscr{S}_2) \in M\,, \qquad g(A_0) = \bigcap_{A\in\mathscr{S}} g(A)\,.$$

For each $\Delta \in I$ let

$$\mathscr{A}_\Delta = \mathscr{A}_{m,n}/\Delta$$

and

$$g_\Delta([A]_\Delta) = g(\Delta)\,, \quad \text{for all} \quad A \in \mathscr{A}_{0,n}\,.$$

Set $i_\Delta = g_\Delta^{-1}$. We need the following results due to Sikorski.

PROPOSITION 2.1. *The ordered pair* $\{i_\Delta, \mathscr{A}_\Delta\}$ *is a* (J, M, m)-*extension of the Boolean algebra* $\mathscr{A}$ *and if* $\{i, \mathscr{B}\}$ *is a* (J, M, m)-*extension of* $\mathscr{A}$ *there is a* $\Delta \in I$ *such that* $\{i_\Delta, \mathscr{A}_\Delta\}$ *is isomorphic to* $\{i, \mathscr{B}\}$. *Further, if* $\Delta, \Delta' \in I$ *then*

$$\{i_\Delta, \mathscr{A}_\Delta\} \leqq \{i_{\Delta'}, \mathscr{A}_{\Delta'}\} \quad \textit{if, and only if,} \quad \Delta \supseteqq \Delta' .$$

LEMMA 2.1. *If* S *is a set of elements in* $\mathscr{K}$ *then the least upper bound* (lub) *of* S *exists in* $\mathscr{K}$.

Now let $\mathscr{K}(J, M, m)$ denote the class of all (J, M, m)-extensions of $\mathscr{A}$.

THEOREM 2.1. *Let* $\mathscr{K}$ *be the class of all* (J, M, m)-*extensions of a Boolean algebra* $\mathscr{A}$. *The following are equivalent*:

1. $\mathscr{K}$ *contains a smallest element*;
2. $\mathscr{K}$ *is a lattice*;
3. $\mathscr{K}$ *is a complete lattice.*

Proof.

$1. \Rightarrow 3.$ It suffices to show that if S is a set of (J, M, m)-extensions of $\mathscr{A}$ then the greatest lower bound (glb) of S exists in $\mathscr{K}$, which follows from noting that if L is the set of all lower bounds for the set S then $L \neq 0$ and by Lemma 2.1 the lub of L exists in $\mathscr{K}$, hence is in L.

$3. \Rightarrow 2.$ By definition.

$2. \Rightarrow 1.$ If $\{i, \mathscr{B}\}$ is an m-completion of $\mathscr{A}$, $\{j, \mathscr{C}\} \in \mathscr{K}$, and $\mathscr{K}$ a lattice, then there is an element $\{j', \mathscr{C}'\} \in \mathscr{K}$ such that

$$\{j', \mathscr{C}'\} \leqq \{j, \mathscr{C}\} .$$

Thus

$$\{j', \mathscr{C}'\} \leqq \{i, \mathscr{B}\} ,$$

so

$$\{j', \mathscr{C}'\} = \{i, \mathscr{B}\} ,$$

implying

$$\{i, \mathscr{B}\} \leqq \{j, \mathscr{C}\} .$$

Hence $\{i, \mathscr{B}\}$ is a smallest element in $\mathscr{K}$.

COROLLARY 2.1. *If* $J' \supseteqq J$ *and* $M' \supseteqq M$ *then the following are equivalent*:

1. $\mathscr{K}(J, M, m)$ *contains a smallest element*;

2. *$\mathscr{K}(J', M', m)$ is a sublattice of $\mathscr{K}(J, M, m)$;*
3. *$\mathscr{K}(J', M', m)$ is a complete sublattice of $\mathscr{K}(J, M, m)$.*

Proof.

1. $\Rightarrow$ 3. Since $\mathscr{K}(J', M', m)$ contains a smallest element, so does $\mathscr{K}(J, M, m)$ hence $\mathscr{K}(J', M', m)$ and $\mathscr{K}(J, M, m)$ are complete lattices. If $\{\{i_t, \mathscr{B}_t\}\}_{t\in T} = S$ is a set of elements in $\mathscr{K}(J', M', m)$, $\{i, \mathscr{C}\}$ is the lub of S in $\mathscr{K}(J, M, m)$ and $\{i', \mathscr{C}'\}$ is the lub of S in $\mathscr{K}(J', M', m)$, then there is an m-homomorphism h mapping $\mathscr{C}'$ onto $\mathscr{C}$ such that $hi' = i$. Hence i is a (J', M', m)-isomorphism. Thus $\{i, \mathscr{C}\} \in \mathscr{K}(J', M', m)$, implying

$$\{i, \mathscr{C}\} = \{i', \mathscr{C}'\} .$$

If $\{i, \mathscr{C}\}$ is the glb of S in $\mathscr{K}(J, M, m)$ and $\{i', \mathscr{C}'\} \in S$, then by a similar argument, i is a (J', M', m)-isomorphism, which implies $\{i, \mathscr{C}\}$ is the glb of S in $\mathscr{K}(J', M', m)$.

3. $\Rightarrow$ 2. By definition.

2. $\Rightarrow$ 1. The proof is the same as that for showing 2. $\Rightarrow$ 1, in Theorem 2.1.

Thus it is of particular interest to know whether $\mathscr{K}(J, M, m)$ contains a smallest element, in general. Although, as it turns out, $\mathscr{K}(J, M, m)$ need not contain a smallest element in general, a minimal (J, M, m)-extension is always an m-completion, hence there is always a unique minimal (J, M, m)-extension in $\mathscr{K}$ (J, M, m).

PROPOSITION 2.2. *An m-completion $\{i, \mathscr{B}\}$ of the Boolean algebra $\mathscr{A}$ is a unique minimal element in $\mathscr{K}$.*

Proof. That a minimal element in $\mathscr{K}$ is an m-completion is clear.

If $\{i', \mathscr{B}'\}$ is another minimal element in $\mathscr{K}$, there are $\Delta, \Delta' \in I$ such that

$$\{i, \mathscr{B}\} = \{i_\Delta, \mathscr{A}_\Delta\}$$

and

$$\{i', \mathscr{B}'\} = \{i_{\Delta'}, \mathscr{A}_{\Delta'}\} .$$

Now $\{i, \mathscr{B}\}$ and $\{i', \mathscr{B}'\}$ minimal in $\mathscr{K}$ imply Δ and Δ' are maximal m-ideals in I, but if $\hat{\Delta}$ is a maximal m-ideal in I then $g_{\hat{\Delta}}(\mathscr{A}_{0,n})$ is dense in $\mathscr{A}_{\hat{\Delta}}$. The ideal $\hat{\Delta}' = \langle \hat{\Delta}, A \rangle$ in $\mathscr{A}_{m,n}$ is an m-ideal and $\hat{\Delta}' \in I$, contradicting the maximality of $\hat{\Delta}$. So $\{i', \mathscr{B}'\}$ is an m-completion of $\mathscr{A}$, hence isomorphic to $\{i, \mathscr{B}\}$, implying

$$\{i', \mathscr{B}'\} = \{i, \mathscr{B}\} .$$

PROPOSITION 2.3. *If $\mathscr{A}$ is a Boolean m-algebra that satisfies the m-chain condition and*

$$\bigcup_{t \in T} A_t$$

is the join of an indexed set $\{A_t\}_{t \in T}$ in $\mathscr{A}$, then there is an indexed set $\{A'_t\}_{t \in T}$ of disjoint elements of $\mathscr{A}$ such that

1. $$\bigcup_{t \in T} A'_t = \bigcup_{t \in T} A_t ;$$
2. $$A'_t \subseteqq A_t \quad \textit{for all} \quad t \in T .$$

Proof. Let $\mathscr{S}$ be the collection of all sets S of disjoint elements in $\mathscr{A}$ such that for each $s \in S$ there is a $t \in T$ with $s \subseteqq A_t$. If

$$S_1 \subseteqq S_2 \subseteqq \cdots \subseteqq S_i \subseteqq \cdots$$

is a chain of sets in $\mathscr{S}$ indexed by I and ordered by set theoretical inclusion, then

$$\bigcup_{i \in I} S_i = S \in \mathscr{S} .$$

By Zorn's lemma there is a maximal set in $\mathscr{S}$, say $S' = \{A_r\}_{r \in R}$, and it immediately follows that

$$\bigcup_{r \in R} A_r \neq A .$$

Now let

$$\varphi: S' \longrightarrow T$$

be a mapping such that if $A_r \in S'$ then

$$A_r \subseteqq A_{\varphi(A_r)} .$$

For each $t \in T$ define

$$A'_t = \bigcup \{A_r \in S' : \varphi(A_r) = t\}$$

if there is an $A_r \in S'$ such that $\varphi(A_r) = t$, otherwise define

$$A'_t = \wedge .$$

Then

$$\{A'_t\}_{t \in T}$$

is the desired set.

PROPOSITION 2.4. *Let $\mathscr{A}$ be a Boolean algebra. The following are equivalent:*

1. *$\mathscr{A}$ satisfies the m-chain condition:*
2. *for all sets S in $\mathscr{A}$ such that $\bigcup_{s \in S} s$ exists,*

$$\bigcup_{s \in S} s = \bigcup_{s \in S'} s$$

for some set $S' \subseteqq S$ with $S' \leqq m$; and dually for meets.

Proof.

$1. \Rightarrow 2.$ Suppose $\mathscr{A}$ satisfies the m-chain condition. It suffices to show that if

$$S = \{A_t\}_{t \in T} \text{ and } \mathbf{V} = \bigcup_{t \in T} A_t\ , \quad \overline{\overline{T}} = m' > m\ ,$$

then there is a set $T' \subseteqq T$, $\overline{\overline{T'}} \leqq m$, such that

$$\bigcup_{t \in T'} A_t = \mathbf{V}\ .$$

Let $\{i, \mathscr{B}\}$ be an m'-completion of $\mathscr{A}$. Then $\mathscr{B}$ satisfies the m-chain condition and

$$\begin{aligned} \mathbf{V}_{\mathscr{B}} &= i(\mathbf{V}_{\mathscr{A}}) \\ &= \bigcup_{t \in T}^{\mathscr{B}} i(A_t)\ . \end{aligned}$$

By Proposition 2.3, there is a set $\{\mathscr{B}_t\}_{t \in T}$ of disjoint elements in $\mathscr{B}$ such that

$$B_t \subseteqq i(A_t) \quad \text{and} \quad \bigcup_{t \in T}^{\mathscr{B}} B_t = \bigcup_{t \in T}^{\mathscr{B}} i(A_t)\ .$$

Since this set contains at most m-distinct elements,

$$\bigcup_{t \in T}^{\mathscr{B}} B_t = \bigcup_{t \in T'}^{\mathscr{B}} B_t\ ,$$

$T' \subseteqq T$ and $\overline{\overline{T'}} \leqq m$. Thus

$$\mathbf{V}_{\mathscr{B}} = \bigcup_{t \in T'}^{\mathscr{B}} i(A_t)$$

or

$$\mathbf{V}_{\mathscr{A}} = \bigcup_{t \in T'}^{\mathscr{A}} A_t\ .$$

$2. \Rightarrow 1.$ Suppose $\{A_t\}_{t \in T}$ is an m'-indexed set of disjoint elements of $\mathscr{A}$, $m' > m$. It may be assumed that $\{A_t\}_{t \in T}$ is a maximal set of disjoint elements of $\mathscr{A}$. Then for some $T' \subseteqq T$, $\overline{\overline{T'}} \leqq m$,

$$\mathbf{V}_{\mathscr{A}} = \bigcup_{t \in T'}^{\mathscr{A}} A_t\ .$$

Since $\overline{\overline{T'}} \neq \overline{\overline{T}}$, there is a $t_0 \in T - T'$ such that

$$A_{t_0} \in \{A_t\}_{t \in T} - \{A_t\}_{t \in T'} \quad \text{and} \quad A_{t_0} \neq \bigwedge_{\mathscr{A}} .$$

Thus

$$\bigcup_{t \in T'}^{\mathscr{A}} A_t \neq \bigvee_{\mathscr{A}} ,$$

a contradiction. Hence $\overline{\overline{T}} \leqq m$.

This gives, as an immediate corollary, the following result due to Sikorski [2].

COROLLARY 2.2. *If $\mathscr{A}$ is a Boolean m-algebra and satisfies the m-chain condition, it is a complete Boolean algebra.*

PROPOSITION 2.5. *The class $\mathscr{K}(J, M, m')$ contains a smallest element if $\mathscr{K}(J, M, m)$ contains a smallest element, $m' < m$.*

Proof. Let $\{i, \mathscr{B}\}$ be the smallest element in $\mathscr{K}(J, M, m)$. If $\{j', \mathscr{C}'\} \in \mathscr{K}(J, M, m')$, let $\{k, \mathscr{C}\}$ be an m-completion of $\mathscr{C}'$. Then $\{kj, \mathscr{C}\} \in \mathscr{K}(J, M, m)$.

By the fact that $\{i, \mathscr{B}\}$ is the smallest element in $\mathscr{K}(J, M, m)$, there is an m-homomorphism h such that

$$h\colon \mathscr{C} \longrightarrow \mathscr{B} \quad \text{and} \quad hkj = i .$$

Also $\{i, \mathscr{B}\}$ an m-completion of $\mathscr{A}$ implies that there is an m'-completion $\{i, \mathscr{B}'\}$ of $\mathscr{A}$ such that $\mathscr{B}' \subseteqq \mathscr{B}$. Thus $hk(\mathscr{C}')$ is an m-subalgebra of $\mathscr{B}$, hence $\mathscr{B}' \subseteqq hk(\mathscr{C}')$ and is an m-subalgebra of $\mathscr{C}$.

Now $kj(\mathscr{A})$ m-generates $k(\mathscr{C}')$ in $\mathscr{C}$ and $kj(\mathscr{A}) \subseteqq h^{-1}(\mathscr{B}')$, hence

$$h^{-1}(\mathscr{B}') \supseteqq k(\mathscr{C}') ,$$

or

$$h(h^{-1}(\mathscr{B}')) \supseteqq hk(\mathscr{C}') .$$

But

$$h(h^{-1}(\mathscr{B}')) = \mathscr{B}' ,$$

thus

$$\mathscr{B}' \supseteqq hk(\mathscr{C}') ,$$

so

$$\mathscr{B}' = hk(\mathscr{C}') .$$

Since $hkj = i$,

$$\{i, \mathscr{B}'\} \leqq \{kj, k(\mathscr{C}')\} .$$

But k a complete isomorphism implies that

$$\{kj, k(\mathscr{C}')\} \cong \{j, \mathscr{C}'\} ,$$

and since isomorphic elements in $\mathscr{K}(J, M, m)$ have been identified,

$$\{i, \mathscr{B}'\} = \{j, \mathscr{C}'\} .$$

LEMMA 2.2. *If* $\overline{\overline{J}} \leqq \sigma$ *and* $\overline{\overline{M}} \leqq \sigma$ *then there is a* (J, M, m)*-isomorphism* i *of a Boolean algebra* $\mathscr{A}$ *into the field* $\mathscr{F}$ *of all subsets of a space.*

PROPOSITION 2.6. *If the Boolean algebra* $\mathscr{A}$ *is* m*-representable but not* m^+*-representable,* m^+ *the smallest cardinal greater than* m*, then* $\mathscr{K}(J, M, m^+)$ *does not contain a smallest element if*

$$\mathscr{K}_r(J, M, m^+) \neq \emptyset .$$

If $\overline{\overline{J}} \leqq \sigma$, $\overline{\overline{M}} \leqq \sigma$ *then* $\mathscr{K}_r(J, M, m^+) \neq \emptyset$.

Proof. Suppose $\{j, \mathscr{C}\} \in \mathscr{K}_r(J, M, m^+)$. Then $\mathscr{C}$ is m-representable and if an m^+-completion $\{i, \mathscr{B}\}$ of $\mathscr{A}$ is a smallest element in $\mathscr{K}(J, M, m^+)$, there is a surjective m^+-homomorphism

$$h: \mathscr{C} \longrightarrow \mathscr{B} ,$$

which implies $\mathscr{B}$ is m^+-representable, hence $\mathscr{A}$ is m^+-representable, a contradiction. Thus $\mathscr{K}(J, M, m^+)$ does not contain a smallest element if $\mathscr{K}_r(J, M, m^+) \neq \emptyset$.

If $\overline{\overline{J}} \leqq \sigma$ and $\overline{\overline{M}} \leqq \sigma$ then $\mathscr{A}$ is (J, M, m^+)-representable by Lemma 2.2, hence $\mathscr{K}_r(J, M, m^+) \neq \emptyset$.

The next proposition is an easy generalization of Sikorski's [2] Proposition 25.2 and will be needed for the last theorem in this section.

PROPOSITION 2.7. *A Boolean algebra* $\mathscr{A}$ *is completely distributive, if, and only if, it is atomic.*

COROLLARY 2.3. *A Boolean algebra* $\mathscr{A}$ *is completely distributive, if, and only if,* $\mathscr{A}$ *is* m*-distributive,* $m = \overline{\overline{\mathscr{A}}}$.

The following proposition is due to Sikorski [2] and will be given without proof.

PROPOSITION 2.8. *If the Boolean algebra* $\mathscr{A}$ *is* m*-distributive, then* $\mathscr{K}(J, M, m)$ *contains a smallest element for arbitrary* J *and* M.

LEMMA 2.3. *If $\{i, \mathscr{B}\}$ is an m-extension of the Boolean algebra $\mathscr{A}$ and $\mathscr{B}$ is m-representable, then $\mathscr{A}$ is m-representable.*

Proof. This follows immediately from the fact that $\mathscr{A}$ is m-regular in $\mathscr{B}$.

Now to prove the main theorem of this section.

THEOREM 2.2. *Let $\mathscr{A}$ be a Boolean algebra. Then the following are equivalent:*

1. *$\mathscr{K}$ contains a smallest element for arbitrary J, M, and m;*
2. *$\mathscr{A}$ is m-representable for all m;*
3. *$\mathscr{A}$ is completely distributive;*
4. *$\mathscr{A}$ is atomic;*
5. *an m-completion of $\mathscr{A}$ is atomic for all m;*
6. *an m-completion of $\mathscr{A}$ is in $\mathscr{K}_r(J, M, m)$ for arbitrary J, M, and m;*
7. *$\mathscr{K}(J, M, 2^{m^*})$ contains a smallest element, where $J = M = \varnothing$ and $\overline{\overline{\mathscr{A}}} = m^*$.*

Proof.

1. $\Rightarrow$ 2. If $\mathscr{A}$ is m-representable but not m^*-representable, then Proposition 2.6 implies $\mathscr{K}(J, M, m^*)$ does not contain a smallest element if $\overline{\overline{J}}, \overline{\overline{M}} < \sigma$.

2. $\Rightarrow$ 3. This follows from the fact that if a Boolean algebra $\mathscr{A}$ is 2^m-representable, it is m-distributive.

3. $\Leftrightarrow$ 4. This follows from Proposition 2.7.

3. $\Rightarrow$ 1. This follows from Proposition 2.8.

4. $\Leftrightarrow$ 5. If $\{i, \mathscr{B}\}$ is an m-completion of $\mathscr{A}$ then $i(\mathscr{A})$ is dense in $\mathscr{B}$, so $\mathscr{B}$ is atomic, and conversely.

2. $\Rightarrow$ 6. This follows from noting that 2. $\Rightarrow$ 3. and $\mathscr{A}$ completely distributive implies an m-completion of $\mathscr{A}$ is completely distributive, hence m-representable for all cardinals m.

6. $\Rightarrow$ 2. This follows from Lemma 2.3.

3. $\Leftrightarrow$ 7. If $J = M = \varnothing$ and $\mathscr{K}(J, M, 2^{m^*})$ contains a smallest element, then by Proposition 2.6, $\mathscr{A}$ is 2^{m^*}-representable, hence m^*-distributive. Since $m^* = \overline{\overline{\mathscr{A}}}$, $\mathscr{A}$ is completely distributive, by Corollary 2.3. The converse is clear.

3. The example in § 2 of a Boolean algebra $\mathscr{A}$ such that the class of all (J, M, m)-extensions of $\mathscr{A}$ does not contain a smallest element depends on the assumption that $\bar{\bar{J}}, \bar{\bar{M}} \leqq \sigma$. Thus it is of interest to know whether an example can be found showing that the class of all m-extensions of $\mathscr{A}$ does not contain a smallest element, since this corresponds to the case where J and M are as large as possible. As it turns out, there are Boolean algebras $\mathscr{A}$ such that the class of all m-extensions $\mathscr{K}$ does not contain a smallest element. In this section such an example will be constructed for each infinite cardinal m and several general types of Boolean algebras such that $\mathscr{K}$ does not contain a smallest element will be given.

Throughout this section $\mathscr{K}$ will denote the class of all m-extensions of a Boolean algebra $\mathscr{A}$ and $\mathscr{K}(J, M, m)$ the class of all (J, M, m)-extensions.

If $\mathscr{A}$ is a Boolean algebra and $\{i, \mathscr{C}\} \in \mathscr{K}(J, M, m)$, let

$$K(\mathscr{C}) = \{C \in \mathscr{C} \colon \text{if } i(A) \subseteqq C,\ A \in \mathscr{A}, \text{ then } A = \Lambda_{\mathscr{A}}\} ,$$

and

$$K_P(\mathscr{C}) = \{C \in \mathscr{C} \colon \text{if } P = \{A \in \mathscr{A} \colon i(A) \supseteqq C\} \text{ then } \bigcap_{A \in P}^{\mathscr{A}} A = \Lambda_{\mathscr{A}}\} .$$

Note that $K_P(\mathscr{C}) \subseteqq K(\mathscr{C})$.

LEMMA 3.1. *The set* $K_P(\mathscr{C})$ *is an ideal and* $K(\mathscr{C}) = K_P(\mathscr{C})$, *if, and only if,* $K(\mathscr{C})$ *is an ideal.*

Proof. It follows easily that $K_P(\mathscr{C})$ is an ideal.

If $K(\mathscr{C})$ is an ideal and $\mathscr{C} \in K(\mathscr{C})$ let

$$P = \{A \in \mathscr{A} \colon i(A) \supseteqq C\} .$$

If $A' \in \mathscr{A}$ and $A' \subseteqq A$ for all $A \in P$, then

$$i(A') - C \in K(\mathscr{C}) .$$

Now $i(A') \cap C \in K(\mathscr{C})$, hence

$$i(A') = (i(A') - C) \cup (i(A') \cap C) \in K(\mathscr{C}) ,$$

which implies $i(A') = \Lambda_{\mathscr{C}}$ or $A' = \Lambda_{\mathscr{A}}$. Thus

$$\bigcap_{A \in P}^{\mathscr{A}} A = \Lambda_{\mathscr{A}} ,$$

so $C \in K_P(\mathscr{C})$, and

$$K_P(\mathscr{C}) = K(\mathscr{C}) .$$

Since $K_P(\mathscr{C})$ is an ideal, the converse is true.

PROPOSITION 3.1. *If $\mathscr{A}$ is a Boolean algebra the following are equivalent:*

1. $\mathscr{K}(J, M, m)$ *contains a smallest element;*
2. $K(\mathscr{C}) = K_P(\mathscr{C})$ *for all* $\{i, \mathscr{C}\} \in \mathscr{K}(J, M, m)$;
3. $K(\mathscr{C}) = K_P(\mathscr{C})$ *if* $\{i, \mathscr{C}\}$ *is the maximum element in* $\mathscr{K}(J, M, m)$.

Proof.

1. $\Rightarrow$ 2. Suppose $\mathscr{K}(J, M, m)$ contains a smallest element $\{i, \mathscr{B}\}$, and there is an element

$$\{j, \mathscr{C}\} \in \mathscr{K}(J, M, m)$$

with the property that

$$K(\mathscr{C}) \neq K_P(\mathscr{C}) .$$

Let h be the unique m-homomorphism mapping $\mathscr{C}$ onto $\mathscr{B}$ such that $hj = i$. Let $\ker h$ be the kernel of this mapping. Then

$$K_P(\mathscr{C}) \subseteqq \ker h \subseteqq K(\mathscr{C}) ,$$

and

$$\ker h \neq K(\mathscr{C}) .$$

Pick $x \in K(\mathscr{C}) - \ker h$ and let

$$\Delta = \langle x \rangle ,$$

so Δ is a complete ideal. Thus

$$\{i_\Delta, \mathscr{C}/\Delta\} \in \mathscr{K}(J, M, m) ,$$

where

$$i_\Delta\colon \mathscr{A} \rightarrow \mathscr{C}/\Delta$$

is defined by

$$i_\Delta(A) = [i(A)]_\Delta .$$

Consequently, there are unique homomorphisms h_Δ and h' mapping $\mathscr{C}$ onto $\mathscr{C}/\Delta$, $\mathscr{C}/\Delta$ onto $\mathscr{B}$, and satisfying $h_\Delta j = i_\Delta$, $h' i_\Delta = i$, respectively. Hence

$$h' h_\Delta j = h' i_\Delta = i$$

and by the uniqueness of h,

$$h = h' h_\Delta .$$

This implies

$$h(x) = h' h_\Delta(x) = \bigwedge\nolimits_{\mathscr{B}} ,$$

a contradiction. Thus

$$K(\mathscr{C}) = K_P(\mathscr{C})\,.$$

2. $\Rightarrow$ 3. Obvious.

3. $\Rightarrow$ 1. To show that $\mathscr{K}(J, M, m)$ contains a smallest element, let $\{j, \mathscr{C}\}$ be the largest element in $\mathscr{K}(J, M, m)$ and suppose $\{j', \mathscr{C}'\} \in \mathscr{K}(J, M, m)$. Let $\{i, \mathscr{B}\}$ be an m-completion of $\mathscr{A}$. Then there is an m-homomorphism h' mapping $\mathscr{C}$ onto $\mathscr{C}'$ such that $h'j = j'$ and an m-homomorphism h mapping $\mathscr{C}$ onto $\mathscr{B}$ such that $hj = i$. Thus

$$K_P(\mathscr{C}) \subseteqq \ker h \subseteqq K(\mathscr{C})\,,$$

which implies, by assumption, that

$$K_P(\mathscr{C}) = \ker h = K(\mathscr{C})\,,$$

so $K_P(\mathscr{C})$ and $K(\mathscr{C})$ are m-ideals in $\mathscr{C}$. Further,

$$h'(K_P(\mathscr{C})) \subseteqq K_P(\mathscr{C}') \subseteqq K(\mathscr{C}') \subseteqq h'(K(\mathscr{C}))\,.$$

This implies that

$$h'(K_P(\mathscr{C})) = K_P(\mathscr{C}') = K(\mathscr{C}') = h'(K(\mathscr{C}))\,,$$

hence $K(\mathscr{C}')$ is an m-ideal. Let

$$\varDelta = K(\mathscr{C}')\,.$$

Then $\mathscr{C}'/\varDelta$ is an m-algebra and

$$j'_\varDelta(\mathscr{A}) = \{[j'(A)]_\varDelta \colon A \in \mathscr{A}\}$$

m-generates $\mathscr{C}'/\varDelta$. Finally, $j'_\varDelta(\mathscr{A})$ is dense in $\mathscr{C}'/\varDelta$. Thus $\{j'_\varDelta, \mathscr{C}'/\varDelta\}$ is an m-completion of $\mathscr{A}$, hence is equal to $\{i, \mathscr{B}\}$, as isomorphic elements of $\mathscr{K}(J, M, m)$ have been identified. The m-homomorphism

$$h_\varDelta \colon \mathscr{C}' \longrightarrow \mathscr{C}'/\varDelta$$

defined by

$$h_\varDelta(C') = [C']_\varDelta$$

has the property that

$$h_\varDelta j = j'_\varDelta \quad \text{for all} \quad A \in \mathscr{A}\,,$$

implying that

$$\{i_\varDelta, \mathscr{C}'/\varDelta\} \leqq \{j', \mathscr{C}'\}\,.$$

Hence $\mathscr{K}(J, M, m)$ contains a smallest element.

This, then, gives a way to construct a Boolean algebra $\mathscr{A}$ such that $\mathscr{K}$ does not contain a smallest element. Namely, by finding a Boolean algebra $\mathscr{A}$ with an m-extension $\{i, \mathscr{C}\}$ such that $K_P(\mathscr{C}) \neq K(\mathscr{C})$. The next task is to construct such a Boolean algebra.

If $\overline{\overline{T}} = m$ and $\mathscr{A} = \mathscr{A}_t$ for all $t \in T$, the Boolean product of $\{\mathscr{A}_t\}_{t \in T}$ will be called the *m-fold product of* $\mathscr{A}$. Note that if $\mathscr{A}$ is a subalgebra of the Boolean algebra $\mathscr{A}'$, $\mathscr{F}$ is the m-fold product of $\mathscr{A}$ and $\mathscr{F}'$ is the m-fold product of $\mathscr{A}'$, then $\mathscr{F} \subseteqq \mathscr{F}'$.

LEMMA 3.2. *If $\mathscr{A}$ is an m-regular subalgebra of the Boolean algebra $\mathscr{A}'$ then the Boolean m-fold product $\mathscr{F}$ of $\mathscr{A}$ is isomorphic to an m-regular subalgebra of the Boolean m-fold product $\mathscr{F}'$ of $\mathscr{A}'$.*

Proof. Since $\mathscr{A}$ is a subalgebra of $\mathscr{A}'$, $\mathscr{F} \subseteqq \mathscr{F}'$. Let $\mathscr{S}(\mathscr{S}')$ be the set of all $\varphi_t(A)$, $A \in \mathscr{A}$ and $t \in T (A \in \mathscr{A}'$ and $t \in T)$. Then $F \in \mathscr{S} (F \in \mathscr{S}')$ implies $-F \in \mathscr{S}(-F \in \mathscr{S}')$ and $\mathscr{S}(\mathscr{S}')$ are sets of generators for $\mathscr{F}(\mathscr{F}')$. For elements $F \in \mathscr{F}'$ of the form

$$F = \bigcap_{i=1}^{N} F_i \,, \qquad F_i \in \mathscr{S} \,,$$

define

$$\lambda_t(F) = \left\{\pi_t(x) \colon x \in \bigcap_{i=1}^{N} F_i\right\} .$$

Note that if $F \in \mathscr{S}'$ and $t \in T$ is such that $\lambda_t(F) \neq \bigvee_{\mathscr{A}'}$ then $\varphi_t(\lambda_t(F)) = F$.

In order to show $\mathscr{F}$ is m-regular in $\mathscr{F}'$, it suffices to prove that if $\{F_t\}_{t \in T}$ is an m-indexed set of elements of $\mathscr{F}$ such that

$$\bigcap_{t \in T}^{\mathscr{F}} F_t = \bigwedge_{\mathscr{F}}$$

then

$$\bigcap_{t \in T}^{\mathscr{F}'} F_t = \bigwedge_{\mathscr{F}'} .$$

Now $F_t \in \mathscr{F}$ so F_t may be rewritten as

$$F_t = \bigcap_{p=1}^{P_t} \bigcup_{q=1}^{Q_t} F_{p,q,t} \,,$$

where P_t, Q_t are finite numbers and $F_{p,q,t} \in \mathscr{S}$, for all $p \in P_t$, $q \in Q_t$, and $t \in T$. Thus

$$\Lambda_{\mathscr{F}} = \bigcap_{t\in T}^{\mathscr{F}} \bigcap_{p=1}^{P_t} \bigcup_{q=1}^{Q_t} F_{p,q,t}$$
$$= \bigcap_{s\in S}^{\mathscr{F}} \bigcup_{q=1}^{Q_s} F_{s,q}$$

after a suitable re-indexing, where $\bar{\bar{\bar{S}}} \leqq m$ and $F_{s,q} = F_{p,q,t}$ for suitable $p \in P_t$, $t \in T$. Without loss of generality, assume that for each $s \in S$, $\lambda_t(F_{s,q}) \neq \Lambda_{\mathscr{A}'}$ implies $\lambda_t(F_{s,q'}) = \mathsf{V}_{\mathscr{A}'}$ for all $t \in T$ and $q' \neq q$, and that $F_{s,q} \neq \mathsf{V}_{\mathscr{F}'}$ for all q, $1 \leqq q \leqq Q_s$, and all $s \in S$. Suppose $F' \in \mathscr{F}'$ and $F' \subseteqq F_t$ for all $t \in T$. Then

$$F' = \bigcup_{m=1}^{M} \bigcap_{n=1}^{N} F'_{m,n}\,, \qquad F'_{m,n} \in \mathscr{S}'\,,$$

so

$$\bigcap_{n=1}^{N} F'_{m,n} \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

for $1 < m \leqq M$, and all $s \in S$. Thus to show $F' = \Lambda_{\mathscr{F}'}$, it suffices to prove that if

$$\bigcap_{n=1}^{N} F'_n \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}\,,$$

for all $s \in S$, where $F'_n \in \mathscr{S}'$, then

$$\bigcap_{n=1}^{N} F'_n = \Lambda_{\mathscr{F}'}\,.$$

It may be assumed that for each n, $1 \leqq n \leqq N$, $\lambda_t(F'_n) \neq \Lambda_{\mathscr{A}'}$ implies $\lambda_t(F'_{n'}) = \mathsf{V}_{\mathscr{A}'}$ for all $t \in T$ and $n' \neq n$, and that $F'_n \neq \mathsf{V}_{\mathscr{F}'}$ for all n, $1 \leqq n \leqq N$.

Now

$$\bigcap_{n=1}^{N} F'_n \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

implies

$$\bigcap_{n=1}^{N} F'_n \cap \bigcup_{q=1}^{Q_s} -F_{s,q} = \Lambda_{\mathscr{F}'}\,,$$

and as each F'_n and $-F_{s,q}$ is of the form $\varphi_t(A)$ for some $A \in \mathscr{A}'$ and $t \in T$, the independence of the indexed set $\{\varphi_t(\mathscr{A}')\}_{t\in T}$ of subalgebras of $\mathscr{F}'$ implies that for some n_s, $1 \leqq n_s \leqq N$, and some q_s, $1 \leqq q_s \leqq Q_s$,

$$F'_{n_s} \cap -F_{s,q_s} = \Lambda_{\mathscr{F}'}\,,$$

which implies $F'_{n_s} \subseteqq F_{s,q_s}$. This argument may be repeated for each $s \in S$.

The set $\{n_s: s \in S\}$ is finite so let $\{n_s: s \in S\} = \{n_i: 1 \leqq i \leqq N'\}$. Let $S_i = \{s \in S: F'_{n_i} \subseteqq F_{s,q_s}\}$. If $t_s \in T$ is such that

$$\lambda_{t_s}(F_{s,q_s}) \neq \bigvee_{\mathscr{A}'} \quad \text{for all} \quad s \in S$$

then $\lambda_{t_s}(F_{s,q_s}) \in \mathscr{A}$ and

$$\bigcap_{s \in S_i}^{\mathscr{A}} \lambda_{t_s}(F_{s,q_s}) \neq \bigwedge_{\mathscr{A}'} .$$

Thus

$$\bigcap_{s \in S_i}^{\mathscr{A}'} \lambda_{t_s}(F_{s,q_s}) \neq \bigwedge_{\mathscr{A}'} ,$$

or

$$\bigcap_{s \in S_i}^{\mathscr{A}} \lambda_{t_s}(F_{s,q_s}) \neq \bigwedge_{\mathscr{A}'} ,$$

hence there is an $A_i \in \mathscr{A}$, $A_i \neq \bigwedge_{\mathscr{A}'}$, with

$$A_i \subseteqq \lambda_{t_s}(F_{s,q_s}) \quad \text{for all} \quad s \in S_i .$$

Let $A_{t,i}$ be the set of all $x \in X$ such that $\pi_{t_s}(x) \in A_i$. Thus $A_{t,i} \in \mathscr{F}$ and this argument may be repeated for each i, $1 \leqq i \leqq N'$. Now

$$\bigwedge_{\mathscr{F}'} \neq \bigcap_{i=1}^{N'} A_{t,i}$$

and

$$\bigcap_{i=1}^{N'} A_{t,i} \subseteqq \bigcup_{q=1}^{Q_s} F_{q,s}$$

for all $s \in S$. But then

$$\bigcap_{i=1}^{N'} A_{t,i} \subseteqq \bigcap_{s \in S}^{\mathscr{F}} \bigcup_{q=1}^{Q_s} F_{q,s} = \bigwedge_{\mathscr{F}} ,$$

a contradiction. Thus $\mathscr{F}$ is m-regular in $\mathscr{F}'$.

The next lemma assumes there is a Boolean algebra $\mathscr{A}$ such that an m-extension is not an m-completion. Sikorski [2] cites an example due to Katětov of such a Boolean algebra for the case $m = \sigma$. As Lemmas 3.5 and 3.6 imply, there is such an $\mathscr{A}$ for all infinite cardinal numbers m.

Assume for the moment that $\mathscr{A}$ is a Boolean algebra such that $\mathscr{K}$ contains more than one element and $\{i, \mathscr{B}\} \in \mathscr{K}$ is an m-extension that is not an m-completion. Thus there is a $B \in \mathscr{B}$ such that $i(A) \subseteqq B$, $A \in \mathscr{A}$, implies $A = \bigwedge_{\mathscr{A}}$. Let $\mathscr{F}'$ be the Boolean m-fold product of $\mathscr{B}$, h_0 an isomorphism of $\mathscr{B}$ onto the Stone space $\mathscr{F}$ of

$\mathscr{B}$, X the Cartesian product of $\mathscr{F}$ with itself m times and indexed by T, and

$$B_t = \varphi_t h_0(B) \quad \text{for all} \quad t \in T .$$

Let

$$B_0 = \bigcup_{t \in T'} B_t ,$$

where T' is a fixed, but arbitrary subset of T such that $\bar{\bar{T}}' \geqq \sigma$, and define

$$\mathscr{F}_0 = \langle \mathscr{F}', B_0 \rangle .$$

Since $\bar{\bar{T}}' \geqq \sigma$, $\mathscr{F}_0 \neq \mathscr{F}'$.

LEMMA 3.3. *If $\mathscr{F}$ is the Boolean m-fold product of $\mathscr{A}$ then $\mathscr{F}$ is isomorphic to an m-regular subalgebra of $\mathscr{F}_0$.*

Proof. It may be assumed, without loss of generality, that $\mathscr{A} \subseteqq \mathscr{B}$. Thus $\mathscr{F} \subseteqq \mathscr{F}_0$. Let $\mathscr{S}(\mathscr{S}')$ be a generating set for $\mathscr{F}(\mathscr{F}')$. Let

$$\mathscr{S}_0 = \mathscr{S}' \cup \{B_0\} ,$$

so $\mathscr{S}_0$ is a generating set for $\mathscr{F}_0$. As in the previous lemma, to prove $\mathscr{F}$ is m-regular in $\mathscr{F}_0$ it suffices to show that if

$$\bigcap_{n=1}^{N} F'_n \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

for all $s \in S$, $\bar{\bar{S}} \leqq m$; and

$$\bigcap_{s \in S}^{\mathscr{F}} \bigcup_{q=1}^{Q_s} F_{s,q} = \Lambda_{\mathscr{F}} ;$$

$F_{s,q} \in \mathscr{S}$ for all $s \in S$ and $1 \leqq q \leqq Q_s$, $F'_n \in \mathscr{S}_0$, $1 \leqq n \leqq N$; then

$$\bigcap_{n=1}^{N} F'_n = \Lambda_{\mathscr{F}'} .$$

Since $F'_n \in \mathscr{S}_0$, there is an n, $1 \leqq n \leqq N$, such that $F'_n = B_0$ or $F'_n = -B_0$, otherwise there is nothing to prove. This may be reduced to two cases:

Case 1.

$$\bigcap_{n=1}^{N} F'_n \cap B_0 \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

for all $s \in S$, where $F'_n \in \mathscr{S}'$ and $F_{s,q} \in \mathscr{S}$.

Case 2.

$$(-B_0) \cap \bigcap_{n=1}^{N} F'_n \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

for all $s \in S$, where $F'_n \in \mathscr{S}'$ and $F_{s,q} \in \mathscr{S}$.

Proof of Case 1. If for each $s \in S$ there is an n_s, $1 \leqq n_s \leqq N$, such that there is a q_s, $1 \leqq q_s \leqq Q_s$, with $F'_{n_s} \subseteqq F_{s,q_s}$, then

$$\bigcap_{n=1}^{N} F'_n \subseteqq \bigcup_{q=1}^{Q_s} F_{s,q}$$

for all $s \in S$, and

$$\bigcap_{n=1}^{N} F'_n \in \mathscr{F}'$$

implies

$$\bigcap_{n=1}^{N} F'_n = \bigwedge_{\mathscr{F}'} .$$

Thus it may be assumed there is an s_0 such that

$$\bigcap_{n=1}^{N} F'_n \not\subseteqq \bigcup_{q=1}^{Q_{s_0}} F_{s_0,q} .$$

Hence for all n, $F'_n \subseteqq F_{s_0,q}$ for some q, is false. If

$$\bigcap_{n=1}^{N} F'_n \cap B_0 \neq \bigwedge_{\mathscr{F}'} ,$$

let $x \in X$ be defined as follows. Let $t_1, \cdots, t_n \in T$ be such that $\lambda_{t_i}(F'_i) \neq \bigvee_{\mathscr{B}}$, $1 \leqq i \leqq N$. Choose an $x \in X$ such that it satisfies the following conditions:

(a)

$$\pi_i(x) \in \begin{cases} \lambda_{t_i}(F'_i) \text{ if } \lambda_{t_i}(F_{s_0,q}) = \bigvee_{\mathscr{B}} \text{ for all } q,\ 1 \leqq q \leqq Q_{s_0} \\ \lambda_{t_i}(F'_i) - \lambda_{t_i}(F_{s_0,q_0}) \text{ if } \lambda_{t_i}(F_{s_0,q_0}) \neq \bigvee_{\mathscr{B}} \end{cases}$$

for $1 \leqq i \leqq N$;

(b) $\pi_{t_q}(s) \in -\lambda_{t_q}(F_{s_0,q})$ for each $t_q \in T$ such that $\lambda_{t_q}(F_{s_0,q}) \neq \bigvee_{\mathscr{B}}$, $1 \leqq q \leqq Q_{s_0}$ and $t_q \neq t_i$, $1 \leqq i \leqq n$;

(c) $\pi_t(x) \in h_0(B)$ for all $t \neq t_q$; $1 \leqq i \leqq N$, $1 \leqq q \leqq Q_{s_0}$.

Now x is well defined,

$$x \in B_0 \quad \text{and} \quad x \in \bigcap_{n=1}^{N} F'_n ,$$

by its definition. But $x \notin F_{s_0,q}$ for all q, $1 \leqq q \leqq Q_{s_0}$, hence

$$x \notin \bigcup_{q=1}^{Q_{s_0}} F_{s_0,q} ,$$

a contradiction.

Proof of Case 2. If

$$-B_0 \cap \bigcap_{n=1}^{N} F'_n \neq \bigwedge_{\mathscr{F}'}$$

and $\lambda_{t_n}(F'_n) \neq \bigvee_{\mathscr{B}}$, $t_n \in T$, let $A_n = \varphi_{t_n}(-B_0)$, $1 \leqq n \leqq N$. Then

$$\bigcap_{n=1}^{N} F'_n \cap (-B_0) = \bigcap_{n=1}^{N} (F'_n \cap A_n) \cap (-B_0)$$

and

$$\bigcap_{n=1}^{N} (F'_n \cap A_n) \in \mathscr{F}' .$$

As before, an $s_0 \in S$ may be found such that

$$\bigcap_{n=1}^{N} (F'_n \cap A_n) \nsubseteq \bigcup_{q=1}^{Q_{s_0}} F_{s_0,q} .$$

Define $t_1, \cdots, t_N$ as before so that $\lambda_{t_i}(F'_i \cap A_i) \neq \bigvee_{\mathscr{B}}$, $1 \leqq i \leqq N$. Choose $x \in X$ satisfying the following conditions:

(a)

$$\pi_{t_i}(x) \in \begin{cases} \lambda_{t_i}(F'_i \cap A_i) \text{ if } \lambda_{t_i}(F_{s_0,q}) = \bigvee_{\mathscr{B}},\ 1 \leqq q \leqq Q_{s_0} \\ \lambda_{t_i}(F'_i \cap A_i) - \lambda_{t_i}(F_{s_0,q}) \text{ if } \lambda_{t_i}(F_{s_0,q_0}) \neq \bigvee_{\mathscr{B}} \end{cases}$$

for $1 \leqq i \leqq N$.

(b) $\pi_{t_q}(x) \in -\lambda_{t_q}(F_{s_0,q})$ for each $t_q \in T$ such that $\lambda_{t_q}(F_{s_0,q}) \neq \bigvee_{\mathscr{B}}$; $1 \leqq q \leqq Q_{s_0}$, and $t_q \neq t_i$, $1 \leqq i \leqq N$.

(c) $\pi_t(x) \in \lambda_t(-B_0)$ if $t \neq t_i, t_q$; $1 \leqq i \leqq n$, $1 \leqq q \leqq Q_{s_0}$.

Now x is well defined and

$$x \in (-B_0) \cap \bigcap_{n=1}^{N} (F'_n \cap A_n) = -B_0 \cap \bigcap_{n=1}^{N} F'_n ,$$

so

$$x \notin \bigcup_{q=1}^{Q_{s_0}} F_{s,q} ,$$

a contradiction.

Consequently, in either case

$$\bigcap_{n=1}^{N} F'_n = \bigwedge_{\mathscr{F}'} .$$

LEMMA 3.4. *If j is the identity isomorphism of $\mathscr{F}$ into $\mathscr{F}_0$ and $\{i, \mathscr{C}\}$ is an m-completion of $\mathscr{F}_0$, then $\{ij, \mathscr{C}\}$ is an m-extension of $\mathscr{F}$.*

Proof. All that needs to be shown is that $ij(\mathscr{F})$ m-generates $\mathscr{C}$. But this follows immediately from the fact that $\mathscr{A}$ m-generates $\mathscr{B}$ and the definition of $\mathscr{F}$ and $\mathscr{F}_0$.

THEOREM 3.1. *If $\mathscr{A}$ m-generates $\mathscr{B}$ then $\mathscr{K}(\mathscr{F})$ does not contain a smallest element.*

Proof. $F \in \mathscr{F}$ and $F \supseteqq B_0$ then $F = \bigvee_{\mathscr{F}_0}$, by definition of B_0. Thus if j and $\{i, \mathscr{C}\}$ are defined as in Lemma 3.4, $\{ij, \mathscr{C}\}$ is an m-extension of $\mathscr{F}$ and $ij(B_0) \in K(\mathscr{C})$. By Proposition 3.1, $\mathscr{K}(\mathscr{F})$ does not contain a smallest element.

The results of this theorem may be generalized as follows. Let $\{\mathscr{A}_t\}_{t \in T}$ be an infinite indexed set of Boolean algebras and $\{\{i_t\}_{t \in T}, \mathscr{B}\}$ be the Boolean product of $\{\mathscr{A}_t\}_{t \in T}$. Let T' be the set of all $t \in T$ such that $\mathscr{K}(\mathscr{A}_t)$ contains more than one element.

THEOREM 3.2. *The class of m-extensions $\mathscr{K}(\mathscr{B})$ does not contain a smallest element if $\bar{\bar{T}}' \geqq \sigma$.*

Proof. Define $\mathscr{F}'$ to be the Boolean product of $\{\{j_t, \mathscr{B}_t\}\}_{t \in T}$, where $\{j_t, \mathscr{B}_t\} \in \mathscr{K}(\mathscr{A}_t)$ for all $t \in T$ and $\{j_t, \mathscr{B}_t\}$ is not an m-completion of $\mathscr{A}_t$ for all $t \in T'$. For each $\mathscr{B}_t$, $t \in T'$, there is a $B_t \in \mathscr{B}_t$ such that $j_t(A) \subseteqq B_t$, $A \in \mathscr{A}_t$, implies $A = \bigwedge_{\mathscr{A}_t}$. Let φ_t map $\mathscr{B}_t$ into $\mathscr{B}$ and set

$$B_0 = \bigcup_{t \in T'}^{\mathscr{B}} \varphi_t(B_t)$$

and

$$\mathscr{F}_0 = \langle \mathscr{F}', B_0 \rangle \,.$$

Then by an argument similar to the proofs of Lemmas 3.2, 3.3, and 3.4, and Theorem 3.1, $\mathscr{K}(\mathscr{B})$ does not contain a smallest element.

COROLLARY 3.1. *If $\mathscr{A}_t = \mathscr{A}_{t'}$ for all $t, t' \in T$ then $\mathscr{K}(\mathscr{B})$ contains a smallest element if, and only if, an m-extension of $\mathscr{B}$ is an m-completion.*

Proof. If $\mathscr{K}(\mathscr{B})$ contains an m-extension which is not an m-completion, let $\mathscr{B}$ play the role of $\mathscr{A}$ in Lemmas 3.2, 3.3, and 3.4. By Theorem 3.1, $\mathscr{K}(\mathscr{F})$ does not contain a smallest element. As the m-fold product $\mathscr{F}$ of $\mathscr{B}$ is isomorphic to $\mathscr{B}$, $\mathscr{K}(\mathscr{B})$ does not contain a smallest element. The converse is clear.

Now to prove the assumption on which these results are based.

LEMMA 3.5. *For each infinite cardinal number m there is a Boolean algebra $\mathscr{A}$ such that an m-completion $\{i, \mathscr{B}\}$ of $\mathscr{A}$ contains an element B with*

$$B \neq \bigcup_{u \in U}^{\mathscr{B}} \bigcap_{v \in V}^{\mathscr{B}} A_{u,v} ,$$

for all m-indexed sets $\{A_{u,v}\}_{u \in U, v \in V}$ in $\mathscr{A}$.

Proof. The proof will be by constructing such an $\mathscr{A}$ for each m. Let S be an indexing set of cardinality m. Let $\mathscr{D}_m$ be the Cartesian product of S with itself m times and indexed by T. Define

$$D_{t,s} = \{d \in \mathscr{D}_m : \pi_t(d) = s\} .$$

Fix $s_1', s_2' \in S$, $s_1' \neq s_2'$, and set $S' = S - \{s_1', s_2'\}$. Let $D = \bigcup_{t \in T}(D_{t,s_1'} \cup D_{t,s_2'})$. Thus $\overline{\overline{D}} = 2^m$ and $d \in \mathscr{D}_m - D$ implies $\pi_t(d) \neq s_k'$, $k = 1, 2$, for all $t \in T$.

Let

$$\mathscr{S} = \{\{d\} : d \in \mathscr{D}_m\} \cup \{D_{t,s} : t \in T, s \in S'\} .$$

Let $\mathscr{A}$ be generated by $\mathscr{S}$ in $\mathscr{D}_m$ and let $\mathscr{B}$ be the m-field of sets m-generated by $\mathscr{S}$ in $\mathscr{D}_m$. Then $\mathscr{A}$ is dense in $\mathscr{B}$ and m-generates $\mathscr{B}$, so if i is the identity map of $\mathscr{A}$ into $\mathscr{B}$, $\{i, \mathscr{B}\}$ is an m-completion of $\mathscr{A}$.

Let

$$B = \mathscr{D}_m - D .$$

Suppose

$$B = \bigcup_{u \in U} \bigcap_{v \in V} A_{u,v} ,$$

$\{A_{u,v}\}_{u \in U, v \in V}$ an m-indexed set in $\mathscr{A}$. This can be written in the form

$$\bigcup_{u \in U} \bigcap_{v \in V} \bigcup_{m \in M_{u,v}} A_{u,v,m} ;$$
$$A_{u,v,m} \text{ or } -A_{u,v,m} \in \mathscr{S}, \quad \overline{\overline{M_{u,v}}} < \sigma .$$

Let $B' = \{d \in \mathscr{D}_m : \{d\} = A_{u,v,m}$ for some $u \in U$, $v \in V$, and $m \in M_{u,v}\}$. Then $\overline{\overline{B'}} \leqq m$, so if

$M_{u,v}' = \{m \in M_{u,v} : A_{u,v,m}$ is not of the form $\{d\}$, $d \in \mathscr{D}_m\}$, it follows that

$$\overline{\overline{B - \bigcup_{u \in U} \bigcap_{v \in V} \bigcup_{m \in M'_{u,v}} A_{u,v,m}}} \leqq m .$$

It will now be shown that in fact

$$\overline{\overline{B - \bigcup_{u \in U} \bigcap_{v \in V} \bigcup_{m \in M'_{u,v}} A_{u,v,m}}} > m ,$$

a contradiction. Hence it may be assumed that $A_{u,v,m}$ is not of the form $\{d\}$, $d \in \mathscr{D}_m$, for all $u \in U$, $v \in V$, and $m \in M_{u,v}$.

If $A_{u,v,m} = -\{d\}$, $d \in \mathscr{D}_m$, for some $m \in M_{u,v}$, then either

$$(1) \qquad \bigcup_{m \in M_{u,v}} A_{u,v,m} = -\{d\}$$

or

$$(2) \qquad \bigcup_{m \in M_{u,v}} A_{u,v,m} = \bigvee .$$

If (1) occurs, it may be assumed that $M_{u,v} = \{1\}$ and $A_{u,v,1} = -\{d\}$. If (2) occurs, the term $\bigcup_{m \in M_{u,v}} A_{u,v,m}$ may be dropped. Thus for all $u \in U$, V may be written as $V_u \cup V'_u$, where (1) $V_u \cap V'_u = \varnothing$; (2) $A_{u,v,m} = -\{d_{u,v}\}$, $d_{u,v} \in \mathscr{D}_m$, for all $v \in V_u$; and (3) $A_{u,v,m}$ is either of the form $-D_{t,s}$ or $D_{t,s}$ for all $v \in V'_u$. Consequently, for all $u \in U$,

$$\bigcap_{v \in V} \bigcup_{m \in M_{u,v}} A_{u,v,m} = \bigcap_{v \in V_u} - \{d_{u,v}\} \cap \bigcap_{v \in V'_u} \bigcup_{m \in M_{u\ v}} A_{u,v,m} .$$

Let

$$C_u = \bigcap_{v \in V} \bigcup_{m \in M_{u,v}} A_{u,v,m} .$$

Suppose U is the set of all ordinals $u < \alpha$, where $\alpha = \overline{\overline{U}}$. Let $D_1 = \{d \in \mathscr{D}_m \colon \pi_t(d) = s'_1, s'_2\}$. Now $\overline{\overline{D}}_1 = 2^m$ implies there is a $d_1 \in D$ such that

$$d_1 \in \bigcap_{v \in V_1} - \{d_{1,v}\} .$$

Since $d_1 \notin B$, this implies

$$d_1 \notin \bigcap_{v \in V'_1} \bigcup_{m \in M_{1,v}} A_{1,v,m} ,$$

hence for some $v_1 \in V'_1$,

$$d_1 \notin \bigcup_{m \in M_{1,v_1}} A_{1,v_1,m} .$$

Also, $D_1 \subseteqq -D_{t,s}$ for all $t \in T$ and $s \in S'$, hence

$$A_{1,v_1,m} = D_{t_{1,m}, s_{t_{1,m}}}$$

for some $t_{1,m} \in T$ and $s_{t_{1,m}} \in S'$, for all $m \in M_{1,v_1}$. Let $T_1 = \{t_{1,m} \colon m \in M_{1,v_1}\}$

and pick $s_1 \in S'$ such that $s_1 \neq s_{t_{1,m}}$ for all $m \in M_{1,v_1}$. Define

$$\varphi(t) = s_1$$

for all $t \in T_1$. Let $B_1 = \varnothing$ and define $B_2 = \{d \in \mathscr{D}_m : \pi_t(d) = \varphi(t)$ for all $t \in T_1\}$.

Note that $B_2 \cap C_1 = \varnothing$.

Suppose $i > 1$ and a finite set $T_{i'}$ has been defined for each $i' < i$ so that $T_{i'} \cap T_{i''} = \varnothing$ if $i', i'' < i$, $i' \neq i''$; $s_{i'} \in S'$ has been chosen; φ has been defined on each $T_{i'}$, $i' < i$, so that $\varphi(t) = s_{i'}$ for all $t \in T_{i'}$; and if

$$B_i = \{d \in \mathscr{D}_m : \pi_t(d) = \varphi(t) \text{ for all } t \in \bigcup_{i'<i} T_{i'}\}$$

then

$$B_i \cap \bigcup_{i'<i} C_{i'} = \varnothing \ .$$

Let

$$\hat{T}_i = \bigcup_{i'<i} T_{i'}$$

and note that $\overline{\overline{\hat{T}_i}} < m$. Let

$$\begin{aligned} D_i = \{d \in \mathscr{D}_m &: \pi_t(d) = \varphi(t) \text{ for all } t \in \hat{T}_i \\ &\text{and } \pi_t(d) = s'_k,\ k = 1, 2, \text{ if } t \in T - \hat{T}_i\} \ . \end{aligned}$$

Then $D_i \subseteqq D$ and $\overline{\overline{D_i}} = 2^m$, hence there is a $d_i \in D_i$ such that

$$d_i \in \bigcap_{v \in V_i} - \{d_{i,v}\} \ .$$

Since $d_i \notin B$, this implies

$$d_i \notin \bigcap_{v \in V'_i} \bigcup_{m \in M_{i,v}} A_{i,v,m} \ ,$$

hence for some $v_i \in V'_i$,

$$d_i \notin \bigcup_{m \in M_{i,v_i}} A_{i,v_i,m} \ .$$

If $B_i \cap C_i = \varnothing$ set $T_i = \varnothing$. If not, there is a $d'_i \in B_i$ such that $d'_i \in C_i$, so

$$d'_i \in \bigcup_{m \in M_{i,v_i}} A_{i,v_i,m} \ .$$

Note that $\pi_t(d'_i) = \pi_t(d_i)$ for all $t \in \hat{T}_i$.

It immediately follows that if

$$d_i' \in \bigcup_{m \in M_{i,v_i}} A_{i,v_i,m}$$

then

$$A_{i,v_i,m} = D_{t_{i,m}, s_{t_{i,m}}} ,$$

where $t_{i,m} \notin \hat{T}_i$ and

$$\pi_{t_{i,m}}(d_i') = s_{t_{i,m}} ,$$

for some $m \in M_{i,v_i}$.

Let

$$T_i = \{t_{i,m} \in T - \hat{T}_i \colon A_{i,v_{i,m}} = D_{t_{i,m}, s_{t_{i,m}}} \text{ for some } m \in M_{i,v_i}\}$$

and pick $s_i \in S'$ such that if $t_{i,m} \in T_i$ then

$$s_i \neq S_{t_{i,m}} ,$$

for all $m \in M_{i,v_i}$. Now define

$$\varphi(t) = s_i \quad \text{for all} \quad t \in T_i .$$

Thus $T_i \cap \hat{T}_i = \varnothing$ which implies $T_i \cap T_{i'} = \varnothing$ for all $i' < i$. If

$$B_{i+1} = \{d \in \mathscr{D}_m \colon \pi_t(d) = \varphi(t) \text{ for all } t \in T_i \cup \hat{T}_i\}$$

then it is clear that

$$B_{i+1} \cap \bigcup_{i' < i} C_i = \varnothing .$$

Now let $\hat{T} = \bigcup_{i<\alpha} T_i$ and set

$$\begin{aligned} \hat{B} = \{d \in \mathscr{D}_m \colon \pi_t(d) = \varphi(t) \text{ for all } t \in \hat{T} \\ \text{and } \pi_t(d) \neq s_1', s_2' \text{ if } t \in T - \hat{T}\} . \end{aligned}$$

Then $\hat{B} \neq \varnothing$ and $\hat{B} \subseteqq B$. But $\hat{B} \cap \bigcup_{u \in U} C_u = \varnothing$ which implies

$$B - \bigcup_{u \in U} C_u \neq \varnothing .$$

If $B' = B - \bigcup_{u \in U} C_u$ then for each $b \in B'$,

$$b = \bigcap_{t \in T} D_{t, s_{t,b}} ,$$

for some m-indexed set $\{s_{t,b}\}_{t \in T}$ in S'. Thus

$$B = \bigcup_{u \in U} \bigcap_{v \in V} \bigcup_{m \in M_{u,v}} A_{u,v,m} \cup \bigcup_{b \in B'} \bigcap_{t \in T} D_{t, s_{t,b}} ,$$

but the above construction shows that

$$B - (\bigcup_{u \in U} \bigcap_{v \in V} \bigcup_{m \in M_{u,v}} A_{u,v,m} \cup \bigcup_{b \in B'} \bigcap_{t \in T} D_{t,s_{t,b}}) \neq \emptyset$$

if $\bar{\bar{B}}' \leqq m$. Hence

$$\overline{\overline{B - \bigcup_{u \in U} C_u}} > m .$$

LEMMA 3.6. *If $\{i, \mathscr{B}\}$ is an m-completion of the Boolean algebra $\mathscr{A}$ and there is a $B \in \mathscr{B}$ such that*

$$B \neq \overset{\mathscr{B}}{\bigcup_{t \in T}} \overset{\mathscr{B}}{\bigcap_{s \in S}} i(A_{t,s})$$

for all m-indexed sets $\{A_{t,s}\}_{t \in T, s \in S}$ in $\mathscr{A}$, then there is an m-ideal Δ in $\mathscr{B}$ such that $\{j, \mathscr{B}_\Delta\}$ is an m-extension of $i_\Delta(\mathscr{A})$ but not an m-completion, where $i_\Delta(A) = [i(A)]_\Delta$ for all $A \in \mathscr{A}$, $\mathscr{B}_\Delta = \mathscr{B}/\Delta$ and j is the identity map of $i_\Delta(\mathscr{A})$ into $\mathscr{B}_\Delta$.

Proof. Let

$$\Delta' = \{B' \in \mathscr{B} : B' \subseteqq B \text{ and } B' = \overset{\mathscr{B}}{\bigcap_{t \in T}} i(A_t) ,$$
$$\text{for some } m\text{-indexed set } \{A_t\}_{t \in T} \text{ in } \mathscr{A}\}$$

and let $\Delta = \langle \Delta' \rangle_m$. Then if $\delta \in \Delta$, $\delta \subseteqq B$, so $B \notin \Delta$. If $A \in \mathscr{A}$ and $[i(A)]_\Delta \subseteqq [B]_\Delta$ then $i(A) - B \in \Delta$ so $i(A) - B \subseteqq B$ which implies $i(A) \subseteqq B$, hence $i(A) \in \Delta$ and $[i(A)]_\Delta = \Lambda_{\mathscr{B}_\Delta}$, implying $i_\Delta(\mathscr{A})$ is not dense in $\mathscr{B}_\Delta$.

It only remains to show that $i_\Delta(\mathscr{A})$ is m-regular in $\mathscr{B}_\Delta$. If

$$\overset{i_\Delta(\mathscr{A})}{\bigcap_{t \in T}} [i(A_t)]_\Delta = \Lambda_{\mathscr{B}_\Delta}$$

then $i(A) \subseteqq i(A_t)$ for all $t \in T$ implies $i(A) \in \Delta$, so $i(A) \subseteqq B$. If

$$\overset{\mathscr{B}}{\bigcap_{t \in T}} i(A_t) \nsubseteqq B ,$$

then there is an $A \neq \Lambda_{\mathscr{A}}$ in $\mathscr{A}$ such that

$$i(A) \subseteqq \overset{\mathscr{B}}{\bigcap_{t \in T}} i(A_t) - B ,$$

contradicting the above statement. Thus

$$\overset{\mathscr{B}}{\bigcap_{t \in T}} i(A_t) \subseteqq B$$

so

$$\overset{\mathscr{B}}{\bigcap_{t \in T}} i(A_t) \in \Delta$$

and

$$\bigwedge_{\mathscr{B}_\Delta} = [\bigcap_{t\in T}^{\mathscr{B}} i(A_t)]_\Delta = \bigcap_{t\in T}^{\mathscr{B}_\Delta} [i(A_t)]_\Delta .$$

Thus if $\mathscr{A}$ is the Boolean algebra constructed in Lemua 3.5, $i_\Delta(\mathscr{A})$ is a Boolean algebra such that $\mathscr{K}(i_\Delta(\mathscr{A}))$ contains more than one element. Hence it is justified to assume that for each infinite cardinal m there is a Boolean algebra $\mathscr{A}$ such that $\mathscr{A}$ has an m-extension which is not an m-completion.

4. Let $\{\mathscr{A}_t\}_{t\in T}$ be a (fixed) indexed set of Boolean algebras. Let h_t be an isomorphism of $\mathscr{A}_t$ onto the field $\mathscr{F}_t$ of all open-closed subsets of the Stone space X_t of $\mathscr{A}_t$. Let X denote the Cartesian product of all the spaces X_t. Let π_t be the projection of X onto $\mathscr{F}_t$ and define

$$\varphi_t: \mathscr{F}_t \longrightarrow X$$

by:

$$\text{if } F\in\mathscr{F}_t \text{ then } \varphi_t(F) = \{x\in X: \pi_t(x)\in F\} .$$

Let $\mathscr{F}$ be the Boolean product of $\{\mathscr{A}_t\}_{t\in T}$. Define $h_t^* = \varphi_t h_t$ and let $\mathscr{S}$ be the set of all sets $\bigcap_{t\in T'} h_t^*(A_t)$; $A_t\in\mathscr{A}_t$, $T'\subseteqq T'$, $\overline{\overline{T'}}\leqq n$. Define $\hat{\mathscr{F}}$ to be the field of sets generated by $\mathscr{S}$. Let J be the set of all sets $S\subseteqq\hat{\mathscr{F}}$ such that

1. $\overline{\overline{S}}\leqq m$;
2. there is a $t\in T$ such that $S\subseteqq h_t^*(\mathscr{A}_t)$;
3. the join $\bigcup_{A\in S}^{\hat{\mathscr{F}}} A$ exists.

Let M' be the set of all sets $S\subseteqq\hat{T}$ such that

1. $\overline{\overline{S}}\leqq m$;
2. there is a $t\in T$ such that $S\subseteqq h_t^*(\mathscr{A}_t)$;
3. the meet $\bigcap_{A\in S}^{\hat{\mathscr{F}}} A$ exists.

Let M'' be the set of all sets $S\subseteqq\hat{T}$ such that

1. $\overline{\overline{S}}\leqq n$;
2. if $A\in S$ then $A\in h_t^*(\mathscr{A}_t)$ for some $t\in T$;
3. if $A, B\in S$, $A\neq B$, then $A\in h_t^*(\mathscr{A}_t)$ implies $B\notin h_t^*(\mathscr{A}_t)$. Let $M = M'\cup M''$.

The following lemma is due to La Grange [1] and will be given without proof.

LEMMA 4.1. *If* $\{\{i_t\}_{t\in T}, \mathscr{B}\}\in\mathscr{P}_n$ *then there is one and only one* (J, M, m)*-isomorphism* h *mapping* $\hat{\mathscr{F}}$ *into* $\mathscr{B}$ *such that*

$$hh_t^* = i_t \quad \textit{for all} \quad t\in T .$$

THEOREM 4.1. *If* $\{\{i_t\}_{t\in T}, \mathscr{B}\} \in \mathscr{P}_n$ *then there is a mapping* h *of* $\hat{\mathscr{F}}$ *into* $\mathscr{B}$ *such that* $\{h, \mathscr{B}\}$ *is a* (J, M, m)*-extension of* $\hat{\mathscr{F}}$*. If* $\{h, \mathscr{B}\}$ *is a* (J, M, m)*-extension of* $\hat{\mathscr{F}}$ *then the ordered pair* $\{\{hh_t^*\}_{t\in T}, \mathscr{B}\} \in \mathscr{P}_n$.

Proof. Let h be the (J, M, m)-isomorphism from $\hat{\mathscr{F}}$ into $\mathscr{B}$ such that $hh_t^* = i_t$ for all $t \in T$. Then $\{h, \mathscr{B}\}$ is a (J, M, m)-extension of $\hat{\mathscr{F}}$.

Conversely, if $\{h, \mathscr{B}\}$ is a (J, M, m)-extension of $\hat{\mathscr{F}}$, it follows immediately that $\{\{hh_t^*\}_{t\in T'}, \mathscr{B}\}$ is an (m, n)-product of $\{\mathscr{A}_t\}_{t\in T}$.

THEOREM 4.2. *If* $\{\{i_t\}_{t\in T}, \mathscr{B}\}$, $\{\{i'_t\}_{t\in T}, \mathscr{B}'\}$ *are two* (m, n)*-products of* $\{\mathscr{A}_t\}_{t\in T}$ *then*

$$\{\{i_t\}_{t\in T}, \mathscr{B}\} \leqq \{\{i'_t\}_{t\in T}, \mathscr{B}'\}$$

if, and only if,

$$\{i, \mathscr{B}\} \leqq \{i', \mathscr{B}'\}$$

where $\{i, \mathscr{B}\}$ *and* $\{i', \mathscr{B}'\}$ *are the* (J, M, m)*-extensions of* $\hat{\mathscr{F}}$ *induced by the* (J, M, m)*-isomorphisms* i' *and* i *of* $\hat{\mathscr{F}}$ *into* $\mathscr{B}'$ *and* $\mathscr{B}$*, respectively, given by Lemma* 4.1.

Proof. Now

$$\{\{i_t\}_{t\in T}, \mathscr{B}\} \leqq \{\{i'_t\}_{t\in T}, \mathscr{B}'\}$$

if, and only if, there is an m-homomorphism h such that

$$h: \mathscr{B}' \longrightarrow \mathscr{B}$$

and $hi'_t = i_t$ for all $t \in T$. Similarly,

$$\{i, \mathscr{B}\} \leqq \{i', \mathscr{B}'\}$$

if, and only if, there is an m-homomorphism

$$h: \mathscr{B}' \longrightarrow \mathscr{B}$$

such that $h'i' = i$. Thus it suffices to show that $hi' = i$, if, and only if, $hi'_t = i_t$. Let h_t^* be defined as above. Then $ih_t^* = i_t$ and $i'h_t^* = i'_t$, so if $hi' = i$,

$$hi'_t = hi'h_t^* = ih_t^* = i_t ,$$

and if $hi'_t = i_t$, then

$$hi' = hi'_t h_t^{*-1} = i_t h_t^{*-1} = i .$$

La Grange [1] has given an example of an $(m, 0)$-product for which $\mathscr{P}$ does not contain a smallest element and an example of an (m, n)-product for which $\mathscr{P}_n$ does not contain a smallest element. Theorem 4.2 extends this result by showing that the question whether $\mathscr{P}$ or $\mathscr{P}_n$ contains a smallest element reduces to asking whether the class of all (J, M, m)-extensions of $\mathscr{A}_0$ or $\hat{\mathscr{F}}$ contains a smallest element for J and M defined appropriately in each case, where $\mathscr{A}_0$ and $\hat{\mathscr{F}}$ are defined as above. Now the class of all (J, M, m)-extensions of $\mathscr{A}_0$ contains a smallest element only if the class of all m-extensions of $\mathscr{A}$ contains a smallest element and Theorem 3.2 shows that the class of all m-extensions of $\mathscr{A}_0$ need not contain a smallest element, which implies the same is true for $\mathscr{P}$. Since Theorem 3.2 may be extended to Boolean algebras of the form $\hat{\mathscr{F}}$, it follows that $\mathscr{P}_n$ need not contain a smallest element.

REFERENCES

1. R. H. La Grange, *On (m, n)-products of Boolean algebras*, Pacific J. Math., **31** (1968), 725-731.
2. R. Sikorski, *Boolean Algebras* (2nd Edition), New York, 1964.

Received November 27, 1972 and in revised form March 28, 1973.

UNIVERSITY OF CALIFORNIA, LOS ANGELES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

MULTIPLICATIVITY-PRESERVING ARITHMETIC POWER SERIES

DAVID REARICK

In the Dirichlet algebra of arithmetic functions let the operator A be represented by an arithmetic power series $Af = \Sigma a(F) f^F$. A condition on the coefficients $a(F)$ is derived which is necessary and sufficient for Af to be multiplicative whenever f is multiplicative.

1. Introduction.

In [2] a *factorization* F was defined to be a nonnegative integer-valued arithmetic function having $F(1) = 0$ and $F(n) \neq 0$ for at most finitely n. The *index* of F was defined by $i(F) = \prod_{j=1}^{\infty} j^{F(j)}$. If f is any arithmetic function, we defined $f^F = \prod_{j=1}^{\infty} [f(j)]^{F(j)}$ with the understanding that $0^0 = 1$. If $a(F)$ is a mapping from factorizations into the real or complex numbers, we wrote

$$Af = \sum a(F) f^F \tag{1}$$

as an abbreviation for the arithmetic function Af whose value on n is equal to $\sum_{i(F)=n} a(F) f^F$. In [2] a series of the form (1) was called an *arithmetic power series*. Since for each n the series is terminating, there is never any question of convergence. Such a series defines an operator A on the Dirichlet algebra of arithmetic functions, and the theory of these operators has been investigated in [1] and [2].

In particular, if r is a real number, the Dirichlet rth power of an arithmetic function f is represented, when $f(1) = 1$, by an arithmetic power series $\sum \binom{r}{F} f^F$. The symbol $\binom{r}{F}$ was defined in [2]. It is known [1, Theorem 5] that f^r is multiplicative whenever f is, and therefore the series $\sum \binom{r}{F} f^F$ is an example of a multiplicativity-preserving arithmetic power series. The present paper is devoted to determining a necessary and sufficient condition on the coefficients $a(F)$ in order that the general series (1) preserve multiplicativity. The method, and the statement of the result (Theorem 1), depend on a certain equivalence relation between factorizations, to be introduced below.

2. Equivalent factorizations.

DEFINITION 1. If F and F' are two factorizations, we say F is *equivalent* to F', written $F \sim F'$, if $f^F = f^{F'}$ for every multiplicative arithmetic function f.

It is obvious that this is an equivalence relation. An example of a pair of nonequal but equivalent factorizations may be constructed by taking $F(2) = F(3) = F'(6) = 1$, with all other values being zero. Then $f^F = f(2)f(3) = f(6) = f^{F'}$ for every multiplicative f. Two equivalent factorizations F and F' necessarily have the same index, for if we choose the particular multiplicative function $f(n) = n$, we have $i(F) = f^F = f^{F'} = i(F')$.

DEFINITION 2. We shall use the letter C to denote an equivalence class of factorizations. The *index* $i(C)$ of an equivalence class C is defined to be the index of the factorizations F belonging to C. If f is multiplicative, we denote by f^C the common value of f^F for all $F \in C$. If $F_1 \in C_1$ and $F_2 \in C_2$, we define $C_1 + C_2$ to be the equivalence class containing the factorization $F_1 + F_2$.

It is obvious that the definition of $C_1 + C_2$ is unambiguous.

If the operator (1) is applied to a multiplicative f, the sum over all factorizations F of index n reduces to a sum over all classes C of index n, thus:

$$Af(n) = \sum_{i(F)=n} a(F)f^F = \sum_{i(C)=n} f^C \sum_{F \in C} a(F) .$$

Therefore, insofar as its action on multiplicative functions is concerned, an arithmetic power series is determined by the *sums* of its coefficients over equivalence classes of factorizations, and it is natural to make the following definition:

DEFINITION 3. $a^*(C) = \sum_{F \in C} a(F)$.

Thus, when f is multiplicative, we may write

$$Af(n) = \sum_{i(C)=n} a^*(C)f^C . \tag{2}$$

The main theorem may now be stated as follows.

THEOREM 1. *The arithmetic function $Af = \sum a(F)f^F$ is multiplicative whenever f is, if and only if the following pair of conditions holds*:

$$a^*(C_1 + C_2) = a^*(C_1)a^*(C_2) \tag{3}$$

for every pair of equivalence classes C_1 *and* C_2 *having relatively prime indices, and*

$$a^*(0) = 1 \tag{4}$$

where 0 *is the class containing the zero factorization.*

3. Lemmas. Let those positive integers which are prime powers be arranged in increasing order. Let $x_1, x_2, \cdots$ be an arbitrary sequence of complex numbers. We may construct a multiplicative function f by setting $f(1) = 1$ and, it p^ν is the kth prime power, defining

$$f(p^\nu) = x_k \, . \tag{5}$$

The requirement that f be multiplicative then defines $f(n)$ for all positive integers n. Furthermore, every multiplicative f arises from exactly one particular choice of the sequence $\{x_k\}$. (Following the usual convention, we do not consider the identically zero function to be multiplicative.)

These observations establish a one-to-one correspondence between the set of all multiplicative functions and the set of all sequences of variables $\{x_k\}$. Under this correspondence we may associate, with each factorization F, an expression f^F which is a monomial (with coefficient 1) in certain of the variables x_k. We note that a given variable x_k cannot appear in this monomial if it does not correspond, in (5), to a prime power divisor of $i(F)$, since, by definition of index $F(j) = 0$ if j does not divide $i(F)$.

LEMMA 1. *Two factorizations* F *and* F' *are equivalent if and only if the two corresponding monomials* f^F *and* $f^{F'}$ *are identical.*

Proof. It is familiar from algebra [3, Chapter 4] that if two polynomials always agree in value while each variable x_k is assigned infinitely many different values, holding the others fixed, then the two polynomials are identical. The converse part of the assertion is trivial.

Lemma 1 shows that equivalence classes of factorizations may be identified with monomials in an arbitrary finite number of variables. Also, it is clear that each equivalence class of prime power index p^ν consists of a single factorization.

LEMMA 2. *Let* $F_1, \cdots, F_r$ *be nonequivalent factorizations. Suppose that, for every multiplicative* f, *the linear combination* $\sum_{j=1}^r b_j f^{F_j}$ *is equal to zero. Then each of the coefficients* b_j *is zero.*

Proof. The linear combination referred to in the lemma is a polynomial in certain of the variables x_k, and the numbers b_j are precisely its coefficients, since by Lemma 1 no two of the monomials f^{F_j} are identical. As in the proof of Lemma 1, each of these coefficients must be zero.

LEMMA 3. *Let F, F', G, and G' be factorizations, with $i(F) = i(F') = m$ and $i(G) = i(G') = n$, and assume m and n are relatively prime. Suppose $F + G \sim F' + G'$. Then $F \sim F'$ and $G \sim G'$.*

Proof. As observed earlier, each variable x_k appearing in the monomial f^F corresponds, in (5), to a prime power divisor of m. Similarly, f^G contains only variables corresponding to prime power divisors of n. Since $(m, n) = 1$, these two sets of variables are disjoint. Applying the same reasoning to F' and G', we see that no variable appearing in either f^F or $f^{F'}$ can appear in either f^G or $f^{G'}$, and conversely. By hypothesis we have $f^F f^G = f^{F+G} = f^{F'+G'} = f^{F'} f^{G'}$ for all multiplicative f, or equivalently $f^F/f^{F'} = f^{G'}/f^G$. Since opposite sides of this identity are rational functions in disjoint sets of independent variables, both sides must be equal to a constant B. In the identity $f^F = Bf^{F'}$, putting $f(k) = 1$ for all k, we obtain $B = 1$. Therefore $f^F = f^{F'}$ and $f^G = f^{G'}$, meaning $F \sim F'$ and $G \sim G'$.

LEMMA 4. *Let $F_1, \cdots, F_r$ be nonequivalent factorizations of index m. Let $G_1, \cdots, G_s$ be nonequivalent factorizations of index n. Assume $(m, n) = 1$. Suppose that, for every multiplicative f, the linear combination $\sum_{j=1}^{r} \sum_{k=1}^{s} b_{jk} f^{F_j + G_k}$ is equal to zero. Then each of the coefficients b_{jk} is zero.*

Proof. By Lemma 3 the factorizations $F_j + G_k$ are all nonequivalent, and the result then follows from Lemma 2.

LEMMA 5. *Let F be a factorization of index mn, where $(m, n) = 1$. Then there exist factorizations F_1 and F_2, of indices m and n respectively, such that $F \sim F_1 + F_2$. Furthermore, if F_1' and F_2' also satisfy these conditions, then $F_1 \sim F_1'$ and $F_2 \sim F_2'$. In other words, if $(m, n) = 1$, then each equivalence class of index mn is the sum of a unique pair of classes of indices m and n respectively.*

Proof. The uniqueness part follows immediately from Lemma 3. As regards the existence of F_1 and F_2, we claim that the pair defined as follows will satisfy the requirements:

$$\begin{aligned} F_1(k) &= 0 && \text{if } k = 1 \\ &= \sum_{(j,m)=k} F(j) && \text{if } k > 1 \\ F_2(k) &= 0 && \text{if } k = 1 \\ &= \sum_{(j,n)=k} F(j) && \text{if } k > 1 \,. \end{aligned}$$

To check this, choose any multiplicative f. Then

$$\begin{aligned} f^{F_1+F_2} = f^{F_1}f^{F_2} &= \prod_{k=1}^{\infty} [f(k)]^{F_1(k)} \prod_{k=1}^{\infty} [f(k)]^{F_2(k)} \\ &= \prod_{j=1}^{\infty} [f((j,\, m))]^{F(j)} \prod_{j=1}^{\infty} [f((j,\, n))]^{F(j)} \\ &= \prod_{j=1}^{\infty} [f((j,\, m))f((j,\, n))]^{F(j)} \\ &= \prod_{j=1}^{\infty} [f((j,\, m)(j,\, n))]^{F(j)} \\ &= \prod_{j=1}^{\infty} [f((j,\, mn))]^{F(j)} = \prod_{j=1}^{\infty} [f(j)]^{F(j)} = f^F \,, \end{aligned}$$

where in the last step we use the fact that $F(j) = 0$ if j does not divide mn. Therefore $F \sim F_1 + F_2$. To find the indices of F_1 and F_2, we first observe that $i(F_1)i(F_2) = i(F_1 + F_2) = i(F) = mn$. Also, if we choose for f the identity function $f(k) = k$, we have $i(F_1) = f^{F_1} = \prod_{j=1}^{\infty} (j,\, m)^{F(j)}$, and each factor in the product is relatively prime to n, so $i(F_1)$ is relatively prime to n. Similarly, $i(F_2)$ is relatively prime to m. Therefore $i(F_1) = m$ and $i(F_2) = n$.

4. **Proof of Theorem 1.** First assume conditions (3) and (4) hold. Choose any multiplicative f, and let m and n be relatively prime. We are to show that $Af(mn) = Af(m)Af(n)$ and $Af(1) = 1$. By Lemma 5, each equivalence class C of index mn is the sum of a unique pair of classes $C_1 + C_2$ where $i(C_1) = m$ and $i(C_2) = n$. Remembering (2), we may evaluate $Af(mn)$ as follows:

$$\begin{aligned} Af(mn) &= \sum_{i(C)=mn} a^*(C)f^C = \sum_{i(C_1)=m} \sum_{i(C_2)=n} a^*(C_1 + C_2)f^{C_1+C_2} \\ &= \sum_{i(C_1)=m} a^*(C_1)f^{C_1} \sum_{i(C_2)=n} a^*(C_2)f^{C_2} = Af(m)Af(n) \,. \end{aligned}$$

Also, $Af(1) = a^*(0)f0 = 1$.

To prove the converse, assume the operator A preserves multiplicativity. Choose m and n relatively prime, and let f be any multiplicative function. Proceeding as in the last computation above, we have

$$\begin{aligned} 0 &= Af(mn) - Af(m)Af(n) \\ &= \sum_{i(C_1)=m} \sum_{i(C_2)=n} f^{C_1+C_2}[a^*(C_1 + C_2) - a^*(C_1)a^*(C_2)] \,. \end{aligned}$$

This double sum is a linear combination of the type considered in Lemma 4, and therefore, by the result of that lemma, the expression in square brackets is equal to zero for all C_1 and C_2 in the sum. That is, equation (3) is satisfied. Also, (4) is satisfied because $1 = Af(1) = a^*(0)f0 = a^*(0)$. This completes the proof of Theorem 1.

5. **Further consequences.** We wish to show how to construct all solutions $a^*(C)$ of (3) which also satisfy (4) (and which we shall refer to as *nontrivial* solutions of (3)). Given a nontrivial solution $a^*(C)$ of (3), we can recover (nonuniquely) by Definition 3 the coefficients $a(F)$ of an arithmetic power series (1) which preserves multiplicativity, and the class of such series will then be completely characterized.

LEMMA 6. *Let C be an equivalence class whose index is greater than 1 and has prime factorization $i(C) = p_1^{\nu_1}, \cdots, p_r^{\nu_r}$. Then there are unique classes $C_1, \cdots, C_r$, of indices $p_1^{\nu_1}, \cdots, p_r^{\nu_r}$ respectively, such that $C = C_1 + \cdots + C_r$.*

Proof. Apply Lemma 5 repeatedly to the r maximal prime power divisors $p_1^{\nu_1}, \cdots, p_r^{\nu_r}$ of $i(C)$.

LEMMA 7. *$a^*(C)$ is a nontrivial solution of* (3) *if and only if $a^*(0) = 1$ and*

$$a^*(C) = \prod_{k=1}^{r} a^*(C_k) \tag{6}$$

whenever $i(C) > 1$, where the classes $C_1, \cdots, C_r$ are related to C as in Lemma 6.

Proof. Equation (6) is obtained from (3) by applying the latter repeatedly to the maximal prime power divisors of $i(C)$. Conversely, (3) is obtained from (6) by applying (6) to the prime decomposition of mn, separating the maximal prime power divisors of m from those of n.

Lemma 7 gives us a process for constructing all nontrivial solutions of (3). The method is analogous to that used at the beginning of § 3 to construct all multiplicative functions, namely:

THEOREM 2. *The nontrivial solutions $a^*(C)$ of* (3) *are exactly those which take the value* 1 *on the zero class and are defined arbitrarily on classes of prime power index, the definition then being extended to all C by the product formula* (6).

Finally, we shall determine the number of equivalence classes of index n. Let this number be denoted by $E(n)$. It follows from Lemma 5 that $E(n)$, as an arithmetic function, is multiplicative. Therefore, it suffices to evaluate this function on prime powers p^ν. Since each class of index p^ν contains only one factorization, $E(p^\nu)$ is equal to the number of factorizations of index p^ν, and this is evidently just the number of unrestricted partitions of ν. These observations yield the following explicit formula for $E(n)$:

THEOREM 3.

$$E(1) = 1$$
$$E(n) = \prod_{p^\nu || n} p(\nu) \quad \textit{if } n > 1 ,$$

where $p(\nu)$ *is the partition function, and the product is extended over all maximal prime power divisors* p^ν *of* n.

REFERENCES

1. D. Rearick, *Operators on algebras of arithmetic functions*, Duke Math J., **35** (1968), 761-766.
2. ———, *The trigonometry of numbers*, Duke Math. J., **35** (1968), 767-776.
3. B. L. Van der Waerden, *Modern Algebra*, Vol. I, 2nd ed., New York, 1953.

Received March 20, 1973.

UNIVERSITY OF COLORADO

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

CHARACTERISTIC IDEALS IN GROUP ALGEBRAS

I. SINHA

If $\mathfrak{F}G$ is the group-algebra of a group G over a field $\mathfrak{F}$, and $\mathfrak{A}$ is any subgroup of the automorphism group of the $\mathfrak{F}$-algebra $\mathfrak{F}G$, then an ideal I of $\mathfrak{F}G$, is called $\mathfrak{A}$-characteristic if $I^\alpha \subseteqq I$, $\forall^\alpha \in A$. If A is the whole automorphism group itself, then we merely say that I is characteristic. Then D.S. Passman has proved the following result:

"Let $H \trianglelefteq G$ such that G/H is $\mathfrak{F}$-complete. Then for each characteristic ideal I of $\mathfrak{F}G$, $I = (I \cap \mathfrak{F}H)\mathfrak{F}G$."

The main concern in this paper is to consider the converse of this result.

2. Some preliminaries. For a given ideal $I \trianglelefteq \mathfrak{F}G$, let $\mathscr{R}(I)$ be the set of all $H \leqq G$ such that $I = (I \cap \mathfrak{F}H)\mathfrak{F}G$. Let $C(I)$ be the set of all H in G such that if for some right $\mathfrak{F}H$-module $\mathfrak{M}$, $I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$, then $I \subseteqq \operatorname{Ann} \mathfrak{M}^G$, the induced $\mathfrak{F}G$-module. We first of all have:

THEOREM 1. (i) *For any $I \trianglelefteq \mathfrak{F}G$, $C(I) \subseteqq \mathscr{R}(I)$.*
(ii) *If $H \trianglelefteq G$, then $H \in \mathscr{R}(I)$ if and only if $H \in C(I)$.*

Proof. (i) Let $I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$ imply that $I \subseteqq \operatorname{Ann} \mathfrak{M}^G$. Let $\sum p_i x_i \in I$ with $p_i \in \mathfrak{F}H$, where $G = \cup Hx_i$ is a coset-decomposition. We have $(\sum \mathfrak{M} \otimes x_i)(\sum p_i x_i) = 0$ if $I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$. In particular $(m \otimes I)(\sum p_i x_i) = 0$, $\forall m \in \mathfrak{M}$, i.e., $\sum m p_i \otimes x_i = 0$, $\forall m \in \mathfrak{M}$. So $\mathfrak{M} \cdot p_i = 0$ for each i. Thus $p_i \in \operatorname{Ann} \mathfrak{M}$. Since $\mathfrak{M}$ is arbitrary with the property that $I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$, so we may take $\mathfrak{M} = \mathfrak{F}H/I \cap \mathfrak{F}H$, and conclude that each $p_i \in \operatorname{Ann} \mathfrak{M} = I \cap \mathfrak{F}H$. Thus $\sum p_i x_i \in (I \cap \mathfrak{F}H)\mathfrak{F}G$.

(ii) Suppose $I = \mathfrak{F}G(I \cap \mathfrak{F}H)$ and $I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$, for some $\mathfrak{F}H$-module $\mathfrak{M}$. Note that $H \trianglelefteq G$ implies that $\mathfrak{F}G(I \cap \mathfrak{F}H) = (I \cap \mathfrak{F}H)\mathfrak{F}G$. Let $a = \sum x_i p_i \in I$ where $p_i \in I \cap \mathfrak{F}H$. So $a\mathfrak{M}^G = (\sum x_i p_i)(\sum x_j \otimes \mathfrak{M}) = \sum x_i x_j \otimes p_i^{x_j} \mathfrak{M} = 0$ since $p_i^{x_j} \in I \cap \mathfrak{F}H \subseteqq \operatorname{Ann} \mathfrak{M}$. Thus $a\, \mathfrak{M}^G = 0$ and $I \subseteqq \operatorname{Ann} \mathfrak{M}^G$.

Theorem 17.4 of [1] then gives us:

COROLLARY 1. *Let $H \trianglelefteq G$ such that G/H is $\mathfrak{F}$-complete. Then $H \in C(I)$ for every characteristic ideal I of $\mathfrak{F}G$.*

Also Theorem 17. 7 of [1] implies:

COROLLARY 2. *If $H \trianglelefteq G \ni G/H$ is abelian and has no elements of order $p =$ Char. $\mathscr{F}$, then $H \in C(J(G))$, where J denotes the*

Jacobson-radical of $\mathfrak{F}G$.

3. Main result. We will prove:

THEOREM 2. *For* $I = [\mathfrak{F}G, \mathfrak{F}G]$, *the commutator ideal and for* $J = J(G)$, *if* $H \leqq G$ *such that* $H \in \mathscr{R}(I)$ *and* $H \in \mathscr{R}(J)$ *then* $H \trianglelefteq G$, G/H *is abelian with no elements of order* p. *In particular,* $\mathfrak{F}(G/H)$ *is semi-simple.*

Further, if $\mathfrak{F}$ *is algebraically closed then* G/H *is* $\mathfrak{F}$*-complete.*

We observe that the last two statements in the theorem follows from 17.8 and 17.1 (i) respectively of [1]. The rest of the theorem will be proved by a series of results proved below.

LEMMA 1. *Let* $H \leqq G$, $I \trianglelefteq \mathfrak{F}G$ *and* $H \in \mathscr{R}(I)$. *Then* $H \supseteqq \mathfrak{A}^{-1}(I) = \{g \in G \mid g - 1 \in I\}$.

Proof. Let $G = \cup Hx_i$ be a coset-decomposition, and $g \in \mathfrak{A}^{-1}(I)$ such that $g \notin H$. Then $g = hx_i$ for some i, where $x_i \neq 1$, and $h \in H$; and $hx_i - 1 \in (I \cap \mathfrak{F}H)\mathfrak{F}G = \sum (I \cap \mathfrak{F}H)x_i$. Since $\{x_i\}$ are linearly independent over $\mathfrak{F}H$, $h \in I \cap \mathfrak{F}H$, and $x_i \neq 1$, so $g \in I$ which implies that $1 \in I$, a contradiction.

LEMMA 2. *If* $I = [\mathfrak{F}G, \mathfrak{F}G]$, *and* $H \in \mathscr{R}(I)$ *then* $H \trianglelefteq G$ *and* G/H *is abelian.*

Proof. Observe that I is a proper ideal in $\mathfrak{F}G$, since $\mathfrak{A}(I) = 0$. Also by Lemma 1, $H \supseteqq \mathfrak{A}^{-1}(I)$. Since $(ghg^{-1}h^{-1} - 1)hg = gh - hg \in I$, for all $g, h \in G$, so $(ghg^{-1}h^{-1} - 1) \in I$. Hence $ghg^{-1}h^{-1} \in \mathfrak{A}^{-1}(I) \subseteqq H$; i.e., G', the commutator-subgroup is in H. Hence $H \trianglelefteq G$ and G/H is abelian.

Now let H satisfy the hypothesis of Lemma 2. Then we have:

LEMMA 3. *Let* $I = J(G)$ *and* $H \in \mathscr{R}(I)$. *Then* $\mathfrak{F}(G/H)$ *is semi-simple and* G/H *has no elements of order* $p =$ Char. $\mathfrak{F}$.

Proof. $J(G) = (J(G) \cap \mathfrak{F}H)\mathfrak{F}G \subseteqq J(H) \cdot \mathfrak{F}G$ by 16.9 of [1]. Now $\mathfrak{F}H[\mathfrak{A}_H(H) \cong \mathfrak{F}$ where $\mathfrak{A}_H(H)$ is the ideal of $\mathfrak{F}H$, generated by $\{h - 1 \mid h \in H\}$. So $\mathfrak{A}_H(H) \supseteqq J(H)$. Hence $\mathfrak{A}_H(H)\mathfrak{F}G = \mathfrak{A}_G(H) \supseteqq J(H) \cdot \mathfrak{F}G \supseteqq J(G)$, where $\mathfrak{A}_G(H)$ is the ideal in $\mathfrak{F}G$, generated by $\{h - 1 \mid h \in H\}$. Now $\mathfrak{A}_G(H)$ is the kernel of the natural map of $\mathfrak{F}G$ onto $\mathfrak{F}(G/H)$; {see for example proof of Theorem 1 in [2]}. Thus $\mathfrak{F}(G/H) \cong \mathfrak{F}G/\mathfrak{A}_G(H)$ is semi-simple. Since G/H is abelian by Lemma

2, so it is clear that it has no elements of order p, as $\mathfrak{F}(G/H)$ is semi-simple.

This also completes the proof of Theorem 2.

REFERENCES

1. D. S. Passman, *Infinite Group Rings*, Marcel Dekkar Inc., N.Y., 1971.
2. I. Sinha, *On the augmentation-maps of subgroups of a group*, Math. Zeitschs., **94** (1966), 193-206.

Received April 18, 1973.

MICHIGAN STATE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1. 1974

HOMOMORPHISMS OF RIESZ SPACES

C. T. TUCKER

If L is a Riesz space (lattice ordered vector space), a Riesz homomorphism of L is an order preserving linear map which preserves the finite operations "$\vee$" and "$\wedge$". It is shown here that if L is one of a large class of spaces and φ is a Riesz homomorphism from L onto an Archimedean Riesz space, then φ preserves the order limits of sequences.

The symbol θ will be used to denote the zero element of any vector space. Suppose L is a Riesz space (lattice ordered vector space). If $f \in L$ then $|f| = f \vee \theta - (f \wedge \theta)$. If M is a linear subspace of L then M is said to be an *ideal* of L if, whenever $|g| \leqq |f|$ and $f \in M$, then $g \in M$. If each of L_1 and L_2 is a Riesz space, a Riesz *homomorphism* φ from L_1 to L_2 is a linear map from L_1 to L_2 which preserves order and the finite operations "$\bigvee$" and "$\bigwedge$". A sequence $f_1, f_2, f_3, \cdots$ of points is said to *order converge* to the point f if there exists a sequence $u_1 \geqq u_2 \geqq u_3 \geqq \cdots$ and a sequence $v_1 \leqq v_2 \leqq v_3 \leqq \cdots$ of points such that $\bigvee v_p = f$, $\bigwedge u_p = f$, and $v_p \leqq f_p \leqq u_p$. Order convergence for nets is defined analogously. A sequence $f_1, f_2, f_3, \cdots$ of elements of the Riesz space L is said to converge *relatively uniformly* to the element f of L if there exists an element g of L (called the regulator) such that if $\varepsilon > 0$, there exists a number N_ε such that if n is a positive integer greater than N_ε, then $|f - f_n| \leqq \varepsilon g$. A Riesz space L is said to be *Archimedean* if, whenever f and g are two points of L such that $\theta \leqq nf \leqq g$ for all positive integers n, then $f = \theta$. Also L is said to be *σ-complete* if each countable set of positive elements has a greatest lower bound and *complete* if each set of positive elements has a greatest lower bound. If φ is a Riesz homomorphism which preserves the order limits of sequences then φ is said to be a Riesz *σ-homomorphism*. If φ preserves the order limits of nets it is said to be a *normal Riesz homomorphism*. A one-to-one onto map which is a Riesz homomorphism is a *Riesz isomorphism*. If H is a subset of L, H^+ will denote the set of all points f of H such that $f \geqq \theta$. If $f \in L$ then f^+ denotes $f \vee \theta$.

Suppose L is a Riesz space, M is an ideal of L, and the algebraic quotient L/M is partially ordered as follows: If each of H and K belongs to L/M and there is an element h of H and k of K such that $h \geqq k$, then $H \geqq K$. It follows that L/M is a Riesz space and the normal map $\pi: L \rightarrow L/M$ is a Riesz homomorphism (Luxemburg and Zaanen [3], p. 102). The coset of L/M containing f will be denoted $[f]$. Further, if M is the kernel of a Riesz homomorphism φ defined on a Riesz space L then the image of φ is Riesz isomorphic to L/M. (Luxemburg and Zaanen [3], p. 102).

If M is a subset of a Riesz space L with the property that whenever $m_1, m_2, m_3, \cdots$ is a sequence of points of M which converges relatively uniformly to a point b of L, b is in M, then M is said to be *uniformly closed.*

In many instances properties of Riesz homomorphisms can be related to properties of their kernels. The following four theorems which are examples of this are listed for future reference.

THEOREM A. *If L is a Riesz space and φ is a Riesz homomorphism defined on L then $\varphi(L)$ is Archimedean if and only if the kernel of φ is uniformly closed.* (See Veksler [8] or Luxemburg and Zaanen [3], Theorem 60.2.)

An ideal M of L is called a *σ-ideal* if, whenever $\{m_i\}$ is a countable subset of M and $b = \bigvee m_i$, then $b \in M$.

THEOREM B. *Suppose L is a Riesz space and φ is a Riesz homomorphism from L onto the Riesz space K. Then φ is a Riesz σ-homomorphism if and only if the kernel of φ is a σ-ideal.* (See Luxemburg and Zaanen [3], Theorem 18.11.)

THEOREM C. *Suppose L is a σ-complete Riesz space and φ is a Riesz σ-homomorphism defined on L. Then $\varphi(L)$ is σ-complete.* (See Veksler [7] or Luxemburg and Zaanen [3], Theorem 59.3.)

An ideal M of L is called a *band* if, whenever $\{m_\alpha\}$, $\alpha \in \lambda$, is a subset of M and $b = \bigvee m_\alpha$, then $b \in M$.

THEOREM D. *Suppose L is a Riesz space and φ is a Riesz homomorphism from L onto the Riesz space K. Then φ is a normal Riesz homomorphism if and only if the kernel of φ is a band.* (See Luxemburg and Zaanen [3], Theorem 18.13.)

A question of interest is when can properties of L imply properties of a class of Riesz homomorphisms defined on L. By combining some known results it can be noted that to place requirements on all the Riesz homomorphisms on L is quite strong.

The sequence $f_1, f_2, f_3, \cdots$ is called a *uniform Cauchy sequence* (with regulator g) if, for each $\varepsilon > 0$, there is a number N such that if n and m are positive integers and $n, m > N$, then $|f_n - f_m| \leqq \varepsilon g$. The Riesz space is *uniformly complete* whenever every uniform Cauchy sequence (with regulator g) converges uniformly (with regulator

g) to a point of L.

PROPOSITION 1. *Suppose L is a uniformly complete Archimedean Riesz space. Each two of the following four statements are equivalent:*

(1) *For each Riesz homomorphism φ defined on L, $\varphi(L)$ is Archimedean,*

(2) *For each Riesz homomorphism φ from L onto a Riesz space K, φ is a Riesz σ-homomorphism,*

(3) *For each Riesz homomorphism φ from L onto a Riesz space K, φ is a normal Riesz homomorphism, and*

(4) *There is a nonempty set X such that L is Riesz isomorphic to the space of all real functions which are zero except on some finite subset of X.*

Proof. By a theorem of Luxemburg and Moore [2], $(1) \rightarrow (4)$. By Theorems A, B, and D, $(4) \rightarrow (3) \rightarrow (2) \rightarrow (1)$.

On the other hand, if requirements are placed on only a subcollection of the collection of all Riesz homomorphisms on L, results of wider applicability can be obtained. In particular, in the following theorems, it is shown that for a large class of Riesz spaces every Riesz homomorphism onto an Archimedean Riesz space is a Riesz σ-homomorphism.

If ω is a subset of L, ω^d denotes the set of all elements g such that $|g| \wedge |f| = \theta$ for each point f of ω. If M is a band in L it is said to be a *projection band* if $L = M \oplus M^d$.

A *principal band* is a band generated by a single element. The Riesz space L is said to have the *principal projection property* if every principal band is a projection band. The Riesz space L has the principal projection property if and only if for each pair of points f and g of L^+, $\bigvee_{n=1}^{\infty} (nf \wedge g)$ exists. (See Luxemburg and Zaanan [3], Theorem 24.7.)

Order convergence in L is said to be *stable* if whenever $f_1, f_2, f_3, \cdots$ is a sequence order converging to θ there is an unbounded, nondecreasing sequence of positive numbers $c_1, c_2, c_3, \cdots$ such that $c_1 f_1, c_2 f_2, c_3 f_3, \cdots$ order converges to θ. Order convergence in the spaces L_p, $1 \leqq p < \infty$; l_p, $1 \leqq p < \infty$; and C_0 is stable.

If order convergence in L is stable then every uniformly closed ideal in L is a σ-ideal. Thus if φ is a Riesz homomorphism from L onto an Archimedean Riesz space K, then φ is a Riesz σ-homomorphism.

For certain sets X order convergence in R^X is not stable. This can be seen as follows: Let X be the set to which x belongs only if x is an unbounded, nondecreasing sequence of positive numbers. Let

f_n be the function defined on X such that if $c_1, c_2, c_3, \cdots$ is a point of X then $f_n(c_1, c_2, c_3, \cdots)$ is $1/c_n$. Then $f_1, f_2, f_3, \cdots$ order converges to θ, but if $c_1, c_2, c_3, \cdots$ is an unbounded, nondecreasing sequence of positive numbers then $c_1 f_1, c_2 f_2, c_3 f_3, \cdots$ does not order converge to θ since $c_n f_n(c_1, c_2, c_3, \cdots) = 1$ for each positive integer n. If X is made of larger cardinality then clearly order convergence in R^X still fails to be stable.

The author, in a paper concerned with the order properties of convergence of Baire functions [6], defined a positive element x of a Riesz space L to have *property c* if for each sequence $h_1 \leqq h_2 \leqq h_3 \leqq \cdots$ of elements of L such that $x = \bigvee h_i$, there exists an element b of L such that for each positive integer n, $b \leqq \sum_{i=1}^{n} h_i$.

EXAMPLE 2. The constant function 1 in R^x has property c.

The constant function 1 in $B[0, 1]$ (the space of all Baire functions on the interval $[0, 1]$) has property c.

Let ω be the set of all functions defined on the interval $[0, 1]$ whose ranges are a subset of the rational numbers and let Q be the vector space generated by ω. Then Q is a Riesz space with the principal projection property but is not uniformly complete. This can be seen as follows: If f is in ω, H is a subset of the interval $[0, 1]$, and $\tilde{f}$ is the function obtained by setting f to zero on H and leaving it unchanged off H, then $\tilde{f}$ is in ω. For Q to be a Riesz space it is sufficient that $f \vee \theta$ exists for each point f of Q. Thus, if f is in Q it is of the form $\sum_{i=1}^{n} c_i f_i$ where the f_i's are in ω. Let H be the set of numbers x for which $f(x) < 0$. Then $f \vee \theta = \sum_{i=1}^{n} c_i \tilde{f}_i$ and $f \vee \theta$ is in Q. Clearly Q has the principal projection property. Each point of Q has as range a countable number set, but a function which fails to have this property, say $g(x) = x$ on the interval $[0, 1]$, is the uniform limit of a sequence of points of Q. Further the constant function 1 in Q has property c.

Let L be a Riesz space and x a positive element of L which has property c and M be a sub Riesz space of L containing x with the property that if f belongs to L then there is a point g to M such that $g \geqq f$. Then x has property c in M.

THEOREM 3. *Suppose L is an Archimedean Riesz space containing a point x which has property c. Then each Riesz homomorphism φ of L into an Archimedean Riesz space K is a Riesz σ-homomorphism.*

Proof. If it can be shown that $f_1 \leqq f_2 \leqq f_3 \leqq \cdots \leqq \theta$ and $\bigvee f_p = \theta$ implies $\bigvee \varphi(f_p) = \theta$, then the theorem is proved.

Now

$$\begin{aligned}
&f_p \vee (-x) + f_p \wedge (-x) = f_p - x \\
&\varphi(f_p \vee (-x)) + \varphi(f_p \wedge (-x)) = \varphi(f_p) - \varphi(x) \\
&\varphi(f_p \wedge (-x)) + \varphi(x) = \varphi(f_p) - \varphi(f_p \vee (-x)) \\
&\qquad = \varphi(f_p \wedge (-x) + x) = \varphi((f_p + x) \wedge \theta) \\
&\sum_{p=1}^{n} \varphi((f_p + x) \wedge \theta) = \sum_{p=1}^{n} \varphi(f_p) - \varphi(f_p \vee (-x)) \\
&\varphi\Big(\sum_{p=1}^{n} (f_p + x) \wedge \theta\Big) = \sum_{p=1}^{n} \varphi(f_p) - \varphi(f_p \vee (-x)) .
\end{aligned}$$

As x has property c there exists an element b such that $b \leqq \sum_{p=1}^{n} (f_p + x) \wedge \theta$ for each positive integer n. Thus,

$$\varphi(b) \leqq \varphi\Big(\sum_{p=1}^{n} (f_p + x) \wedge \theta\Big) = \sum_{p=1}^{n} \varphi(f_p) - \varphi(f_p \vee (-x)) .$$

Suppose that $u \leqq \theta$ is an upper bound for $\{\varphi(f_p)\}$. Then

$$\varphi(b) \leqq \sum_{p=1}^{n} (u - \varphi(f_p \vee (-x))) \leqq \sum_{p=1}^{n} (u - \varphi(-x)) = n(u - \varphi(-x)) .$$

Thus, $u - \varphi(-x) \geqq \theta$ as K is Archimedean and $u \geqq \varphi(-x)$.

But if x has property c, $(1/n)x$ has property c for each positive integer n. Therefore, $u \geqq (1/n)\varphi(-x)$ and $u = \theta$ as K is Archimedean. So $\bigvee \varphi(f_p) = \theta$ and φ is a Riesz σ-homomorphism.

Frequently inclusion maps do not preserve the order limits of sequences. For instance the inclusion map of the space of continuous functions on the interval [0, 1] into the space of all functions on the interval [0, 1] fails to preserve the order limits of sequences. For this reason most theorems which guarantee that a Riesz homomorphism is a Riesz σ-homomorphism require that the mappings be onto. Theorem B would not be true if φ was not specified to be an onto map because of the example just noted. However in view of Theorem 3, no such problem can arise in a space that contains an element with property c. Any embedding of such a space into an Archimedean space must preserve the order limits of sequences.

If in Theorem 3, x is assumed to be a strong unit (a point with the property that if $f \in L$ there is a number r such that $rx \geqq |f|$) rather than have property c, then the statement is no longer true. For instance, let L consist of the set of all bounded sequences and M be the set of all sequences $s_1, s_2, s_3, \cdots$ with the property that if $\varepsilon > 0$ there is only a finite number of positive integers n such that $|s_n| > \varepsilon$. Then M is a uniformly closed ideal but not a σ-ideal.

The Riesz space L is σ-complete if and only if it is uniformly complete and has the principal projection property (Luxemburg and Zaanan [3], Theorem 42.5). If L is uniformly complete and φ is a

Riesz homomorphism defined on L then $\varphi(L)$ is uniformly complete (Luxemburg and Moore [2]).

Thus the question of when the operation of taking a quotient preserves the property of σ-completeness can be included in the question of when this operation preserves the principal projection property.

The Riesz space L has the *quasi principal projection property* if for each point f of L, $L = \{f\}^d \oplus \{f\}^{dd}$. Then L has the principal projection property if and only if it has the quasi principal projection property and is Archimedean. If L has the quasi principal projection property then for each point f of L and g of L there is a unique element g_1 of $\{f\}^d$ and a unique element g_2 of $\{f\}^{dd}$ such that $g = g_1 + g_2$. Denote g_2 by $P_f(g)$.

THEOREM 4. *Suppose L is a Riesz space with the quasi principal projection property, M is an ideal of L, and π is the natural map of L onto L/M. Then the following two conditions are equivalent*:

(1) *If m is a point of M, $P_m L$ is a subset of M and*

(2) (a) *L/M has the quasi principal projection property and*

(b) *$\pi P_f = P_{\pi f}\pi$ for each point f of L.*

Proof. Suppose Condition 1 is true and each of H and K belongs to $(L/M)^+$. We wish to show that there exist points H_1 and H_2 belonging to K^d and K^{dd} respectively such that $H = H_1 + H_2$. There exist points h and k in L^+ such that $H = [h]$ and $K = [k]$. As L has the quasi principal projection property there exist points h_1 and h_2 of $\{k\}^d$ and $\{k\}^{dd}$ respectively such that $h = h_1 + h_2$. Now $H = [h_1] + [h_2]$ and $[h_1] \wedge [h_2] = \theta$. Since h_1 is in $\{k\}^d$, $h_1 \wedge k = \theta$, so $[h_1] \wedge [k] = [h_1 \wedge k] = \theta$ and $[h_1]$ belongs to $\{K\}^d$. Suppose $J \geqq \theta$ is in $\{K\}^d$, i.e., $J \wedge K = \theta$. There is a point j of L^+ such that $[j] = J$. There is a point m of M such that $j \wedge k = m$. By hypothesis there exists a point m_1 of M such that $P_m(j) = m_1$. Thus there is a point $j_1 \geqq \theta$ and a point $m_1 \geqq \theta$ such that $j_1 + m_1 = j$, j_1 is in $\{j \wedge k\}^d$, and m_1 is in $\{j \wedge k\}^{dd}$. Since $j_1 + m_1 = j$ and $m_1 \geqq \theta$, $j_1 \leqq j$ and $j_1 \wedge j = j_1$. Therefore, $\theta = j_1 \wedge (j \wedge k) = (j_1 \wedge j) \wedge k = j_1 \wedge k$ or $(j - m_1) \wedge k = \theta$. So $j - m_1$ is in $\{k\}^d$ and hence $(j - m_1) \wedge h_2 = \theta$. It follows that $[j] \wedge [h_2] = \theta$ and $[h_2]$ is in $\{K\}^{dd}$.

Also $\pi P_k(h) = \pi(h_2) = [h_2] = P_K(H) = P_{\pi k}\pi(h)$.

Suppose Condition 2 is true. If m is a point of M and h is a point of L

$$\theta = P_\theta \pi(h) = P_{\pi m}\pi(h) = \pi P_m(h) .$$

Thus $P_m(h)$ belongs to M.

COROLLARY 5. *Suppose L is a Riesz space with the quasi* *principal projection property, M is an ideal of L, and π is the natural map of L onto L/M. Then the following two conditions are equivalent:*

(1) (a) *If m is a point of M, P_mL is a subset of M and*
(b) *M is relatively uniformly closed, and*

(2) (a) *L/M has the principal projection property and*
(b) *$\pi P_f = P_{\pi f}\pi$ for each point f of L.*

Proof. For L/M to have the principal projection property it is equivalent that L/M have the quasi principal projection property and be Archimedean. By Theorem A it is necessary and sufficient for L/M to be Archimedean that M be uniformly closed.

THEOREM 6. *Suppose L is a Riesz space with the quasi principal projection property and M is an ideal of L. Consider the following two properties:*

(1) (a) *If m is a point of M, P_mL is a subset of M and*
(b) *M is relatively uniformly closed, and*

(2) *M is a σ-ideal.*

Then properties 1 *and* 2 *are independent. If L is assumed to have the principal projection property then property* 2 *implies property* 1 *but property* 1 *does not necessarily imply property* 2. *If L is assumed to be uniformly complete then property* 1 *implies property* 2, *but property* 2 *does not necessarily imply property* 1.

Proof. Suppose L is assumed to have the principal projection property and property 2. For each positive integer n and point m of M, $nm \wedge h$ belongs to M as M is an ideal. Now $P_mh = \bigvee (nm \wedge h)$, P_mh belongs to M since M is a σ-ideal, and property 1(a) holds. Property 1(b) is clearly true.

An example of a space with the principal projection property in which property 1 does not imply property 2 is the following: Let L be the subspace of the space of all sequences generated by the collection of all constant sequences and all sequences which are zero except for a finite number of terms. Let M be the ideal consisting of the collection of all sequences which are zero except for a finite number of terms. Then M satisfies property 1 but not property 2.

Assume L is uniformly complete and property 1 is true. Suppose $\{m_1, m_2, m_3, \cdots\}$ is a subset of M^+ and $h = \bigvee_{i=1}^{\infty} m_i$. Let $r_p = \bigvee_{i=1}^{p} m_i$. Then $\theta \leqq r_1 \leqq r_2 \leqq r_3 \leqq \cdots$ and $\bigvee_{i=1}^{\infty} r_i = h$. Let j be a positive integer, $f_1 = P_{r_{j+1}}h$, $f_2 = h - f_1$, $g_1 = P_{r_j}h$, $g_2 = h - g_1$, and $d_j = f_1 - g_1$. Note that d_j is in M. Since $f_1 + f_2 = g_1 + g_2$, $d_j = g_2 - f_2$. As each of g_2 and f_2 is in $\{r_j\}^d$, d_j is in $\{r_j\}^d$ and $d_j \wedge g_1 = \theta$. Thus $d_j \vee g_1 = f_1$.

Therefore, there exists a countable pairwise disjoint subset $\{d_1, d_2,$

$d_3, \cdots\}$ of M such that $h = \bigvee_{i=1}^{\infty} d_i$. Now the sequence d_1, $d_1 + (1/2)d_2$, $d_1 + (1/2)d_2 + (1/3)d_3$, $d_1 + (1/2)d_2 + (1/3)d_3 + (1/4)d_4, \cdots$ converges relatively uniformly to a point m of M. Then h belongs to the band generated by m, $P_m h = h$, and it follows that h is in M.

An example of a uniformly complete space with the quasi principal projection property in which property 2 does not imply property 1 is the lexiographically ordered plane. The vertical axis is a σ-ideal but does not have property 1 (a).

Suppose L is a Riesz space and $e \geqq \theta$ is a point of L. Then e will be called a *weak unit* if $e \wedge |f| = \theta$ only in case $f = \theta$.

When necessary, it will be assumed that L is a subspace of the set of all almost finite extended real valued continuous functions on an extremally disconnected compact Hausdorff space S. Further if L has a weak unit e, this subspace may be chosen so that e is the function identically to 1.

Suppose e is a weak unit of the Riesz space L. The pair (L, e) will be said to be a *Vulikh algebra* if a multiplication can be defined on L which makes it an associative, commutative algebra with multiplicative unit e which is positive in the sense that if $f \geqq \theta$ and $g \geqq \theta$ then $fg \geqq \theta$. For some properties of Vulikh algebras see Rice [4], Tucker [5], or Vulikh [9], [10].

Suppose that it is assumed that L is a subspace of the set of all almost finite extended real valued continuous functions on an extremally disconnected compact Hausdorff space S and that e is the function identically equal to 1. If each of f and g belong to L their pointwise product will be defined as follows: Both f and g are finite on a dense subset Q of S. Their pointwise product on Q is a continuous function on Q and can be extended uniquely to a continuous function on S, since S is extremally disconnected.

There is at most one multiplication which makes (L, e) a Vulikh algebra (Kantorovitch, Vulikh, and Pinsker [1]). If (L, e) is a Vulikh algebra and it is represented as a Riesz space as a subspace of the set of all almost finite extended real valued continuous functions on an extremally disconnected compact Hausdorff space with e the constant function 1, then the Vulikh algebra multiplication will be the same as the pointwise multiplication described above.

THEOREM 7. *Suppose L is a Riesz space with the principal projection property, M is a uniformly closed ideal of L, π is the natural map of L onto L/M and for each m in M^+, if K is the principal band generated by m, (K, m) is a Vulikh algebra. Then L/M has the principal projection property and $\pi P_f = P_{\pi f}\pi$ for each point f of L.*

Proof. By Theorem 4 it is sufficient to show that for each point m of M^+ and f of L^+ that $\bigvee (nm \wedge f)$ belongs to M. Let K be the principal band generated by m.

By the representation theorem for Riesz spaces K can be assumed to consist of almost finite continuous extended real valued functions on a compact Hausdorff space S, where m is the constant function with value 1 everywhere.

Let $h = \bigvee (nm \wedge f)$. The point h belongs to K. By hypothesis (K, m) is a Vulikh algebra. Thus h^2 belongs to K.

Suppose x is a point of S. If $h(x) \geqq n$, then

$$(h - (nm \wedge f))(x) \leqq h(x) \leqq \frac{1}{n} h^2(x) .$$

If $h(x) < n$, then

$$(h - (nm \wedge f))(x) = 0 \leqq \frac{1}{n} h^2(x) .$$

Thus $m \wedge f$, $2m \wedge f$, $3m \wedge f$, $\cdots$ converges relatively uniformly to h with regulator h^2. As M is uniformly closed, h is in M.

If α is a subset of L^+ with the property that for each two points f and g of α, $f \wedge g = \theta$, then α is said to be *orthogonal.*

THEOREM 8. *Suppose L is a Riesz space with the principal projection property, M is a uniformly closed ideal of L with the property that if $\{f_1, f_2, f_3, \cdots\}$ is a bounded countable orthogonal subset of M^+ there is an unbounded nondecreasing positive number sequence $c_1, c_2, c_3, \cdots$ such that $\{c_1 f_1, c_2 f_2, c_3 f_3, \cdots\}$ is bounded, and π is the natural map of L onto L/M. Then L/M has the principal projection property and $\pi P_f = P_{\pi f}\pi$ for each point f of L.*

Proof. By Theorem 4 it is sufficient to show that for each point m of M^+ and f of L^+ that $\bigvee (nm \wedge f)$ belongs to M.

Let K be the principal band generated by m. By hypothesis K is a projection band, let $h = \bigvee (nm \wedge f)$. The point h belongs to K. Also $\bigvee (nm \wedge f) = \bigvee (nm \wedge h)$.

If k is in K^+, let $\chi(k) = \bigvee (nk \wedge m)$. This supremum exists as K has the principal projection property. Let

$$d_n = \chi((nm \wedge h - (n-1)m)^+) - \chi(((n+1)m \wedge h - nm)^+) .$$

By the representation theorem for Riesz spaces K can be assumed to consist of almost finite continuous extended real valued functions on a compact Hausdorff space S, where m is the constant function with value 1 everywhere.

Suppose x is a point of S. If $h(x) > n$, then $d_n(x) = 0$, if

$n \geqq h(x) > n-1$, then $d_n(x) = 1$, and if $h(x) \leqq n-1$, then $d_n(x) = 0$. Let $h_n = (nm \wedge h - (n-1)m)^+ - \chi((h - nm)^+) + (n-1)d_n$. If $h(x) > n$, then $h_n(x) = 0$, if $n \geqq h(x) > n-1$, then $h_n(x) = h(x)$, and if $h(x) \leqq n-1$, then $h_n(x) = 0$.

Therefore $\{h_1, h_2, h_3, \cdots\}$ is an orthogonal subset of M^+ bounded above by h. By hypothesis there is an unbounded nondecreasing positive number sequence $c_1, c_2, c_3, \cdots$ such that $\{c_1h_1, c_2h_2, c_3h_3, \cdots\}$ is bounded above by a point b of L. Then if i is a positive integer, $h - (h_1 + h_2 + \cdots + h_i) \leqq (1/c_{i+1})b$, and the sequence $h_1, h_1 + h_2, h_1 + h_2 + h_3, \cdots$ converges relatively uniformly to h. As M is uniformly closed, h is in M.

COROLLARY 9. *Suppose L is a Riesz space which is σ-complete and with the property that if $\{f_1, f_2, f_3, \cdots\}$ is a bounded countable orthogonal subset of L^+ there is an unbounded nondecreasing positive number sequence $c_1, c_2, c_3, \cdots$ such that $\{c_1f_1, c_2f_2, c_3f_3, \cdots\}$ is bounded then every Riesz homomorphism φ from L onto an Archimedean Riesz space is a Riesz σ-homomorphism.*

EXAMPLE 10. Suppose L is one of the space L_p, $1 \leqq p < \infty$; l_p, $1 \leqq p < \infty$; or C_0 in which order convergence is stable or L is one of the spaces R^X or $B[0, 1]$ which has a point with property c as described in Example 2. Then L satisfies the conditions of Corollary 9. On the other hand, let L be the space of all functions defined on the x-axis with compact support. In this case L satisfies the hypothesis of Corollary 9, while L neither contains a point with property c nor is order convergence stable in L.

By what has just been shown, if L is a σ-complete Riesz space with the property that if $\{f_1, f_2, f_3, \cdots\}$ is a bounded countable orthogonal subset of L^+ then there is an unbounded nondecreasing positive number sequence $c_1, c_2, c_3, \cdots$ such that $\{c_1f_1, c_2f_2, c_3f_3, \cdots\}$ is bounded is sufficient to imply that every uniformly closed ideal is a σ-ideal, but this condition is not necessary, as the following example shows.

EXAMPLE 11. Let S be the set of all ordered pairs of positive integers. Let L be the collection to which f belongs only in case f is a real valued function on S with the property that there is a set ω which includes all but at most a finite number of positive integers such that if k is a positive integer in ω, $f(1, k)$, $f(2, k)$, $f(3, k)$, $\cdots$ is a bounded number sequence.

The space L is a complete Riesz space.

Suppose M is an ideal which is uniformly closed. Let f be the l.u.b. of a countable subset α of M. Let β be the collection to which

g belongs only in case there is a positive integer k and a member h of α such that $g(k, p) = h(k, p)$ for each positive integer p and if i is a positive integer not k then $g(i, p) = 0$ for each positive integer p. Then f is the l.u.b. of β. For each positive integer k, let f_k be the function such that $f_k(k, p) = f(k, p)$ for each positive integer p and if i is a positive integer not k then $f_k(i, p) = 0$ for each positive integer p.

The function which is equal to $f(i, j)$ at (i, j) and zero elsewhere is in M. Then since the function which is $pf_k(i, p)$ at (i, p) is in L, f_k is in M. Since the function which is $if(i, j)$ at (i, j) is in L, f is in M.

Thus each uniformly closed ideal of M is a σ-ideal. For each positive integer i let g_i be the function such that $g_i(p, q) = 1$ if $p = i$ and $g_i(p, q) = 0$ if $i \neq p$. Then $\{g_1, g_2, g_3, \cdots\}$ is an orthogonal subset of L which is bounded above by the constant function 1 but there is no nondecreasing unbounded positive number sequence $c_1, c_2, c_3, \cdots$ such that $\{c_1g_1, c_2g_2, c_3g_3, \cdots\}$ is bounded above.

The Riesz space L has the *projection property* if every band in L is a projection band. Suppose L has the projection property, ω is a subset of L, H is the band generated by ω, and f is a point of L. There is a unique point f_1 of H^d and a unique point f_2 of H such that $f = f_1 + f_2$. Denote f_2 by $P_\omega(f)$.

The analogous question of when can the projection property be preserved in a natural manner can be answered easily.

THEOREM 12. *Suppose L is a Riesz space with the projection property, M is an ideal of L, and π is the natural map of L onto L/M. Then the following two properties are equivalent*:

(1) *π is a normal Riesz homomorphism, and*

(2) (a) *L/M has the projection property, and*

(b) *$\pi P_\omega = P_{\pi\omega}\pi$ for each subset ω of L.*

Proof. If (1) is true then the kernel of π, M, is a projection band and 2 (a) and (b) clearly hold. If (2) is true and ω is a subset of M with the point f as least upper bound, then $\pi P_\omega f = \pi f$, but $P_{\pi\omega}\pi f = P_\theta \pi f = \theta$.

Also, several answers to the question of when is every Riesz σ-homomorphism from an Archimedean Riesz space L onto a Riesz space K a normal Riesz homomorphism are given in Theorem 29.3 of Luxemburg and Zaanen [3].

REFERENCES

1. L. Kantorovitch, B. Vulikh, and A. Pinsker, *Functional Analysis in Partially Ordered Spaces*, Gostekhizdat, Moscow, 1950, (Russian).
2. W. A. J. Luxemburg and L. C. Moore, Jr., *Archimedean quotient Riesz spaces*, Duke Math. J., **34** (1967), 725-740.
3. W. A. J. Luxemburg and A. C. Zaanen, *Riesz Spaces I*, North-Holland, Amsterdam, 1971.
4. Norman M. Rice, *Multiplication in vector lattices*, Canad. J. Math., **20** (1968), 1136-1149.
5. C. T. Tucker, *Sequentially relatively uniformly complete Riesz spaces and Vulikh algebras*, Canad. J. Math., **34** (1972), 1110-1113.
6. ———, *On a theorem concerning Baire functions*, Proc. Amer. Math. Soc., **41** (1973), 173-178.
7. A. I. Veksler, *On factor lineals in vector lattices*, A.I. Gercen Ped. Inst., Leningrad, **183** (1958), 107-127.
8. ———, *The Archimedean principle in homomorphic images of l-groups and of vector lattices*, 1zv. Vysš. Učebn Zaved. Mathematika 1966, **5** (53), 33-38.
9. B. Z. Vulikh, *The product in linear partially ordered spaces and its applications to the theory of operators*, Mat. Sb. (N.S.), **22** (64) (1948); I, 27-78; II, 267-317, (Russian).
10. ———, *Introduction to the Theory of Partially Ordered Spaces*, Wolters-Noordhoff, Groningen, 1967.

Received March 5, 1973.

UNIVERSITY OF HOUSTON

PACIFIC JOURNAL OF MATHEMATICS
Vol. 55, No. 1, 1974

THE EXCHANGE PROPERTY AND DIRECT SUMS OF INDECOMPOSABLE INJECTIVE MODULES

KUNIO YAMAGATA

This paper contains two main results. The first gives a necessary and sufficient condition for a direct sum of indecomposable injective modules to have the exchange property. It is seen that the class of these modules satisfying the condition is a new one of modules having the exchange property. The second gives a necessary and sufficient condition on a ring for all direct sums of indecomposable injective modules to have the exchange property.

Throughout this paper R will be an associative ring with identity and all modules will be right R-modules.

A module M has *the exchange property* [5] if for any module A and any two direct sum decompositions

$$A = M' \oplus N = \sum_{i \in I} \oplus A_i$$

with $M' \cong M$, there exist submodules $A'_i \subseteq A_i$ such that

$$A = M' \oplus \sum_{i \in I} \oplus A'_i .$$

The module M has *the finite exchange property* if this holds whenever the index set I is finite. As examples of modules which have the exchange property, we know quasi-injective modules and modules whose endomorphism rings are local (see [16], [7], [15] and for the other ones [5]).

It is well known that a finite direct sum $M = \bigoplus_{i=1}^{n} M_i$ has the exchange property if and only if each of the modules M_i has the same property ([5, Lemma 3.10]). In general, however, an infinite direct sum $M = \bigoplus_{i \in I} M_i$ has not the exchange property even if each of M_i's has the same property. On the other hand, Fuller [8] has recently proved that every module over a generalized uniserial ring has the exchange property (c.f., see [9, Theorem 9 and corollary to Lemma 12]).

Therefore, two interesting questions arise:

(1) When does the infinite direct sum $M = \bigoplus_{i \in I} M_i$ of modules $M_i (i \in I)$ with the exchange property have the same property?

(2) What ring R has the property that every module M has the exchange property?

In this paper we consider these two problems for the class of modules M which are direct sums of indecomposable injectives and completely make answers to them for such a class of modules. In §1 we show a sufficient condition for a direct sum of modules with local endomorphism rings to have the finite exchange property. In §2 we prove the following results (1′) and (2′).

(1′) A module M which is a direct sum of indecomposable injective modules has the exchange property if and only if it has the finite exchange property, and moreover any of these assertions is equivalent to that the Jacobson radical of the endomorphism ring $\mathrm{End}_R(M)$ of M is $\{f \in \mathrm{End}_R(M) \mid \mathrm{Ker}\, f$ is essential in $M\}$.

(2′) A ring R satisfies the ascending chain condition for (meet-) irreducible right ideals if and only if every direct sum of indecomposable injective modules has the exchange property.

It is not known whether the exchange and finite exchange properties coincide, so the first equivalence in (1′) is meaningful. Since any direct summand of a module with the exchange property has also the same property as mentioned above, the second equivalence in (1′) trivially includes [2, Corollaire 5] concerning a problem on an indecomposable decomposition of a direct summand of the module which is a direct sum of indecomposable injectives (this is a problem of Matlis). (2′) is a strengthening of [19, Theorem 1] and, as seen in it, such a ring in (2′) has interesting properties concerning the Krull-Remak-Schmidt-Azumaya's theorem and a problem of Matlis. If a module M is quasi-injective, all properties in (1′) are also valid for M, but conversely neither of them implies the quasi-injectivity of M. In §3 we show this fact with an example which means that the class of all modules with the exchange property which are direct sums of indecomposable injectives is a new one of modules with the same property. In §4 we generalize the results of Chamard [3, Théorème 3] and Yamagata [17, Theorem 4] which are obtained from the point of view of a problem of Matlis.

The author wishes to express hearty thanks to Prof. Tachikawa for his advices.

1. A semi-T-nilpotent system. We will recall some definitions and elementary results from [9] and [10]. A family $\{M_i\}_{i \in I}$, with an infinite index set I, which consists of modules M_i whose endomorphism rings are local is called (resp. *semi-*) *T-nilpotent system* if for any family of nonisomorphisms $\{f_{i_n}: M_{i_n} \to M_{i_{n+1}} \mid n \geqq 1\}$ (resp. $i_n \neq i_{n'}$ for $n \neq n'$) and any element $x_{i_1} \in M_{i_1}$, there is an integer m depending on x_{i_1} such that $f_{i_m} f_{i_{m-1}} \cdots f_{i_1}(x_{i_1}) = 0$. If $\mathscr{A}$ is the full subcategory of the category of all right modules whose objects are isomorphic to direct sums of M_i's, then it is said to be the induced category from $\{M_i\}_{i \in I}$ and we denote by $\mathscr{J}$ the class of all morphisms f in $\mathscr{A}$ such that for two objects $X = \bigoplus_{j \in J} X_j$ and $Y = \bigoplus_{k \in K} Y_k$ of $\mathscr{A}$ with $f: X \to$

Y and indecomposable modules X_j and Y_k, each $\pi_k f \kappa_j$ is a nonisomorphism where κ_j is the canonical injection of X_j to X and π_k the projection of Y to Y_k. In [9] we then know the quotient category $\bar{\mathscr{A}} = \mathscr{A}/\mathscr{J}$ is C_3-completely reducible abelian.

For a morphism $f\colon M \to N$ and a submodule M_0 of M, $f \mid M_0\colon M_0 \to N$ denotes the restriction of f to M_0. We denote by $\mathrm{End}_R(M)$ an endomorphism ring of a right module M_R over a ring R.

Now we write the proposition, without proof, which will play an important role in our proofs.

PROPOSITION 1.1 ([12], [13]). *Let $\{M_i\}_{i \in I}$ be an infinite family of modules with local endomorphism rings and $M = \bigoplus_{i \in I} M_i$. Then the following conditions are equivalent.*

(i) *$\{M_i\}_{i \in I}$ is a semi-T-nilpotent system.*

(ii) *$\mathscr{J} \cap \mathrm{End}_R(M)$ is the Jacobson radical of $\mathrm{End}_R(M)$.*

In this case, each direct summand of M is also a direct sum of indecomposable modules which are isomorphic to some M_i.

LEMMA 1.2. *For two modules M_1 and M_2, let*

$$M = M_1 \oplus M_2$$

and ρ a projection of M to M_1. Then for a nonzero submodule N of M with $N \cap M_2 = 0$ the restriction $\rho \mid N$ is a monomorphism. If, further, $\rho(N)$ is a direct summand of M, then there exists a submodule N_1 of M_1 such that $M = N \oplus N_1 \oplus M_2$.

Proof. The first assertion is clear. For the rest let $\rho(N)$ be a direct summand of M, $M = \rho(N) \oplus M'$ and ρ a monomorphism on N. By the modular law, we then have

$$M_1 = \rho(N) \oplus N_1$$

with a projection π of M_1 to $\rho(N)$ where $N_1 = M_1 \cap M'$. We consider the decomposition

$$M = \rho(N) \oplus N_1 \oplus M_2 .$$

It is then easy to see that the projection of M to $\rho(N)$ be $\pi\rho$ and the restriction $\pi\rho \mid N$ of $\pi\rho$ to N is an isomorphism by the first part of this lemma. As a consequence, we obtain the desired decomposition

$$M = N \oplus N_1 \oplus M_2 .$$

The following corollaries are essentially proved in [9] but we include proofs for completeness. In them, without proofs, we will use some properties for completely reducible objects in $\bar{\mathscr{A}}$ but they are easily proved in the same way as for completely reducible modules (see [9, p. 331-332]).

COROLLARY 1.3. *Let M be a direct sum of indecomposable modules M_i ($i \in I$), where each M_i has a local endomorphism ring, and $\{N_j\}_{j \in J}$ an independent set of indecomposable submodules of M with local endomorphism rings such that it is a semi-T-nilpotent system. Then, if $\sum_{j \in F} \oplus N_j$ is a direct summand of M for every finite subset $F \subset J$, there exists a subset $K \subset I$ such that*

$$M = \sum_{j \in J} \oplus N_j \oplus \sum_{k \in K} \oplus M_k .$$

REMARK. If J is finite, the finite direct sum $\sum_{j \in J} \oplus N_j$ has the exchange property by [15, Proposition 1] and [5, Lemma 3.10] and is a direct summand of M by hypothesis. Hence there exists a subset $K \subset I$ such that $M = \sum_{j \in J} \oplus N_j \oplus \sum_{k \in K} \oplus M_k$.

Proof. We assume J is infinite. Let $\mathscr{A}$ and $\mathscr{J}$ be as above and $\kappa: N = \sum_{j \in J} \oplus N_j \to M$ an inclusion map. For a morphism f in $\mathscr{A}$ we denote by $\bar{f}$ the induced morphism of f in the quotient category $\bar{\mathscr{A}} = \mathscr{A}/\mathscr{J}$. Since $N_{j_1} \oplus \cdots \oplus N_{j_n}$ is a direct summand for any finite subset $\{j_1, \cdots, j_n\}$ of J by assumption, the restriction of $\bar{\kappa}$ to $\bar{N}_{j_1} \oplus \cdots \oplus \bar{N}_{j_n}$ is then an injection in $\bar{\mathscr{A}}$. This will imply that $\bar{\kappa}$ is an injection in $\bar{\mathscr{A}}$.

To show this we suppose that the kernel $\bar{K} = \operatorname{Ker} \bar{\kappa}$ is not zero. Then there is a finite subset $\{j_1, \cdots, j_n\} \subset J$ such that $\bar{K} \cap (\bar{N}_{j_1} \oplus \cdots \oplus \bar{N}_{j_n}) \neq 0$, because $\bar{\mathscr{A}}$ is a C_3-abelian category and $\bar{N} = \bigoplus_{j \in J} \bar{N}_j$ in $\bar{\mathscr{A}}$ ([9, Theorem 7]). Hence $\bar{\kappa}(\bar{K} \cap \sum_{k=1}^{n} \oplus \bar{N}_{j_k}) \neq 0$ by the fact that $\bar{\kappa} \mid \sum_{k=1}^{n} \oplus \bar{N}_{j_k}$ is injective in $\bar{\mathscr{A}}$, a contradiction.

Then, since the category $\bar{\mathscr{A}}$ is C_3-completely reducible abelian, the morphism $\bar{\kappa}: \bar{N} \to \bar{M}$ splits and by the note just before this corollary there is a subset $K \subset I$ such that

$$M = \sum_{i \in I-K} \oplus M_i \oplus \sum_{k \in K} \oplus M_k \tag{1}$$

and

$$\bar{M} = \bar{N} \oplus \sum_{k \in K} \oplus \bar{M}_k \tag{2}$$

$$= \sum_{i \in I-K} \oplus \bar{M}_i \oplus \sum_{k \in K} \oplus \bar{M}_k . \tag{3}$$

Let the projection of M to $\sum_{j \in I-K} \oplus M_i$ be ρ. Then in (3) the projection of $\bar{M}$ to $\sum_{i \in I-K} \bar{M}_i$ is clearly $\bar{\rho}$ and so $\bar{\rho} \circ \bar{\kappa}$ is a bijection of $\bar{N}$

onto $\sum_{i\in I-K}\oplus \bar{M}_i$ in view of (2) and (3). This means that there is a morphism ϕ of $\sum_{i\in I-K}\oplus M_i$ to N such that $\bar{\phi}\circ\bar{\rho}\circ\bar{\kappa}=1_{\bar{N}}$ and $\bar{\rho}\circ\bar{\kappa}\circ\bar{\phi}=1_{\sum_{i\in I-K}\oplus \bar{M}_i}$. Hence we obtain that $\phi\circ(\rho\circ\kappa)-1$ and $(\rho\circ\kappa)\circ\phi-1$ belong to $\mathscr{J}(\mathrm{End}_R(N))=\mathscr{J}\cap\mathrm{End}_R(N)$ and $\mathscr{J}(\mathrm{End}_R(\sum_{i\in I-K}\oplus M_i))=\mathscr{J}\cap\mathrm{End}_R(\sum_{i\in I-K}\oplus M_i)$ respectively. We will show that $\rho\circ\kappa$ is an isomorphism of N to $\sum_{i-I-K\in K}\oplus M_i$.

First, $\phi\circ(\rho\circ\kappa)-1\in\mathscr{J}(\mathrm{End}_R(N))$ implies that $\phi\circ(\rho\circ\kappa)$ is invertible, because $\mathscr{J}(\mathrm{End}_R(N))$ is the Jacobson radical by Proposition 1.1. The morphism $\rho\circ\kappa$ is hence a monomorphism.

Secondly, to show that $\rho\circ\kappa$ is an epimorphism it suffices to show that the family $\{M_i\}_{i\in I-K}$ is a semi-T-nilpotent system by the same reason in the first part. Now since $\bar{N}=\sum_{j\in J}\oplus\bar{N}_j$ is isomorphic to $\sum_{i\in I-K}\oplus\bar{M}_i$, there is a bijection $\sigma: J\to I-K$ such that $\bar{N}_j\cong\bar{M}_{\sigma(j)}$ for every $j\in J$ because $\mathscr{A}$ is a completely reducible C_3-abelian category (see the note before this corollary). It is therefore easy to see that N_j is isomorphic to $M_{\sigma(j)}$ for every $j\in J$ on account of the facts that $\mathscr{J}\cap\mathrm{End}_R(N_j)$ and $\mathscr{J}\cap\mathrm{End}_R(M_{\sigma(j)})$ are the Jacobson radicals of $\mathrm{End}_R(N_j)$ and $\mathrm{End}_R(M_{\sigma(j)})$ respectively. Hence the assumption that $\{N_j\}_{j\in J}$ is a semi-T-nilpotent system implies that the family $\{M_i\}_{i\in I-K}$ is also semi-T-nilpotent, as desired.

Now then, since $(\rho\circ\kappa)(N)=\rho(N)=\sum_{i\in I-K}\oplus M_i$ is a direct summand of M, we can apply Lemma 1.2 to our case and have that

$$M=N\oplus\sum_{k\in K}\oplus M_k\,,$$

which completes the proof of the corollary.

COROLLARY 1.4. *Let M be an infinite direct sum of indecomposable modules M_i $(i\in I)$ with local endomorphism rings. Assume the family $\{M_i\}_{i\in I}$ is a semi-T-nilpotent system and let $\{N_n\}_{n\geqq 1}$ be a family of direct summands of M such that $N_n\subseteqq N_{n+1}$ for all integers $n\geqq 1$. Then the union $\bigcup_{n\geqq 1}N_n$ of the family $\{N_n\}_{n\geqq 1}$ is also a direct summand of M.*

Proof. Since, according to Proposition 1.1, the union $\bigcup_{n\geqq 1}N_n$ is also a direct sum of indecomposable modules with local endomorphism rings, it is an immediate consequence of Corollary 1.3.

For two modules $M=\bigoplus_{i\in I}M_i$ and $N=\bigoplus_{j\in J}N_j$ we can represent every homomorphism f of M to N as a *column summable* matrix (f_{ji}), that is, for the injections κ_i of M_i to M and projections π_j of N to N_j $(i\in I,\ j\in J)$, $f_{ji}=\pi_j f\kappa_i: M_i\to N_j$ and, for any $x\in M$ and $i\in I$ $f_{ji}(\rho_i(x))=0$ for almost all $j\in J$ where ρ_i is the projection of M onto M_i. Hence, in this case we may denote that $f(x)=\sum_j\pi_j f(x)=\sum_{i,j}f_{ji}(\rho_i(x))$ for any $x\in M$ and $f_i=\sum_{j\in J}f_{ji}$ (see [9], p. 332).

A submodule N of M is *essential* in M ($N \subseteq' M$) if $N \cap L \neq 0$ for all nonzero submodules L of M and M is *uniform* if every nonzero submodule is essential in M. In the following we will denote the kernel of a morphism f by $\operatorname{Ker} f$.

LEMMA 1.5. *Let M be a direct sum of uniform modules M_i ($i \in I$) and $f = (f_{ji}) \in \operatorname{End}_R(M)$. Then $\operatorname{Ker} f$ is essential in M if and only if each $\operatorname{Ker} f_{ji}$ is essential in M_i for all $i, j \in I$.*

Proof. Suppose that $\operatorname{Ker} f_{ji}$ is essential in M_i for all $i, j \in I$. Then to show that $\operatorname{Ker} f \subseteq' M$ it suffices to show that $\operatorname{Ker} f \cap M_i \subseteq' M_i$ for all $i \in I$. Now contrary to it, suppose that for some $i \in I$, $\operatorname{Ker} f \cap M_i$ is not essential in M_i or equivalently $\operatorname{Ker} f \cap M_i = 0$ by reason of the uniformity of M_i. Then for $0 \neq x_i \in M_i$ there exists a finite subset $\{j_1, \cdots, j_n\} \subseteq I$ such that $0 \neq f_i(x_i) = \sum_{k=1}^{n} f_{j_k i}(x_i)$ and $f_{ji}(x_i) = 0$ for all $j \neq j_k$, where $f_i = \sum_{j \in J} f_{ji}$. Because the restriction $f_i = f \mid M_i \colon M_i \to M$ is a monomorphism. On the other hand, by hypothesis, $(\bigcap_{k=1}^{n} \operatorname{Ker} f_{j_k i}) \cap x_i R \neq 0$ and $f_i((\bigcap_{k=1}^{n} \operatorname{Ker} f_{j_k i}) \cap x_i R) = (\sum_{k=1}^{n} f_{j_k i})((\bigcap_{k=1}^{n} \operatorname{Ker} f_{j_k i}) \cap x_i R) = 0$, a contradiction. Thus $\operatorname{Ker} f \cap M_i$ is essential in M_i for every $i \in I$.

Conversely, we assume that $\operatorname{Ker} f \subseteq' M$. Clearly this implies that $\operatorname{Ker} f \cap M_i \subseteq' M_i$ by the uniformity of M_i ($i \in I$). On the other hand, since $f_i(x_i) = \sum_{j \in J} f_{ji}(x_i)$ and $f_{ji}(x_i) \in M_j$ for every $x_i \in M_i$, that $f_i(x_i) = 0$ implies that $f_{ji}(x_i) = 0$ for all $j \in I$. Therefore, $\operatorname{Ker} f_{ji} \neq 0$ for all $i, j \in I$, because $\operatorname{Ker} f_i = \operatorname{Ker} f \cap M_i \neq 0$. As a consequence, $\operatorname{Ker} f_{ji} \subseteq' M_i$ for $i, j \in I$.

LEMMA 1.6 ([9], [10]). *Let $\{M_i\}_{i \in I}$ be a family of a semi-T-nilpotent system of modules with local endomorphism rings and $M = \bigoplus_{i \in I} M_i$. Then S/J is a regular ring in the sense of von Neumann and an idempotent of S/J can be lifted to S, where S is an endomorphism ring of M and J its Jacobson radical.*

This follows from Proposition 1.1, [9, Theorem 7] and [10, Theorem 3].

PROPOSITION 1.7. *Let $\{M_i\}_{i \in I}$ be a family of a semi-T-nilpotent system of modules with local endomorphism rings. Then $M = \bigoplus_{i \in I} M_i$ has the finite exchange property.*

Proof. Let $S = \operatorname{End}_R(M)$ and J the Jacobson radical of S. Then S/J is a regular ring and every idempotent is lifted to S by Lemma 1.6. Hence, for every element $s \in S$ there exists an idempotent $e \in S$ such that $sS + J = eS + J$. This shows that S has the exchange property as a S-module and so M_R has the finite exchange property by [17, Theorems 3 and 4].

2. The exchange property. In this section we prove our main theorems being concerned with modules which are direct sums of indecomposable injectives.

First we will continue to consider a general case of modules with local endomorphism rings instead of indecomposable injectives.

LEMMA 2.1. *Let M, N, and A_i $(i \in I)$ be submodules of a module A such that*

$$A = \sum_{i \in I} \oplus A_i = M \oplus N$$

and, furthermore, let M be a direct sum of indecomposable submodules M_j $(j \in J)$ with local endomorphism rings. If $M \cap \sum_{i \in F} \oplus A_i \neq 0$ for some finite subset F of I, then there exist elements $i_0 \in F$ and $j_0 \in J$ such that

$$A = M_{j_0} \oplus A'_{i_0} \oplus \sum_{i \in I - \{i_0\}} \oplus A_i$$

for a suitable submodule A'_{i_0} of A_{i_0}.

Proof. First we remark that, since each M_j $(j \in J)$ has a local endomorphism ring, it has the exchange property by [15, Proposition 1], so that any finite direct sum of M_j's has also the exchange property ([5, Lemma 3.10]).

Now by hypothesis there exists a finite subset J_0 of J such that $\sum_{j \in J_0} \oplus M_j \cap \sum_{i \in F} \oplus A_i \neq 0$. Hence applying the exchange property of $\sum_{j \in J_0} \oplus M_j$ to the given decomposition $A = \sum_{i \in I} \oplus A_i$, we have decompositions such that

$$A_i = B_i \oplus C_i \ (i \in I) \ ,$$

$$A = \sum_{i \in I} \oplus B_i \oplus \sum_{i \in I} \oplus C_i \tag{1}$$

and

$$= \sum_{j \in J_0} \oplus M_j \oplus \sum_{i \in I} \oplus C_i \ . \tag{2}$$

Here there exists at least one element i_0 of F such that $B_{i_0} \neq 0$. For, if the contrary were true, $\sum_{i \in F} \oplus B_i = 0$ and hence $\sum_{i \in F} \oplus A_i = \sum_{i \in F} \oplus C_i$. So $\sum_{j \in J_0} \oplus M_j \cap \sum_{i \in F} \oplus C_i = \sum_{j \in J_0} \oplus M_j \cap \sum_{i \in F} \oplus A_i \neq 0$ by the definition of J_0, which contradicts the decomposition (2).

Now it is clear that $M' = \sum_{j \in J_0} \oplus M_j$ is isomorphic to $\sum_{i \in I} \oplus B_i$ via the restriction $\pi \mid M'$ of π to M', where π is the projection of A onto $\sum_{i \in I} \oplus B_i$ in the formula (1). It follows that

$$(3) \qquad \pi(M') = \sum_{j \in J_0} \pi(M_j) = B_{i_0} \oplus \sum_{i \in I - \{i_0\}} \oplus B_i \, .$$

Since each $\pi(M_j)$ for $j \in J_0$ is isomorphic to M_j, it has a local endomorphism ring. We can thus apply the Krull-Remak-Schmidt-Azumaya's theorem [1, Theorem 1] to this module $\pi(M')$ and the projection ξ of $\pi(M')$ onto B_{i_0} in the formula (3). As a consequence, there exists an element $j_0 \in J_0$ such that the restriction $\xi \mid \pi(M_{j_0})$ is a monomorphism and $\xi\pi(M_{j_0})$ is a direct summand of $\pi(M')$ and hence of B_{i_0}. On the other hand, a simple computation shows that the projection of A to B_{i_0} in the decomposition (1) is $\xi\pi$. Thus from these facts and Lemma 1.2 there is a submodule D_{i_0} of B_{i_0} such that

$$A = M_{j_0} \oplus D_{i_0} \oplus \sum_{i \in I - \{i_0\}} \oplus B_i \oplus \sum_{i \in I} \oplus C_i$$

because the restriction $\xi\pi \mid M_{j_0}$ is clearly a monomorphism. Setting $A'_{i_0} = C_{i_0} \oplus D_{i_0}$, we finally have a desired decomposition

$$A = M_{j_0} \oplus A'_{i_0} \oplus \sum_{i \in I - \{i_0\}} \oplus A_i \, .$$

From now on we will consider indecomposable injectives.

LEMMA 2.2. *Every indecomposable injective module is uniform and has a local endomorphism ring.*

This is well known (c.f., see [6, §5 Proposition 8]).

Assume M_1 and M_2 are indecomposable injectives and f a morphism of M_1 to M_2. If f is a nonmonomorphism, then its kernel $\operatorname{Ker} f$ is essential in M_1 by Lemma 2.2 and the converse is, of course, true. This shows that f is a nonisomorphism if and only if $\operatorname{Ker} f$ is essential in M_1. Under this observation we have

PROPOSITION 2.3. *Let $\{M_i\}_{i \in I}$ be an infinite family of indecomposable injective modules and $M = \bigoplus_{i \in I} M_i$. Let S be an endomorphism ring of M_R and J the Jacobson radical of S. Then $J = \{f \in S \mid \operatorname{Ker} f \subseteqq' M\}$ if and only if the family $\{M_i\}_{i \in I}$ is a semi-T-nilpotent system.*

Proof. We will represent every endomorphism f of M as a column summable matrix: $f = (f_{ji})$, where $f_{ji} = \pi_j f \kappa_i$ for the projections π_j of M onto M_j and injections κ_i of M_i into $M (i, j \in I)$. Then, in accordance with our earlier notations (see § 1), by the above remark we have

$$\mathscr{J} \cap S = \{f = (f_{ji}) \in S \mid \operatorname{Ker} f_{ji} \subseteqq' M_i\}$$

and by Lemma 1.5

$$\{f = (f_{ji}) \in S \mid \operatorname{Ker} f_{ji} \subseteqq' M_i\} = \{f \in S \mid \operatorname{Ker} f \subseteqq' M\} .$$

On the other hand, we know by Proposition 1.1 that the family $\{M_i\}_{i \in I}$ is a semi-T-nilpotent system if and only if $J = \mathscr{J} \cap S$. It follows from them that $\{M_i\}_{i \in I}$ is a semi-T-nilpotent system if and only if $J = \{f \in S \mid \operatorname{Ker} f \subseteqq' M\}$, which proves the proposition.

We need more lemmas for the main theorems.

LEMMA 2.4. *A module M has the exchange property if for any modules $A_i (i \in I)$ which are isomorphic to submodules of M and any decomposition $A = \bigoplus_{i \in I} A_i = M' \oplus N$ where $M' \cong M$, there exist submodules $A_i' \subseteqq A_i$ such that $A = M' \oplus \sum_{i \in I} \oplus A_i'$.*

This is well known in [5, Theorem 8.2] and its proof will be omitted.

LEMMA 2.5. *Let $G = M \oplus N$ for submodules M and N of a given module G. We moreover assume $M = \sum_{i \in I} \oplus M_i$, where $\{M_i\}_{i \in I}$ is an infinite family of indecomposable injective submodules of G and a semi-T-nilpotent system. Then if a module A is isomorphic to a submodule of M and contains an injective submodule, there exists a maximal submodule A_0 of A with the property that A_0 is a direct sum of indecomposable injective submodules. In this case such a module A_0 is a direct summand of A.*

Proof. Let the monomorphism of A to M be f and E an injective submodule of A. Then by [1, Theorem 1] and Lemma 2.2, $f(E)$ contains an indecomposable injective submodule isomorphic to some M_i in view of that $f(E)$ is a direct summand of M. This implies that A contains a submodule isomorphic to some M_i. Now then we can take a family $\{A_n\}_{n \geqq 1}$ of submodules of A such that each A_n is a direct sum of indecomposable injectives and $A_n \subseteqq A_{n+1}$ for any $n \geqq 1$. Then, by Zorn's lemma, we will be done if we can show that the union $A_0 = \bigcup_n A_n$ is also a direct sum of indecomposable injectives and, furthermore, a direct summand of A.

Since f is a monomorphism, the image $f(A_n)$ of A_n by f is also a direct sum of indecomposable injectives and hence $f(A_n)$ is a direct summand of M for any $j \in J$ by Corollary 1.3 and Lemma 2.2. Thus the union $\bigcup_n f(A_n)$ is also a direct summand of M and a direct sum of indecomposable injectives by Corollary 1.4. Taking account of $f(A_0) = \bigcup_n f(A_n)$, we have $M = f(A_0) \oplus N$ for a submodule N of M and A_0 is a direct sum of indecomposable injectives since $A_0 \cong f(A_0)$. By the modular law, $f(A) = f(A_0) \oplus f(A) \cap N$. We therefore have $A = A_0 \oplus f^{-1}(f(A) \cap N)$, where $f^{-1}(f(A) \cap N)$ is the inverse image of $f(A) \cap N$ by f, which proves the lemma.

It is clear that the exchange property implies the finite exchange property, but it is not known whether the converse is true in general. However, in our case that modules are direct sums of indecomposable injectives we can conclude this question affirmatively.

THEOREM 2.6. *Let M be a module which is a direct sum of indecomposable injective modules and let S be an endomorphism ring of M_R. Then the following assertions are equivalent.*

(i) *M has the exchange property.*

(ii) *M has the finite exchange property.*

(iii) *The Jacobson radical of S is $\{f \in S \mid \operatorname{Ker} f \subseteq' M\}$.*

Proof. Let $M = \sum_{i \in I} \oplus M_i$, where every submodule M_i is indecomposable injective. If the index set I is finite, then M is clearly injective, so all of the above assertions (i), (ii), and (iii) are true. It therefore suffices to show the theorem for only the case with the infinite index set I.

Now let I be an infinite index set. By Proposition 2.3 the assertion (iii) is then equivalent to

(iii′) *The family $\{M_i\}_{i \in I}$ is a semi-T-nilpotent system.*

Thus we will consider (iii′) instead of (iii) in the following.

The implication (i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii′). The idea of the proof is due to [9, Lemma 9]. Assume that M has the finite exchange property. Take an arbitrary countable subfamily of $\{M_i\}_{i \in I}$, say $\{M_n\}_{n \geqq 1}$, and nonisomorphisms $f_n\colon M_n \rightarrow M_{n+1}$ $(n \geqq 1)$. For every $x \in M_1$ we will find an integer $n(x)$ depending on x such that $f_{n(x)} f_{n(x)-1} \cdots f_1(x) = 0$.

For this put $M_n' = \{x + f_n(x) \mid x \in M_n\}$. It is then clear that $M_n' \oplus M_{n+1} = M_n \oplus M_{n+1}$ for $n \geqq 1$. Since each M_i is indecomposable injective, every nonisomorphism f_n is only nonmonomorphism, i.e., $\operatorname{Ker} f_n \neq 0$. This implies that $M_n' \cap M_n \neq 0$ for every $n \geqq 1$.

It is clear that

$$(1)\quad \sum_{i=1}^{\infty} \oplus M_i = M_1' \oplus M_2 \oplus M_3' \oplus M_4 \oplus \cdots \oplus M_{2n-1}' \oplus M_{2n} \oplus \cdots$$

$$(2)\quad \qquad = M_1 \oplus M_2' \oplus M_3 \oplus M_4' \oplus \cdots \oplus M_{2n-1} \oplus M_{2n}' \oplus \cdots$$

and we put

$$N = M_1' \oplus M_3' \oplus \cdots \oplus M_{2n-1}' \oplus \cdots .$$

Then, applying the fact that N has also the finite exchange property ([5, Lemma 3.10]) to the decomposition (2), we have that $\sum_{i=1}^{\infty} \oplus M_i = N \oplus X \oplus Y$ for some submodules X and Y of $\sum_{n=1}^{\infty} \oplus M_{2n-1}$ and $\sum_{n=1}^{\infty} \oplus M_{2n}'$ respectively. Here, in fact, it will hold $X = 0$.

To show this, suppose that $X \neq 0$ contrary. Then by Lemma 2.1 there exists M_{2m-1} such that

$$\sum_{i=1}^{\infty} \oplus M_i = N \oplus M_{2m-1} \oplus X' \oplus Y$$

for some submodule X' of X.

This however contradicts that $0 \neq M_{2m-1} \cap M_{2m-1}' \subseteqq M_{2m-1} \cap N$. Thus it holds

$$\sum_{i=1}^{\infty} \oplus M_i = N \oplus Y .$$

Now we take an arbitrary nonzero element $x \in M_1$ and we let $x = y + z$ with $y \in N$ and $z \in Y$. Considering these y and z in the decompositions $N = \sum_{n=1}^{\infty} \oplus M_{2n-1}'$ and $\sum_{n=1}^{\infty} \oplus M_{2n}'$ respectively, we have

$$y = \sum_{i=1}^{s} (x_{2i-1} + f_{2i-1}(x_{2i-1}))$$

and

$$z = \sum_{i=1}^{s} (x_{2i} + f_{2i}(x_{2i})) ,$$

and substituting these expressions for y and z, we have

$$\begin{aligned} x &= \sum_{i=1}^{s} (x_{2i-1} + f_{2i-1}(x_{2i-1})) + \sum_{i=1}^{s} (x_{2i} + f_{2i}(x_{2i})) \\ &= x_1 + \sum_{i=1}^{2s-1} (f_i(x_i) + x_{i+1}) + f_{2s}(x_{2s}) . \end{aligned}$$

Therefore, $x = x_1$, $f_i(x_i) + x_{i+1} = 0$ $(1 \leqq i \leqq 2s-1)$ and $f_{2s}(x_{2s}) = 0$, that is, $x_1 = x$, $x_2 = -f_1(x)$, $\cdots$, $x_{2s} = f_{2s}(x_{2s-1})$ and $f_{2s-1}(x_{2s}) = 0$. By successive substitutions, we obtain $x_{2s} = (-1)^{2s-1} f_{2s-1} \cdots f_1(x)$ and, finally, $f_{2s} f_{2s-1} \cdots f_1(x) = 0$. Thus we can put $n(x) = 2s$, which completes the proof of (ii) $\Longrightarrow$ (iii′).

(iii$'$) $\Rightarrow$ (i). We assume the family $\{M_i\}_{i\in I}$ is a semi-T-nilpotent system. Suppose $A = \sum_{j\in J} \oplus A_j = M' \oplus N$, where $M' \cong M$ and each A_j is isomorphic to a submodule of M'. Then, taking account of Lemma 2.4, we will be done if we can find submodules A'_j of A_j ($j \in J$) such that $A = M' \oplus \sum_{j\in J} \oplus A'_j$.

For this, we will first refine the given decomposition $A = \sum_{j\in J} \oplus A_j$. We should note that M' is also a direct sum of indecomposable injective submodules $M'_i (i \in I)$. By Lemma 2.1 there exists at least one element $j_0 \in J$ such that A_{j_0} has a nonzero submodule isomorphic to some M'_i. Let the subset of J of such elements $j_0 \in J$ be J_0. By Lemma 2.5 there exist maximal submodules B_j of A_j ($j \in J_0$) such that each B_j is a direct sum of indecomposable injective submodules of A_j, in which case every B_j is a direct summand of A_j, say $A_j = B_j \oplus C_j$ for a submodule $C_j \subset A_j$ for $j \in J_0$. Consequently, we have such a refinement of $A = \sum_{j\in J} \oplus A_j$ that

$$(1) \qquad A = \sum_{j\in J_0} \oplus B_j \oplus \sum_{j\in J_0} \oplus C_j \oplus \sum_{j\in J-J_0} A_j\ ,$$

where $J - J_0$ is the complement of J_0 in J and if $J - J_0$ is empty, we put A_j in the formula (1) to be zero submodule of A for convenience.

Next we will have that

$$(2) \qquad M' \cap (\sum_{j\in J_0} \oplus C_j \oplus \sum_{j\in J-J_0} \oplus A_j) = 0\ .$$

For this we suppose that $M' \cap (\sum_{j\in J_0} \oplus C_j \oplus \sum_{j\in J-J_0} \oplus A_j) \neq 0$. Then by Lemma 2.1 and the choice of J_0 there exists M'_{i_0} such that for a submodule $X_{j_0} \subsetneqq C_{j_0}$,

$$A = M'_{i_0} \oplus X_{j_0} \oplus B_{j_0} \oplus \sum_{j\in J-\{j_0\}} \oplus A_j\ .$$

This implies there exists an injective submodule C'_{j_0} of C_{j_0} which is isomorphic to M'_{i_0}. However, in this case we have that $B_{j_0} \oplus C'_{j_0}$ is a direct summand of A_{j_0} and a direct sum of indecomposable injective submodules, which contradicts the maximality of B_{j_0}.

Now we can exchange the complement N of M' for a direct sum of submodules of $A_j (j \in J)$. For this let the projection of A onto $\sum_{j\in J_0} \oplus B_j$ in (1) be ρ. The family $\{M'_i\}_{i\in I}$ is semi-T-nilpotent by hypothesis, and so is $\{\rho(M'_i)\}_{i\in I}$ because the restriction $\rho\,|\,M'$ of ρ to M' is a monomorphism by (2) and Lemma 1.2. Using Corollary 1.3 the image $\rho(M')$ therefore is a direct summand of $\sum_{j\in J_0} \oplus B_j$ and there is a subset $K \subset J_0$ such that $\sum_{j\in J_0} \oplus B_j = \rho(M') \oplus \sum_{k\in K} \oplus B_k$ and, consequently we have $A = \rho(M') \oplus \sum_{k\in K} \oplus B_k \oplus \sum_{j\in J_0} \oplus C_j \oplus$

$\sum_{j \in J-J_0} \oplus A_j$. Computing the projection of A to $\rho(M')$ and by Lemma 1.2, we therefore have a decomposition

$$A = M' \oplus \sum_{k \in K} \oplus B_k \oplus \sum_{j \in J_0} \oplus C_j \oplus \sum_{j \in J-J_0} \oplus A_j ,$$

which completes the proof of the implication (iii′) $\Rightarrow$ (i). Thus we conclude the theorem.

The original definition of the exchange property given in the introduction is due to Crawly and Jónsson [5]. However, we will consider the following weaker exchange property, too ([10]).

DEFINITION. A direct summand M of a module A has *the exchange property in A* if for any direct sum decomposition $A = \sum_{i \in I} \oplus A_i$, there exist submodules $A'_i \subseteqq A_i$ such that $A = M \oplus \sum_{i \in I} A'_i$.

We recall that for a ring R a right ideal I is *(meet-) irreducible* provided $I \neq R$ and $I = I_1 \cap I_2$ implies $I = I_1$ or $I = I_2$ for all right ideals I_1 and I_2 or R.

THEOREM 2.7. *The following conditions are equivalent.*

(i) *A ring R satisfies the ascending chain condition for irreducible right ideals.*

(ii) *Any direct sum of indecomposable injective modules has the exchange property.*

(iii) *Any direct sum of indecomposable injective modules has the finite exchange property.*

(iv) *Any direct summand of the module M which is a direct sum of indecomposable injective modules has the exchange property in M.*

(v) *For any direct sum M of indecomposable injective modules, the Jacobson radical of the endomorphism ring* $\mathrm{End}_R(M)$ *is* $\{f \in \mathrm{End}_R(M) \mid \mathrm{Ker}\, f \subseteqq' M\}$.

Proof. The equivalences (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (v) are trivial from Theorem 2.6, and (ii) $\Rightarrow$ (iv) follows from [5, Lemma 3.10]. The implication (iv) $\Rightarrow$ (i) is contained in [19, Theorem 1].

(i) $\Rightarrow$ (ii): Let $M = \sum_{i \in I} \oplus M_i$, where M_i is indecomposable injective for any $i \in I$. If I is finite, then M is clearly injective, so it has the exchange property ([16, Lemma 2]). If I is infinite, the family $\{M_i\}_{i \in I}$ is a semi-T-nilpotent system by [19, Theorem 1 and Lemma 2]. Therefore, M has the exchange property by Proposition 2.3 and Theorem 2.6.

3. Example. Here we show the existence of modules which are not quasi-injective but isomorphic to direct sums of indecomposable injectives and have the exchange property.

We first note that a quasi-injective module M over a ring R is injective by the criterion of Fuchs [7, Lemma 2] provided that M has the property that some finite direct sum of copies of M contains an element with a zero annihilator right ideal or, equivalently, contains a submodule isomorphic to the ring R.

The ring R regarded as a right (left) module over itself will be written $R_R({}_RR)$.

LEMMA 3.1. *For a ring R the following conditions are equivalent.*

(i) *R is right perfect and its injective hull $E(R_R)$ is projective, Σ-(quasi-) injective.*

(ii) *R is left perfect and its injective hull $E({}_RR)$ is projective, Σ-(quasi-) injective.*

REMARK. By the above note the "Σ-quasi-injective" and "Σ-injective" are coincident in Lemma 3.1.

Proof. We will only prove that (i) implies (ii) as the converse follows by symmetry.

Assume (i). Since R is right perfect, $E(R_R)$ has an indecomposable direct sum decomposition, $E(R_R)=\sum_{i=1}^{m}\oplus P_i$, where each P_i is injective projective right module. Let $R=e_1R\oplus\cdots\oplus e_nR$ for primitive idempotents e_i. Then there is an integer $\kappa(i)$ such that $P_i\cong e_{\kappa(i)}R$ for any $1\leqq i\leqq m$. Let $\{P_j\}_{j=1}^{s}$ be a subclass of mutually nonisomorphic projective modules of $\{P_i\}_{i=1}^{m}$ such that each $P_i(1\leqq i\leqq m)$ is isomorphic to some $P_j(1\leqq j\leqq s)$ (here, if need, the indecies are renumbered) and we put $M=P_1\oplus\cdots\oplus P_s$, then a right ideal $I=e_{\kappa(1)}R\oplus\cdots\oplus e_{\kappa(s)}R$ is isomorphic to M. Since M is clearly Σ-injective and faithful, so is then also I. Thus, by [4, Theorem 1.3], $E({}_RR)$ is projective, and R is left perfect and contains faithful, Σ-injective left ideal $\sum_{i=1}^{l}\oplus E(S_i)$, where $\{S_i\}_{i=1}^{l}$ is the representative class of simple left ideals which are nonisomorphic mutually and $E(S_i)$ an injective hull contained in R. As a consequence, $E({}_RR)$ is Σ-injective because $E({}_RR)$ is isomorphic to a submodule of a finite direct sum of copies of $\sum_{i=1}^{l}\oplus E(S_i)$. This completes the proof.

Now then, we suppose R is a (left and right) perfect ring such that $E(R_R)$ is projective and $E({}_RR)$ is not projective (for the existence of such a ring, see Müller [14] and Colby and Rutter [4]). Then, $E(R_R)=\sum_{i=1}^{m}\oplus P_i$, where each P_i is indecomposable injective for $1\leqq i\leqq m$ and, since the radical of every projective right module over a right perfect ring is small, any infinite family of modules each of which is isomorphic to some P_i is a T-nilpotent system ([12, Theorem 3]). On the other hand, an infinite direct sum $M=\bigoplus_{i\in I}M_i$ with

$M_i \cong E(R_R)$ is not quasi-injective by Lemma 3.1. Thus M is the desirable module having the exchange property by Proposition 2.3 and Theorem 2.6.

4. Applications.

We will generalize the theorems of Chamard [3, Théorème 3] and Yamagata [18, Theorem 4].

We recall definitions. A submodule N of a module M is said to be *closed* if it has no proper essential extension in M, that is, if $N \subseteqq' X$ for any submodule X of M, then $N = X$. A module M is said to be *well-complemented* in case any finite intersection of closed submodules of M is also closed.

LEMMA 4.1. *Let M be a direct sum of indecomposable injective modules $M_i (i \in I)$ and N a direct summand of M. If N is well complemented, then N is also a direct sum of indecomposable injective submodules.*

Proof. By [1, Theorem 1] it is clear N has a nonzero indecomposable injective submodule, so we can choose a maximal independent set $\{N_j\}_{j \in J}$ of indecomposable injective submodules of N. Put $N_0 = \sum_{j \in J} \oplus N_j$.

We will show $N = N_0$. To show this take an arbitrary nonzero element $x \in N$. Then there exists an injective hull $E(xR)$ of xR in N by [18, Lemma 2] and it is a finite direct sum of indecomposable injectives by [1, Theorem 1], say $E(xR) = E_1 \oplus \cdots \oplus E_n$. By the maximality of $\{N_j\}_{j \in J}$, it is evident that $N_0 \cap E_i \neq 0$ for $1 \leqq i \leqq n$. Then, since N is well-complemented by hypothesis, this will imply $E_i \subseteqq N_0$ for $1 \leqq i \leqq n$ and so $x \in E(xR) \subseteqq N_0$, which means $N = N_0$.

Because there exists a finite subset $\{j_1, \cdots, j_m\} \subseteqq J$ such that $\sum_{k=1}^{m} \oplus N_{j_k} \cap E_i \neq 0$ for $1 \leqq i \leqq n$. Since $\sum_{k=1}^{m} \oplus N_{j_k}$ and E_i are injective, they are closed in N and so is $\sum_{k=1}^{m} \oplus N_{j_k} \cap E_i$ by hypothesis of N for any $1 \leqq i \leqq n$. Then, since E_i is an essential extension of $\sum_{k=1}^{m} \oplus N_{j_k} \cap E_i$ by Lemma 2.2, it must be that $E_i = \sum_{k=1}^{m} \oplus N_{j_k} \cap E_i$ and therefore $E_i \subseteqq \sum_{k=1}^{m} \oplus N_{j_k}$ for any i. Consequently $x \in E(xR) \subseteqq \sum_{k=1}^{m} \oplus N_{j_k} \subseteqq N$, which concludes the lemma.

Under the same assumptions as in Lemma 4.1, we remark that N has no proper essential submodule which is a direct sum of indecomposable injectives from the proof of Lemma 4.1. This is first shown by Chamard [3, Lemma 4.1].

PROPOSITION 4.2. *Let M be a direct sum of indecomposable injective modules $M_i (i \in I)$ and N a direct summand of M; $M = N \oplus N'$. If N is well-complemented, then N has the exchange property* *and N and N' are also direct sums of indecomposable injective submodules.*

Proof. By Lemma 4.1, N is a direct sum of indecomposable injective submodules $N_j (j \in J)$. To show that N has the exchange property we will check the property (iii) in Theorem 2.6.

Let S be an endomorphism ring of N_R and J its Jacobson radical. We must show that $J = \{f \in S \mid \mathrm{Ker}\, f \subseteqq' N\}$. The inclusion $J \subseteqq \{f \in S \mid \mathrm{Ker}\, f \subseteqq' N\}$ is known in [2, p. 564]. Conversely take an arbitrary element $f \in S$ with $\mathrm{Ker}\, f \subseteqq' N$. To show that $f \in J$, it is enough to show that $1 - f$ is an isomorphism.

First we will prove that $1 - f$ is a monomorphism. If $\mathrm{Ker}\,(1 - f) \neq 0$, $xR \cap \mathrm{Ker}\, f \neq 0$ for any nonzero element $x \in \mathrm{Ker}\ (1 - f)$ since $\mathrm{Ker}\, f \subseteqq' N$. There is hence a nonzero element y of xR with $f(y) = 0$ and so $y = (1 - f)(y)$ which must imply $y = 0$, because $y \in \mathrm{Ker}\,(1 - f)$, a contradiction.

Next we will prove that $1 - f$ is an epimorphism. Since $1 - f$ is a monomorphism, $(1 - f)(N)$ is also a direct sum of indecomposable injectives. Take an arbitrary nonzero element $x \in N$. Then $xR \cap \mathrm{Ker}\, f \neq 0$, that is, there is a nonzero element $y \in xR \cap \mathrm{Ker}\, f$. We therefore have $xR \cap (1 - f)(N) \neq 0$, because $y = (1 - f)(y) \in xR \cap (1 - f)(N)$. This shows that $(1 - f)(N)$ is essential in N, so that $N = (1 - f)(N)$. Because N has no proper essential submodule which is a direct sum of indecomposable injectives by the remark just before this proposition.

Thus we have shown that N has the exchange property. We can then exchange N' for $\sum_{k \in K} \oplus M_k$ for some subset $K \subset I$, $M = N \oplus \sum_{k \in K} \oplus M_k$. This implies that $N' \cong \sum_{k \in K} \oplus M_k$, which completes the proof of the proposition.

Let M_R be any *nonsingular* module over a ring R, that is, $M \neq 0$ and if $xI = 0$ for $x \in M$ and essential right ideal I of R, then $x = 0$. It is then well known that the lattice of all closed submodules of M is complete and so M is clearly well-complemented (c.f., see [6, Corollary 8, p. 61]). Thus we can sharpen [18, Theorem 4] and [11, Proposition 4].

COROLLARY 4.3. *Let M, N, and N' be as above. If N is nonsingular, then it has the exchange property and so N and N' are also direct sums of indecomposable injective submodules.*

REFERENCES

1. G. Azumaya, *Corrections and supplementaries to my paper concerning Krull-*

Remak-Schmidt's theorem, Nagoya Math. J., **1** (1950), 117-124.
2. A. Cailleau and G. Renault, *Anneau associé à une somme directe infinite de modules quasi-injectifs*, Arch. der Math., **21** (1970), 561-566.
3. J. Y. Chamard, *Modules riches en co-irréductible et sous-modules compléments*, C. R. Acad. Sc. Paris, t. **264** (1967), 987-990.
4. R. R. Colby and E. A. Rutter, Jr., *Generalizations of QF-3 algebras*, Trans. Amer. Math. Soc., **153** (1971), 371-386.
5. P. Crawly and B. Jónsson, *Refinements for infinite direct decompositions of algebraic systems*, Pacific J. Math., **14** (1964), 797-855.
6. C. Faith, *Lectures on Injective Modules and Quotient Rings*, Lecture Notes in Mathematics, 49, Springer-Verlag, 1967.
7. L. Fuchs, *On quasi-injective modules*, Annali della Scuola Norm. Sup. Pisa, **23** (1969), 541-546.
8. K. R. Fuller, *On generalized uniserial rings and decompositions that complement direct summands*, Math. Ann., **200** (1973), 175-178.
9. M. Harada and Y. Sai, *On categories of indecomposable modules I*, Osaka J. Math., **7** (1970), 323-344.
10. M. Harada, *On categories of indecomposable modules II*, ibid., **8** (1971), 309-321.
11. ———, *Note on categories of indecomposable modules*, (to appear).
12. M. Harada and H. Kanbara, *On categories of projective modules*, Osaka J. Math., **8** (1971), 471-483.
13. H. Kanbara, *Note on Krull-Remak-Schmidt-Azumaya's theorem*, ibid., **9** (1972), 409-413.
14. B. J. Müller, *Dominant dimension of semi-primary rings*, J. Reine Angew Math., **232** (1968), 173-179.
15. R. B. Warfield, Jr., *A Krull-Schmidt theorem for infinite sums of modules*, Proc. Amer. Math. Soc., **22** (1969), 460-465.
16. ———, *Decompositions of injective modules*, Pacific J. Math., **31** (1969), 263-276.
17. ———, *Exchange rings and decompositions of modules*, Math. Ann., **199** (1972), 31-36.
18. K. Yamagata, *Nonsingular rings and Matlis' problem*, Sci. Rep. Tokyo Kyoiku Daigaku, Sect. A, **11** (1972), 114-121.
19. ———, *A note on a problem of Matlis*, Proc. Japan Acad., **49** (1973), 145-147.

Received July 17, 1973 and in revised form April 9, 1974.

TOKYO UNIVERSITY OF EDUCATION
OTSUKA, BUNKYO-KU, TOKYO, JAPAN

Pacific Journal of Mathematics

Vol. 55, No. 1 September, 1974