Pacific Journal of Mathematics

IN THIS ISSUE—

Vol. 2, No. 1 January, 1952

Vari-Type Composition by
Elaine Barth
Delores Gilbertson

With the cooperation of
E. F. Beckenbach
E. G. Straus

Printed in the United States of America by
Edwards Brothers, Inc., Ann Arbor, Michigan

UNIVERSITY OF CALIFORNIA PRESS * BERKELEY AND LOS ANGELES

THEOREMS ON GENERALIZED DEDEKIND SUMS

T. M. Apostol

1. Introduction. Generalized Dedekind sums $s_p(h,k)$, defined by

$$(1)\qquad s_p(h,k) = \sum_{\mu=1}^{k-1} \frac{\mu}{k} \overline{B}_p\left(\frac{h\mu}{k}\right) = \sum_{\mu=1}^{k-1} \frac{\mu}{k} B_p\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right]\right),$$

were introduced by the author [1]. The integers h and k are assumed relatively prime, $\overline{B}_p(x)$ is the p-th Bernoulli function, $B_p(x)$ the p-th Bernoulli polynomial (for definitions see [1;(2.11), (2.12)]), and $[x]$ is the greatest integer $\leq x$. For even values of the integer p the sums (1) are trivial (see [1;(4.13)]) and we assume in what follows that p is odd. These sums enjoy a reciprocity law, namely;

$$(2)\qquad \begin{aligned}&(p+1)\,(hk^p\, s_p(h,k) + kh^p\, s_p(k,h))\\ &\qquad = p\, B_{p+1} + \sum_{s=0}^{p+1} \binom{p+1}{s} (-1)^s\, B_s\, B_{p+1-s}\, h^s\, k^{p+1-s}.\end{aligned}$$

The B's being Bernoulli numbers*. An arithmetic proof of (2) is given in [1] by a method closely related to a general summation technique recently developed by Mordell [5]. When $p = 1$, the sums

$$(3)\qquad s_1(h,k) = \sum_{\mu=1}^{k-1} \frac{\mu}{k}\left(\frac{h\mu}{k} - \left[\frac{h\mu}{k}\right] - \frac{1}{2}\right)$$

are known as Dedekind sums and are usually denoted by $s(h,k)$. Aside from being of interest from an arithmetical standpoint, these sums also occur in the asymptotic theory of partitions and have been studied in a large number of papers, for example [1], [3], [5], [6], [7], [8], [9], [10], and [11].

In this paper we establish a connection between the sums (1) and certain finite sums involving Hurwitz zeta functions which makes it possible to give a short analytic proof of (2).

* When $p > 1$, the factor $(-1)^s$ may be suppressed in the summand in (2) because the terms corresponding to odd values of s vanish.

Received September 11, 1951.

Pacific J. Math. 2 (1952), 1-9

2. The Hurwitz zeta function and Dedekind sums. The Hurwitz zeta function $\zeta(s,a)$ is defined for $\Re(s)>1$ and $a \neq 0, -1, -2, \cdots$, by the series

$$\zeta(s,a) = \sum_{n=0}^{\infty} (n+a)^{-s}$$

and its relation to $s_p(h,k)$ is given in the following theorem.

THEOREM 1. *For odd* $p > 1$ *we have*

$$s_p(h,k) = i\ p!\ (2\pi ik)^{-p} \sum_{\mu=1}^{k-1} \cot \frac{\pi h\mu}{k}\, \zeta\left(p, \frac{\mu}{k}\right), \tag{4}$$

while for $p=1$ *we have the two equivalent expressions*

$$s(h,k) = \frac{1}{4k} \sum_{\mu=1}^{k-1} \cot \frac{\pi h\mu}{k} \cot \frac{\pi\mu}{k} \tag{5}$$

and

$$s(h,k) = \frac{-1}{2\pi k} \sum_{\mu=1}^{k-1} \cot \frac{\pi h\mu}{k} \frac{\Gamma'(\mu/k)}{\Gamma(\mu/k)}. \tag{6}$$

Formula (5) is due to Rademacher [8], who derived it from the Fourier series expansion of (3). We will give here a purely arithmetic proof of (5) based on *finite* rather than infinite Fourier series. Secondly, we establish the equivalence of (5) and (6) and then prove (4). Finally, we indicate how (5) and (6) can be thought of as limiting cases of (4).

Proof of (5): The function $\overline{B}_1(x)$ is given by

$$\overline{B}_1(x) = \begin{cases} x - [x] - 1/2 & \text{if } x \neq \text{integer}, \\ 0 & \text{otherwise}. \end{cases}$$

Therefore, by formula (2.5) of [10] we may write

$$s(h,k) = \sum_{\mu \bmod k} \overline{B}_1(\mu/k)\, \overline{B}_1(h\mu/k). \tag{7}$$

From Eisenstein's finite Fourier series expansion [4; p. 318] we have

$$\overline{B}_1(h\mu/k) = -\frac{1}{2k}\sum_{\nu=1}^{k-1} \sin\frac{2\pi h\nu\mu}{k}\cot\frac{\pi\nu}{k}. \tag{8}$$

Using (8) in each factor of the summand in (7), we obtain

$$s(h,k) = \frac{1}{8k^2}\sum_{\lambda=1}^{k-1}\sum_{\nu=1}^{k-1}\cot\frac{\pi\nu}{k}\cot\frac{\pi\lambda}{k}\sum_{\mu \bmod k}\left(\cos\frac{2\pi\mu(\lambda-\nu h)}{k} - \cos\frac{2\pi\mu(\lambda+\nu h)}{k}\right) \tag{9}$$

because of the identity $2\sin x \sin y = \cos(x-y) - \cos(x+y)$. Since we have

$$\sum_{\mu \bmod k} \cos\frac{2\pi\mu(\lambda\pm\nu h)}{k} = \begin{cases} k & \text{if } \lambda \pm \nu h \equiv 0 \pmod{k}, \\ 0 & \text{otherwise}, \end{cases}$$

for each fixed ν only one value of λ gives a nonzero contribution to each sum in the second member of (9), namely $\lambda \equiv \nu h \pmod{k}$ in the first sum and $\lambda \equiv -\nu h \pmod{k}$ in the second. Therefore we have

$$s(h,k) = \frac{1}{8k}\sum_{\nu=1}^{k-1}\cot\frac{\pi\nu}{k}\cot\frac{\pi h\nu}{k} - \frac{1}{8k}\sum_{\nu=1}^{k-1}\cot\frac{\pi\nu}{k}\cot\frac{-\pi h\nu}{k},$$

and this is the same as (5).

Proof that (5) *and* (6) *are equivalent:* The relation [2; p. 163]

$$\frac{\Gamma'(\mu/k)}{\Gamma(\mu/k)} = -\gamma - \log k - \frac{\pi}{2}\cot\frac{\pi\mu}{k} + {\sum_{n \le k/2}}' \cos\frac{2\pi n\mu}{k}\log\left(2 - 2\cos\frac{2\pi n}{k}\right), \tag{10}$$

where γ is Euler's constant and the prime indicates that when k is even the last term is to be multiplied by $1/2$, is due to Gauss. Multiplying both sides of (10) by $\cot(\pi h\mu/k)$ and summing on μ shows the equivalence of (5) and (6) upon observing that we have

$$\sum_{\mu=1}^{k-1} f(\mu) = 0 \tag{11}$$

whenever f is an odd function of μ which is also periodic mod k.

Proof of (4): Formula (4.11) of [1] gives a representation of $s_p(h,k)$ as an infinite series which, with some simplification, can be written in the form

$$s_p(h,k) = i\; p!\; (2\pi i)^{-p} \sum_{\substack{n=1 \\ n \not\equiv 0 \pmod{k}}}^{\infty} n^{-p} \cot \frac{n\pi h}{k}.$$

Writing $n = qk + \mu$, with $q = 0, 1, 2, \cdots, \infty$, and $\mu = 1, 2, \cdots, k-1$, we obtain

$$s_p(h,k) = i\; p!\; (2\pi i)^{-p} \sum_{\mu=1}^{k-1} \sum_{q=0}^{\infty} (qk+\mu)^{-p} \cot \frac{\pi h \mu}{k}$$

$$= i\; p!\; (2\pi i k)^{-p} \sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \zeta(p, \mu/k),$$

where we must assume $p > 1$ in order to insure that the series involved should be absolutely convergent and the rearrangements valid. This proves (4). We cannot hope for a proof of (4) along the lines of our proof of (5) because of the non-elementary nature of $\zeta(s,a)$.

If in (4) we replace $p!$ by $\Gamma(p+1)$ and let p be a complex variable which tends to 1, then we can show that the two expressions for $s(h,k)$ in (5) and (6) occur naturally as limiting cases of the right member of (4). We first observe that, although the function $\zeta(s,a)$ has a pole at $s = 1$, the sum

$$\sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \zeta(s, \mu/k) \tag{12}$$

is *regular* at $s = 1$. This is easily seen by using the expansion

$$\zeta(s,a) = \frac{1}{s-1} - \frac{\Gamma'(a)}{\Gamma(a)} + O(s-1) \qquad (\text{as } s \longrightarrow 1)$$

obtained from Whittaker and Watson [12; p. 271], substituting in (12) and using (11) to obtain

$$\lim_{s \to 1} \sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \zeta(s, \mu/k) = - \sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \frac{\Gamma'(\mu/k)}{\Gamma(\mu/k)},$$

which shows that the right member of (4) tends to the right member of (6) as $p \to 1$.

The connection between (5) and (4) can be obtained by using Hurwitz's functional equation in the form given by Rademacher [6; (1.24)], namely:

$$\zeta(s, \mu/k) = 2\Gamma(1-s)(2\pi k)^{s-1} \sum_{\lambda=1}^{k} \left(\cos \frac{\pi s}{2} \sin \frac{2\pi\lambda\mu}{k} \right.$$
$$\left. + \sin \frac{\pi s}{2} \cos \frac{2\pi\lambda\mu}{k} \right) \zeta\left(1 - s, \frac{\lambda}{k}\right),$$

this being valid for $s \neq 1$, $1 \leq \mu \leq k$. Multiplying by $\cot(\pi h \mu/k)$, summing on μ and using (11) leads to

$$\sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \zeta(s, \mu/k)$$
(13)
$$= 2\Gamma(1-s)(2\pi k)^{s-1} \cos \frac{\pi s}{2} \sum_{\lambda, \mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \sin \frac{2\pi\lambda\mu}{k} \zeta(1-s, \lambda/k).$$

Since $\zeta(0, a) = 1/2 - a$, when s tends to 1 the right member of (13) approaches the value

$$\frac{1}{2k} \sum_{\lambda=1}^{k-1} \sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \sin \frac{2\pi\lambda\mu}{k} \left(\frac{1}{2} - \frac{\lambda}{k} \right)$$
$$= \frac{-1}{2k^2} \sum_{\mu=1}^{k-1} \cot \frac{\pi h \mu}{k} \sum_{\lambda=1}^{k-1} \lambda \sin \frac{2\pi\lambda\mu}{k}$$

because of (11). Noticing that the last sum on λ is the imaginary part of the sum

$$\sum_{\lambda=1}^{k-1} \lambda e^{2\pi i \lambda\mu/k} = k/(e^{2\pi i \mu/k} - 1) = - \frac{ki}{2} \cot(\pi\mu/k) - \frac{k}{2},$$

we see that the right member of (4) also tends to the right member of (5) when $p \longrightarrow 1$.

3. **Proof of the reciprocity law.** We can now give a proof of the reciprocity law (2) using complex integration. This proof is of additional interest in that we use properties of $\zeta(s,a)$ for fixed s and variable a. We will need the following facts about $\zeta(s,a)$:

(14) $$\zeta(s,a) = \zeta(s,a+1) + a^{-s},$$

(15) $$\zeta(s,a+1) = \zeta(s) - s\,\zeta(s+1)\,a + O(a^2) \qquad as\ a \longrightarrow 0,$$

(16) *for* $0 \leq \Re(a) \leq M$, *(M fixed),* $\zeta(s,a)$ *tends uniformly to* 0 *as* $\Im(a) \to \pm\infty$. *(The uniformity is with respect to* $\Re(a)$*).*

Equation (14) follows at once from the definition of $\zeta(s,a)$ and (15) is merely the beginning of the Taylor series for $\zeta(s,a+1)$ near $a = 0$. Here $\zeta(s) = \zeta(s,1)$ is Riemann's zeta function. Relation (16) can be readily obtained, for example, by applying the Riemann-Lebesgue theorem to the integral representation [2; p. 266]:

$$\Gamma(s)\,\zeta(s,a) = \int_0^\infty \frac{t^{s-1}\,e^{-t\,\Re(a)}}{1-e^{-t}}\;e^{-it\,\Im(a)}\,dt,$$

valid for $\Re(s) > 1$ and $\Re(a) > 0$. This gives (16) for $0 < \Re(a) \leq M$ and (14) proves it for $0 \leq \Re(a) \leq M$.

Because of (4), the reciprocity formula (2) can now be put into the following form:

THEOREM 2. *For odd* $p > 1$ *we have*

(17) $$\frac{i(p+1)!}{(2\pi i)^p}\left\{h\sum_{\mu=1}^{k-1}\cot\frac{\pi h\mu}{k}\,\zeta(p,\mu/k) + k\sum_{\nu=1}^{h-1}\cot\frac{\pi k\nu}{h}\,\zeta(p,\nu/h)\right\}$$
$$= p\,B_{p+1} + \sum_{s=0}^{p+1}\binom{p+1}{s}B_s\,B_{p+1-s}\,h^s\,k^{p+1-s}.$$

Proof. We apply Cauchy's residue theorem to the function

$$f(z) = \cot\pi z\,\cot(\pi hz/k)\,\zeta(p,z/k),$$

Integrating in the positive sense around a contour C consisting of a rectangle whose vertices are the points $\pm iT$, $k \pm iT$, with small semi-circular detours C_0 and C_k around the points $z = 0$, $z = k$, traversed along the arcs $z = \epsilon e^{i\theta}$ and $z = k + \epsilon e^{i\theta}$, respectively, where $\pi/2 \leq \theta \leq 3\pi/2$, and $0 < \epsilon < 1/h$. Ultimately, ϵ will tend to 0 and $T > 1/2$ will tend to ∞. The integrand $f(z)$ has first order poles at the points $z = 1, 2, \cdots, k-1$ due to the factor $\cot \pi z$, and at the points $z = k/h, 2k/h, \cdots, (h-1)k/h$ because of the factor $\cot(\pi hz/k)$. By (14) we have

$$\zeta(p, z/k) = \zeta(p, z/k + 1) + (k/z)^p,$$

so that the point $z = 0$ is a pole of order $p + 2$ for $f(z)$. Using the power series expansion

$$\pi z \cot \pi z = \sum_{n=0}^{\infty} \frac{(2\pi i)^n B_n}{n!} z^n \tag{18}$$

in the neighborhood of $z = 0$ (with the understanding that B_1 should be replaced by 0), and (15) with $a = z/k$ we find that Cauchy's theorem gives us

$$\begin{aligned}\frac{1}{2\pi i}\int_C f(z)\,dz &= \frac{1}{\pi}\sum_{\mu=1}^{k-1} \cot\frac{\pi h\mu}{k}\,\zeta(p, \mu/k) \\ &+ \frac{k}{\pi h}\sum_{\nu=1}^{h-1} \cot\frac{\pi k\nu}{h}\,\zeta(p, \nu/h) - \frac{p}{\pi^2 h}\,\zeta(p+1) + R_0,\end{aligned} \tag{19}$$

where

$$R_0 = \operatorname*{Res}_{z=0} \cot \pi z \cot(\pi hz/k)\,(k/z)^p.$$

We now observe that by periodicity of the cotangent and by (14), the contribution to the integral from the part of C consisting of vertical line segments is

$$\left(\int_{iT}^{i\epsilon} + \int_{-i\epsilon}^{-iT}\right) \cot \pi z \cot(\pi hz/k)\,(k/z)^p\,dz,$$

and this vanishes since the integrand is an odd function of z. Next, the integrals along the horizontal segments tend to zero as $T \longrightarrow \infty$ since, for $0 \leq x \leq k$ we have $\cot \pi(x + iy) \longrightarrow \mp i$ and, by (16), $\zeta(p, (x + iy)/k)$ tends to 0 uniformly in x as $y \longrightarrow \pm\infty$. Finally, combining the integrals over C_0 and C_k by means of (14) and letting $T \longrightarrow \infty$ we obtain

$$\lim_{T\to\infty} \int_C f(z)\, dz = \int_{C_0} \cot \pi z \cot(\pi hz/k)\,(k/z)^p\, dz.$$

When $\epsilon \longrightarrow 0$ we find

$$\lim_{\epsilon\to 0} \int_{C_0} = \pi i\, R_0$$

so that equation (19) leads to the result

$$\frac{1}{\pi}\sum_{\mu=1}^{k-1} \cot\frac{\pi h\mu}{k}\,\zeta(p,\mu/k) + \frac{k}{\pi h}\sum_{\nu=1}^{h-1} \cot\frac{\pi k\nu}{h}\,\zeta(p,\nu/h) \tag{20}$$
$$= \frac{p}{\pi^2 h}\,\zeta(p+1) - \frac{1}{2}R_0.$$

From (18) we easily calculate that

$$R_0 = \frac{2i\,(2\pi i)^p}{\pi h\,(p+1)\,!}\sum_{s=0}^{p+1}\binom{p+1}{s} B_s\, B_{p+1-s}\, h^s\, k^{p+1-s},$$

and, since we have

$$\zeta(p+1) = -\frac{(2\pi i)^{p+1}\,B_{p+1}}{2(p+1)\,!},$$

equation (20) yields (17) and the proof is complete.

In [8], Rademacher gives a proof for the case $p = 1$ using (5) instead of (4). Apparently unaware of [8], K. Iseki [3] has given a proof very much like Rademacher's analytic proof for the case $p = 1$ in a recent paper.

References

1. T. M. Apostol, *Generalized Dedekind sums and transformation formulae of certain Lambert series*, Duke Math. J., 17 (1950), 147-157.

2. G. Brunel, *Bestimmte Integrale*, Encyklopädie der Mathematischen Wissenschaften, II A 3, 135-188.

3. Kanesiroo Iseki, *Analytic proof for the reciprocity law of Dedekind sums*, Sugaku, 2, no. 3 (1950), 240-241. (Japanese)

4. G. B. Mathews, *Theory of Numbers*, Cambridge, 1892.

5. L. J. Mordell, *The reciprocity formula for Dedekind sums*, Amer. J. Math., **73** (1951), 593-598.

6. Hans Rademacher, *Zur Theorie der Modulfunktionen*, J. Reine Angew. Math. **167** (1932), 312-336.

7. ———, *Eine arithmetische Summenformel*, Monatsh. Math. Phys., **39** (1932), 221-228.

8. ———, *Egy reciprocitásképletröl a modülfuggvények elméletéböl*, Mat. Fiz. Lapok, **40** (1933), 24-34.

9. ———, *Die Reziprozitätsformel für Dedekindsche Summen*, Acta Univ. Szeged. Sect. Sci. Math., **12**, B (1950), 57-60.

10. ——— and Albert Whiteman, *Theorems on Dedekind sums*, Amer. J. Math., **63** (1941), 377-407.

11. L. Rédei, *Elementarer Beweis und Verallgemeinerung einer Reziprozitätsformel von Dedekind*, Acta Univ. Szeged. Sect. Sci. Math., **12**, B (1950), 236-239.

12. E. T. Whittaker and G. N. Watson, *Modern Analysis*, University Press, Cambridge, 1945.

California Institute of Technology

ADDENDUM TO 'ON THE LERCH ZETA FUNCTION'

T. M. APOSTOL

Professor L. Carlitz has been kind enough to point out that the functions $\beta_n(a, \alpha)$ which were used in [1] to evaluate the Lerch zeta function $\phi(x, a, s)$ for negative integer values of s have occurred elsewhere in the literature in other connections, for example in [2] and [3]. As Carlitz points out, formula (3.3) of [1] leads to the result

$$\alpha^m \ \beta_n(m, \alpha) - \beta_n(0, \alpha) = n \sum_{a=0}^{m-1} a^{n-1} \ \alpha^a$$

which, for integer values of the variable a, makes apparent the relation of the functions $\beta_n(a, \alpha)$ with the Mirimanoff polynomials discussed by Vandiver in [3].

There is a misprint in the next to last equation on p. 164 of [1]. The coefficient of $a^2/2$ in the expression for $\phi(x, a, -2)$ should read $i \cot \pi x + 1$ instead of $i \cot \pi x + 1/4$.

REFERENCES

1. T. M. Apostol, *On the Lerch zeta function*, Pacific J. Math., **1**, (1951), 161-167.

2. L. Euler, *Institutiones calculi differentialis*, (II) 487-491.

3. H. S. Vandiver, *An arithmetical theory of the Bernoulli numbers*, Trans. Amer. Math. Soc., **51** (1942), 506.

4. G. Frobenius, *Uber die Bernoulli'schen Zahlen und die Euler'schen Polynome*, Sitzungsber. Akad. Wissensch. Berlin, **2** (1910), 826.

CALIFORNIA INSTITUTE OF TECHNOLOGY

Received November 5, 1951.
Pacific J. Math. 2 (1952), 10.

EXTENSION OF FUNCTIONS ON FULLY NORMAL SPACES

RICHARD ARENS

1. **Introduction.** Starting from the recent discovery of A. H. Stone that metric spaces are "paracompact" [12] (paracompactness means that every open covering has a refinement only a finite number of whose members meet a suitable neighborhood of each point [5]), J. Dugundji has been able to extend to metric spaces certain techniques in the theory of retracts which were hitherto applicable at most to separable metric spaces [6]. The cornerstone of his method is a theorem (see 2.4, below) according to which a continuous function on a closed set A of a metric space X with values in a convex (= "locally convex") topological linear space L may be extended to the whole space X, indeed without enlarging the convex hull of the image. Essentially, the possibility of doing this for a locally separable metric space X is implicitly given by a procedure for the real valued case in [10].

One of the problems to which we address ourselves in this paper is that of determining whether the assumption that X is metric can be reduced to X is merely paracompact. The answer (see 6, below) is *no*. However, we have fairly general results which imply that if L is metric and complete (and X is paracompact) then the extension is possible (4.1, below). Our proof utilizes a process of extending a pseudo-metric on A to all of X, which is ultimately based on a theorem of Hausdorff. We generalize Hausdorff's theorem (3.2 and 3.4) and incidentally show how Dugundji's result enables one to construct a short proof of Hausdorff's theorem.

None of these extension theorems can properly be regarded as a true generalization of Tietze's extension theorem, which deals with mappings on normal spaces with values on the line or in the Hilbert cube, since there exist normal, not fully normal spaces. In order to provide a generalization of Tietze's theorem, we have shown by way of application that the Hilbert cube may be replaced by any compact convex subset of a normed linear space (4.3).

S. Kakutani [10] has introduced the notion of "simultaneous extension regarded as a linear positive operation," in his case of real valued functions on a locally separable metric space X: this means that it is possible so to extend

Received March 10, 1951.

Pacific J. Math. 2 (1952), 11-22

all continuous functions such that $(\lambda f + g)^e = \lambda f^e + g^e$ and $f^e \geq 0$ whenever $f \geq 0$, where the superscribed e indicates the extension. We show that this is possible in the more general case in which X is metric and L convex, the order preserving feature of Kakutani's formulation being naturally reformulated as a nonenlargement of the convex hull of the image (2.6). What is perhaps more surprising is that the "simultaneous extension," while possible in more general cases (4.2), is not possible under as general conditions as those under which the individual extension (as in 4.1) is possible. As a matter of fact, we tie up the notion of simultaneous linear and order relation preserving extension with that of "sweeping" a measure from X onto A, and thus show that it is not always possible even for compact Hausdorff spaces X (6.1).

The question arises whether among the simultaneous extensions which preserve linear and order relations, there are not some which preserve quadratic polynomial relations as well. It has already been shown by Yoshizawa just when this is possible: at least when X is compact, A must be a retract of X.

We have inserted a section (5) showing that the "simultaneous extension" for real valued functions can be derived from the "individual" extension of a suitable continuous function with values in an infinite dimensional space, as well as from the fundamental Lemma (2.1) directly.

In formulating our results, we shall always speak of "fully normal" [13] rather than "paracompact" spaces, although it is known that these two properties coincide [12]. We do this because we use the full normality as such, using Stone's result only for the metric case.

2. **Extension of functions on metric space.** One of the main geometric ideas underlying the process of extension involved here is contained in the following construction (cf. [6, 4.3]).

2.1. LEMMA. *Let X be a metric space, and A a closed subset of X. Then there exists a family g_V of continuous real-valued functions defined on X, and a similarly indexed family of points a_V of A such that*

2.2. *each g_V vanishes on A, all but a finite number of the g_V vanish in some neighborhood of each point of X, the sum $\sum_V g_V(x) = 1$ for all x in $X - A$ and each $g_V(x)$ is nonnegative;*

2.3. *for each a in A and each V, if $g_V(x) > 0$ then $d(a, a_V) < 3\, d(a, x)$ (d is the metric in X) and $d(a, x) < d(a, a_V) + 2\, d(x, A)$.*

Proof. To each x in $X - A$, assign the open sphere of radius $d(x, A)/4$. Since $X - A$ is metric it is paracompact [12], so that this covering of $X - A$ has a refinement R which is "neighborhood finite" (see [12]). For each V in R there is a point x such that V is contained in the open sphere about x of radius $d(x,A)/4$; select such a point and call it x_V. Also select a_V in A such that $d(a_V, x_V) < (5/4)\, d(x_V, A)$. Let $f_V(x) = d(x, X - V)$ for every x in X. Since R is neighborhood finite, each x in X has a neighborhood on which all but finitely many f_V vanish, and so $s(x) = \sum_{V \in R} f_V(x)$ is both finite and continuous. Since $s(x)$ is never 0, the functions $g_V = f_V/s$ are continuous. It is easy to see that they provide the kind of "Dieudonné partition" required by 2.2.

We now turn to 2.3. Suppose $g_V(x) > 0$. Then x belongs to the open set V. From what has been said about V, x_V, and a_V follows

$$d(a_V, x) \leq d(a_V, x_V) + d(x_V, x) < (5/4)\, d(x_V, A) + (1/4)\, d(x_V, A).$$

Of course

$$\begin{aligned} d(x_V, A) &\leq d(x_V, a) \\ &\leq d(x_V, x) + d(x, a) < (1/4)\, d(x_V, A) + d(x, a), \end{aligned}$$

and so $(3/4)\ d(x_V, A) < d(x, a)$. Thus $d(a_V, x) < (6/4)\, d(x_V, A) < 2\, d(x, a)$. Thus finally we have half of 2.3, since

$$d(a, a_V) \leq d(a, x) + d(x, a_V) < d(a, x) + 2\, d(x, a) = 3\, d(a, x).$$

For the second half of 2.3, we note first that $d(x, A) > 3/4\, d(x_V, A)$. On the other hand,

$$d(a_V, x) \leq d(a_V, x_V) + d(x_V, x) < (6/4)\ d(x_V, A),$$

which is thus less than $2\ d(x, A)$. Finally,

$$d(a, x) \leq d(a, a_V) + d(a_V, x) < d(a, a_V) + 2\ d(x, A).$$

Thus the proof of 2.3 is complete.

Geometrically, the lemma given above says that $X - A$ can be so mapped into the finite dimensional faces of the "formal simplex" with vertices equal to the points of A, in such a way that as x tends to a point a_0 of A, the vertices of the carrier of the image of x all tend to a_0, in the topology of A. With this picture in mind, it is easy to imagine how functions on A with values in a convex set E, can be extended to all of X. The next result [6, 4.1] makes this precise.

2.4. THEOREM. *Let K be a convex subset of a convex topological linear space L* (cf. [14]). *Let A be a closed subset of a metric space X. Let f be continuous on A with values in K. Then f may be extended continuously to a function f^e defined on X with values in K.*

Proof. Using the result and notation of 2.1, we define f^e at once by setting $f^e(x) = \sum_V g_V(x) f(a_V)$, for x in $X - A$, and $f^e(a) = f(a)$ for a in A. There remains only the proof of continuity. Now the topology of L (and thus K) can be based on neighborhoods of 0 defined by relations $||k|| < 1$ where $||\cdots||$ is one of the pseudo-norms of L, according to von Neumann's idea (cf. [14]). Select any point a of A. There is a positive r such that $d(a, b) < r$ implies $|| f(a) - f(b)|| < 1$ for b in A. Now suppose $d(a, x) < r/3$. For those finitely many V for which $g_V(x)$ is not 0, we have $d(a, a_V) < r$, so that

$$||f^e(x) - f(a)|| \leq \sum g_V(x) ||f(a_V) - f(a)|| < 1.$$

This shows the continuity of f^e at any point of A. At points x of $X - A$ we can find a neighborhood in which only finitely many g_V do not vanish, so that f^e is continuous there also. The rest of 2.4 is obvious. The second half of 2.3 is not needed for this proof.

The fact that a single formula, so to speak, can be chosen to perform the extension can be expressed in several ways. Suppose K_1 and K_2 are convex subsets of two convex topological linear spaces, and let there be an affine mapping m of K_1 into K_2. Suppose f_1, f_2 are functions as in 2.4 with values in K_1, K_2 respectively, but satisfying the condition $m(f_1(a)) = f_2(a)$ for all a in A. If we use the same system g_V, a_V in extending f_2 as in extending f_1 then we surely obtain $m(f_1^e(x)) = f_2^e(x)$ for all x in X. We shall abbreviate this by saying that the process of extension when applied to all possible f is *consistent*, and note the result:

2.5. THEOREM. *Each f satisfying the hypothesis of* 2.4 *with K variable but A and X constant may be so extended that the entire process is consistent.*

Another kind of consistency or simultaneity is expressed as follows.

2.6. THEOREM. *Let K be a* (*linear or possibly merely convex*) *subset of a convex topological linear space, and let A be a closed subset of the metric space X. Let F be the class of continuous functions on A with values in K. Then each f may be extended by an f^e*(*using* 2.4) *in such a way that, for $f_1, \cdots, f_n$ in F and $c_1, \cdots, c_n$ real numbers* (*nonnegative with sum* 1 *when K is merely* *convex), we have*

$$(c_1 f_1 + \cdots + c_n f_n)^e = c_1 f_1^e + \cdots + c_n f_n^e .$$

This result is a generalization of Kakutani's theorem [10] on "simultaneous extension of continuous functions considered as a positive linear operation". The only real advance of 2.6 over Kakutani's theorem is the removal of separability, although Kakutani limits F to the space of bounded real valued continuous functions $C(A)$.

An addendum to 2.4 and 2.6 is of interest:

2.7. *Under the conditions of* 2.4 *or* 2.6, *if there is an* f *and a point* a *of* A *such that* f *is constant on a neighborhood* (*relative to* A) *of* a, *then* f^e *is constant on a neighborhood* (*relative to* X) *of* a.

In fact, suppose $f(a')$ is constant for $d(a, a') \leq 3e$ and a' in A. Then $f^e(x)$ is constant for $d(a, x) \leq e$, since then $d(a, a_V) \leq 3e$.

3. **Extension of pseudo-metrics.** Let X be a topological space. Let s be a real-valued function of two variables defined in X such that

$$s(y, x) = s(x, y) \geq 0, \; s(x, z) \leq s(x, y) + s(y, z);$$

and such that the set of x such that $s(x, y) < e$ is open for each $e > 0$ and y. Then s is a *pseudo-metric*. It falls short of being a metric in that $s(x_n, y) \longrightarrow 0$ $[s(x, y) = 0]$ does not necessarily imply $x_n \longrightarrow y$ $(x = y)$. Our first result is in the direction of an extension of a pseudo-metric from a closed set to the whole space.

3.1. LEMMA. *Let* X *be a fully normal* [13] *topological space, and let* q *be a pseudo-metric defined on a closed subset* A *of* X. *Then there is a pseudo-metric* s *defined in all of* X *such that for* x, y *in* A *and* $k = 4, 5, \cdots$, *if* $s(x, y) < 2^{-k}$ *then* $q(x, y) < 2^{-k}$.

Proof. Select a positive integer n. Construct an open covering U consisting of those open sets V which intersect A in a set of q-diameter less than 2^{-n}. Using the terminology, notation, and results of [13] we obtain $U \overset{*}{>} U_1 \overset{*}{>} U_2 \overset{*}{>} \cdots$ [13, V–7.4], and a pseudo-metric r_n such that [13, V–7.5, correcting $\in$ to $\notin$] $x \notin S(y, U_p)$ implies $r_n(x, y) > 2^{-(p+2)}$ for $p = 1, 2, \cdots$. We may also assume $r_n(x, y) \leq 1$. We can thus form $r(x, y) = \sum_n 2^{-n} r_n(x, y)$. This is clearly a pseudo-metric. Suppose $r(x, y) < 2^{-k}$, for $k \geq 4$ and x, y in A. Then $r_{k-3}(x, y) < 2^{-3}$. Hence $x \in S(y, U_1)$, this covering being the one obtained for $n = k - 3$.

(We have omitted an index showing dependence on n.) Since $U_1 \overset{*}{<} U$ we obtain $q(x, y) < 2^{-k+3}$. Setting $s = 2^{-3} r$ gives the required pseudo-metric, completing the proof of 3.1.

We remind the reader that compact Hausdorff spaces and metric spaces are fully normal [13, V–8.14, VI–4.5].

This lemma is actually all we need in order to extend the results of §2 as we shall do later. However, by an application of a theorem of Hausdorff, we can improve 3.1 aesthetically by obtaining a pseudo-metric s which is an extension to X of the original q. In fact, rather than refer to Hausdorff's theorem, we first give a new proof since it is an interesting application of 2.1, is much shorter than Hausdorff's, and shows in passing how a metric space may be isometrically imbedded in its own space of bounded continuous functions (cf. [4, p.187]). The present proof resembles that in [11] more than that in [9]. However, Kuratowski's proof, besides requiring separability, generally does not provide an isometric, but merely topological imbedding (see below, and also [9, p.47]).

3.2. THEOREM. [Hausdorff]. *Let A be a closed subset of a metric space X, and let f be a continuous mapping of A into another metric space B. Then B can be isometrically imbedded in a metric space Y such that f can be continuously extended to X with values in Y, such that f is a homeomorphism of $X - A$ with $Y - B$, and such that B is closed in Y.*

Proof: For any space S let $C(S)$ denote the Banach space of real-valued continuous bounded functions g on S, with $||g|| = \sup_{x \in S} |g(x)|$.

To begin the proof, obtain for X a bounded metric d. The metric r in B we must not alter, of course. For b in B, let r_b denote the function with values $r_b(b') = r(b, b')$. This function is not necessarily bounded, but $r_b - r_c$ is bounded (cf. [4, p.187]) and $||r_b - r_c|| = r(b, c)$, where b, c are points of B. Select a point o in A, to be held constant. The function ϕ defined for a in A by $\phi(a) = r_{f(a)} - r_{f(o)}$ evidently maps A into $C(A)$. Indeed, since

$$||\phi(a) - \phi(a')|| = ||r_{f(a)} - r_{f(a')}|| = r(f(a), f(a')),$$

the map ϕ is continuous. It may therefore by extended to all of X by 2.4, and we denote the extension also be ϕ. Now form $L = C(B) \times R \times C(X)$, where

$$||(h, j, k)|| = \max(||h||, |j|, ||k||)$$

and R is the real number system. For x, y in X define $d_x(y) = d(x, y)$ as earlier, and let $d(x) = d(x, A)$. For x in X, define $F(x) = [\phi(x), d(x), d(x)\, d_x]$ in L.

This F is obviously continuous. Define $B_1 = F(A)$ and $Y = F(X)$, both subsets of L. Clearly B_1 is closed relative to Y. Now for each a in A we obtain $f(a)$ on one hand and $F(a)$ on the other. We now show that it sets up an isometry between B and B_1. In fact,

$$||F(a) - F(a')|| = ||\phi(a) - \phi(a')|| = ||r_{f(a)} - r_{f(a')}|| = r[f(a), f(a')],$$

as mentioned earlier. If we identify B with B_1, then F becomes an extension of f, continuous on all of X. Suppose $F(x) = F(y)$, where y belongs to $X - A$. Then $d(x) = d(y) > 0$; hence $d_x = d_y$, which means $x = y$. Thus F has an inverse on $Y - B_1$. We shall show that it is an homeomorphism. Let $y \in X - A$ and suppose $F(x) \longrightarrow F(y)$. Then $d(x) \longrightarrow d(y) > 0$, and $d(x)\, d_x \longrightarrow d(y)\, d_y$. From this we conclude $d_x \longrightarrow d_y$ or $d(x, y) = ||d_x - d_y|| \longrightarrow 0$. Thus 3.2, Hausdorff's theorem, is proved. It is to be borne in mind that it was not known in 1938 that metric spaces were paracompact.

We go on to establish a refinement of 3.2 also due to Hausdorff.

3.3. THEOREM. *If the f in 3.2 is a homeomorphism of A with B then it can be arranged that F also is a homeomorphism.*

To establish 3.3, Hausdorff [9, p.46] modifies the construction of F. It is an interesting fact that the F we construct automatically satisfies 3.3. The only nontrivial part of the proof of 3.3 is that if $F(x) \longrightarrow F(a)$ for x in $X - A$ and a in A, then $x \longrightarrow a$ in X. Therefore, suppose $F(x) \longrightarrow F(a)$. Let

$$h = g(x) - g(a) = \sum g_V(x)\, [r_{f(a_V)} - r_{f(a)}],$$

where the g_V and a_V are described in 2.1. Now $|h(f(a))| \leq ||h|| \longrightarrow 0$. But

$$h[f(a)] = \sum g_V(x)\, r[f(a_V), f(a)]$$

is not less than the least of those $r[f(a_V), f(a)]$ which appear in this sum, that is, for which $g_V(x)$ is not 0. Denote the a_V in question by a_W, where of course W depends on x. Since $r(f(a_W), f(a)]$ tends to 0 and f is a homeomorphism on A, we see that $a_W \longrightarrow a$. From 2.3 we obtain $d(a, x) < d(a, a_W) + 2\,d(x)$, and so $x \longrightarrow a$, as desired.

These two results have the following consequence.

3.4. COROLLARY. *Let A be a closed subset of a metric space X. Let r be a pseudo-metric defined on A. Then this pseudo-metric may be extended to all of X in such a way that*

3.41 *in* $X - A$ *it is equivalent to the metric* d *of* X;

3.42 *in* $x \in X - A$ *then for some positive* e, $r(x, y) < e$ *implies* $y \in X - A$;

3.43 *if* r *is a metric equivalent to* d *on* A, *the extension is equivalent to* d *on all of* X.

Proof. In A, form the equivalence classes for the relation $r(x, y) = 0$, and metrize in the obvious way using r. Call the resulting space B. The natural mapping of A onto B satisfies the hypothesis of 3.2. Let m be the metric in Y. Then $m[F(x), F(y)]$ gives the desired extension of $r(x, y)$.

We can now provide the finishing touch to 3.1.

3.5. THEOREM. *Let* X *be a fully normal topological space, and let* q *be a pseudo-metric defined on a closed subset of* X. *Then* q *can be extended to be a pseudo-metric on* X.

The proof is based on 3.1 and 3.4 as follows. Using the s of 3.1, partition X into a set X^* of equivalence classes according to the relation $s(x, y) = 0$, denoting the class containing x by x^*, and so on. Define $s^*(x^*, y^*) = s(x, y)$; this is a valid definition, which makes X^* a metric space, and the natural mapping of X onto X^* is continuous. Let A^* be the closure in X^* of the image of A. The conclusion of 3.1 shows that q may be carried over in unique fashion to A^*, to form a pseudo-metric q^*. An appeal to 3.4 extends q^* to X^*, and $q(x, y) = q^*(x^*, y^*)$ provides the desired extension.

Note that we have no use for 3.41 - 3.43 in 3.5 because the s was not given to us in advance.

4. **Extension of functions on fully normal spaces.** In the next result, the metric for X in 2.4 is shifted to K.

4.1. THEOREM. *Let* A *be a closed subset of a fully normal space* X. *Let* f *be continuous on* A *with values in a complete convex metric subset* K *of a convex topological linear space* L. *Then* f *can be continuously extended to* X *with all values still in* K.

Proof. In A define the pseudo-metric $q(a, a') = m[f(a), f(a')]$, where m is the metric, and extend q to X by 3.5. *Let* A_0 be the set of x such that $q(x, A) = 0$. Given $e > 0$ and x in A_0, let S_e be the set of a in A such that $q(x, a) < e$. The $f(S_e)$ form a nested system in K, and their diameters shrink to 0. Hence there is just one point, which we call $f(x)$, common to all. This provides an extension of f to A_0. Now partition X into a set X^* of equivalence classes under the relation $q(x, y) = 0$, denoting the class containing x by x^*, and so on. Define $q^*(x^*, y^*) = q(x, y)$; this makes X^* into a metric space and the continuous natural image of X. In this natural mapping, A_0 passes onto a closed subset A^* of X^*. The function $f^*(a^*) = f(a)$, a in A_0, is continuous (indeed isometric) on A^*. It can be extended to all of X^*, by 2.4. Going back and defining $f(x) = f^*(x^*)$, we get an extension of f with the desired properties.

We shall show in 6 that a "simultaneous extension" of the type of 2.6 cannot always be obtained if the hypothesis is merely that of 4.1 for each of the functions involved. However, using the procedure of 4.1 and the result of 2.6, the reader may prove the following:

4.2. THEOREM. *Let A be a closed subset of a fully normal space X. Let F be a linear (convex) set of functions each defined on A and with values in a complete metric linear (convex) subset K of a convex topological linear space L. Furthermore let there be defined on A a pseudo-metric q such that for each f in F and for each positive r there is a positive s such that $q(a, a') < s$ implies $m[f(a), f(a')] < r$, where m is the metric in K. Then a simultaneous extension (in the sense of 2.6) can be made for all the f in F.*

None of the preceeding results can properly be claimed to be a generalization of Tietze's extension theorem, since we always require more than normality of X. We do not know whether the following is true: if A is a closed subset of a normal space X, and f maps A continuously into a bounded closed convex subset K of a Banach space L, then f can be continuously extended to X with values in K. Of course, in the finite dimensional case of L, this result is an easy consequence of the original theorem. In this case, we can replace "bounded" by "compact", and in this form the theorem does admit generalization.

4.3. THEOREM. *Let A be a closed subset of a normal space X. Let K be a compact convex subset of a normed linear space L. Let f be a continuous function on A with values in K. Then f can be continuously extended to X with values in K (see note added in proof).*

Proof: Since K is separable, we can find a countable family $v_1, v_2, \cdots$ of bounded linear functionals on L such that if u, u' belong to K and $u(v_n) = u'(v_n)$ for all n, then $u = u'$. (cf. [2, p.484, "Note"]). We now imbed K in the space (s) of [4]. For u in K, define $U(u)_n = v_n(u)$. This mapping is continuous and one-to-one, and hence a homeomorphism. We may therefore forget about the original L and regard K as a compact convex subset of (s). By 2.4, since (s) is metrizable, we can obtain a retraction of (s) on K, that is, a continuous r such that $r(u) \in K$

for u in (s), and $r(u) = u$ for u in K. Let $f_n(a) = f(a)_n$, the n-th coordinate of $f(a)$ in (s). By Tietze's original theorem, this f_n may be extended continuously to all of X. Defining $f_0(x) = [f_1(x), f_2(x), \cdots]$ in (s) we obtain a mapping of X into (s). Setting $f^{-}(x) = r[f_0(x)]$ gives the desired extension.

5. Simultaneous extension of real-valued functions. This section merely shows that special cases of 2.6 and 4.2 in which the linear space is the real number system R (or any finite dimensional linear space) can be reduced to 2.4 or 4.1, respectively, without further inquiry into the method of extension. In other words the possibility of "simultaneous extension" of real-valued functions is a direct consequence of the possibility of a single extension of a function with values in a suitable infinite dimensional space. This sounds quite plausible, but it is perhaps surprising that we must consider conjugate spaces.

Consider first a closed subset A of a metric space X, and the spaces $C(A)$ and $C(X)$ of continuous real-valued functions on them. Let L be $C(A)^{-}$, the conjugate space, with the weak topology (see [14], for example). Let K be the set of ξ in L with norm not exceeding 1 and with $\xi \geq 0$ (that is, $\xi(f) \geq 0$ for $f \geq 0$). For a in A, define $F(a)$ in K by $F(a)(f) = f(a)$. This F is continuous since we are using the weak topology, and K is convex. By 2.4 this F can be extended to X. For f in $C(A)$, define f^e by $f^e(x) = F(x)(f)$. We leave to the reader the completion of the proof of the following:

5.1. THEOREM. *The operation $f \longrightarrow f^e$ is a linear, isometric, nonnegative transformation of $C(A)$ into $C(X)$, and f^e is an extension of f.*

In the next section we shall show that 5.1 cannot be generalized for nonmetric X even if X is compact. However, the following is true.

5.2. THEOREM. *Let A be a closed subset of a fully normal space X. Let S be a separable (in the norm topology) closed linear subspace of $C(A)$. Then there is a linear isometric nonnegative transformation $f \longrightarrow f^e$ of S into $C(X)$ such that f^e is an extension of f.*

The proof is just like that of 5.1, except that we appeal to 4.1. To do this we must observe that since S is separable, K in S^{-} with the weak topology is metrizable (as is well known), for example with the metric

$$m(\xi, \eta) = \sum 2^{-n} \, |(\xi - \eta)(f_n)|,$$

where the f_n are dense in the unit ball of S; and that K is compact (Alaoglu-Bourbaki [1]) and thus complete.

6. Applications to measure in topological spaces. Let X be a topological space with a measure, and let A be a subset such that every function of a fixed linear set F of real-valued functions on A can be extended to a summable function on X by a positive linear operation P. By defining $J(f) = \int P(f)(x)\, m(dx)$ for f in F, we obtain a functional which may sometimes be represented by an integral (cf. [2, 3] or any of the references given there). When this is true, one obtains a measure m' on A which is generally not the mere restriction [$m'(E) = m(E)$ for $E \in A$] of m to A.

Unfortunately we have not been able to apply this process to any situation to obtain measures in A of a class not more easily obtainable by other methods. This is because of the requirement of the existence of a pseudo-metric in 4.2 with the stated properties, or of the separability of S in 5.2. The interest of the present section lies mainly in the fact that it is shown that one cannot avoid limitations of this sort. For this purpose we present only one of a variety of theorems, and then show why it cannot be generalized.

6.1. THEOREM. *Let X be a fully normal Hausdorff space and let m be a finite Baire measure* [8] *such that $m(V) = 0$ for an open V only if V is void. Let A be a compact subset of X. Let S be a separable subset of $C(A)$. Then there exists a strongly regular measure m' in A such that all functions in S are measurable and if $f \in S$ and $f \geq 0$, $f \neq 0$ then*

$$\int_A f(a)\, m'(da) > 0.$$

Proof. Let Q be the normed linear algebra generated by S and 1. By [3, 4.4] we can obtain a measure as described such that

$$\int_A f(a)\, m'(da) = \int_X f^e(x)\, m(dx).$$

The point to observe is that if $f \geq 0$, $f \neq 0$, then the same thing is true for f^e, and thus the right integral is positive.

Why can we not ignore the separability of S in 6.1? Let A_0 by any uncountable discrete set. By adding a "point at infinity" we obtain a compact space A. This space A can be imbedded in a cartesian product X of unit intervals. The obvious product measure [8, p.158 (2)] has the properties needed for 6.1. Let $S = C(A)$, and, forgetting that S is not separable, apply 6.1. The resulting measure would make every one of the points A_0 have nonzero measure, and so A itself would not be measurable. This shows why the separability of S in 6.1 cannot be ignored; and it also shows that one cannot ignore the pseudo-metric q in 4.2 or the separability of S in 5.2.

Added in proof: We have recently found, and shall soon publish, a stronger form of 4.3, namely in which "compact" is replaced by "separable".

REFERENCES

1. L. Alaoglu, *Weak topologies of normed linear spaces*, Ann. Math. 41 (1940), 252-267.

2. R. Arens, *A topology for spaces of transformations*, Ann. Math. 47 (1946), 480-495.

3. R. Arens, *Representation of functionals by integrals*, Duke Math. 17(1950) 499-506.

4. S. Banch, *Théorie des Opérations Linéaires*, Warsaw, 1932.

5. J. Dieudonné, *Une généralization des espaces compacts*, J. Math. Pure Appl. 23 (1944), 65-76.

6. J. Dugundji, *A generalization of Tietze's theorem*, Pacific J. Math. 1 (1951), 353-367.

7. I. Gelfand and G. Silov, *Über verschiedene Methoden der Einführung der Topologie in die Menge der maximalen Ideale eines normierten Ringes*, Rec. Math., N.S. 9 (1941), 25-39.

8. P. R. Halmos, *Measure Theory*, New York, 1950.

9. F. Hausdorff, *Erweiterung einer stetigen Abbildung*, Fund. Math., 30 (1938), 40-47.

10. S. Kakutani, *Simultaneous extension of continuous functions considered as a positive linear operation*, Japan. Math. 17 (1940), 1-4.

11. C. Kuratowski, *Remarques sur les transformations continus des espaces metriques*, Fund. Math. 30 (1938), 48-49.

12. A. H. Stone, *Paracompactness and product spaces*, Bull. Amer. Math. Soc. 54 (1948), 969-977.

A SHORT PROOF OF PILLAI'S THEOREM ON NORMAL NUMBERS

JOHN E. MAXFIELD

1. Introduction. The object of this paper is to give a short proof of the Pillai theorem [2] on normal numbers using the Niven-Zuckerman result [1] as a tool.

DEFINITION 1. A number σ is *simply normal* to the base r if, in the expansion to the base r of the fractional part of σ, we have $\lim_{n \to \infty} n_c/n = 1/r$ for all c, where n_c is the number of occurrences of the digit c in the first n digits of σ.

DEFINITION 2. A number σ is *normal* to the base r if $\sigma, r\sigma, r^2\sigma, \cdots$ are each simply normal to all the bases $r, r^2, r^3, \cdots$.

THEOREM (Pillai). *A necessary and sufficient condition that a number σ be normal to the base r is that it be simply normal to the bases $r, r^2, r^3, \cdots$.*

2. Proof. The necessity of the condition follows from the definition of normality.

To prove sufficiency, assume that σ is simply normal to the bases $r, r^2, \cdots$. Let $A = (a_1 a_2 \cdots a_v)$ be any fixed sequence of digits (to base r), where $v = hr - s$, $h > 0$, $0 \leq s < r$; and consider the occurrence of A in σ. Count the number of occurrences of A in the collection of sequences of length hr. There are s digits free after v of the hr digits are fixed. Thus there are $(s + 1)r^s$ different occurrences of A in these sequences.

For any positive integer n, define $f_n(A)$ to be the frequency of the occurrences of A in σ except in places where A will straddle the middle of sequences of length $2h2^{n-1}r$ starting in places congruent to 1 (mod $2h2^{n-1}r$), or where A will straddle the middle of sequences of length $4h2^{n-1}r$ starting in places congruent to 1 (mod $4h2^{n-1}r$), or $\cdots$, or where A will straddle the middle of sequences of length $2^s h2^{n-1}r$ starting in places congruent to 1 (mod $2^s h2^{n-1}r$), and so on.

Certainly $\lim_{n \to \infty} f_n(A)$, if it exists, will be equal to $f(A)$, the frequency of A in σ.

We have

$$f_1(A) = \frac{(s + 1)\, r^s}{hr\, r^{hr}} = \frac{1}{r^v} - \frac{v - 1}{hr^{v+1}},$$

Received July 5, 1951.
Pacific J. Math. 2 (1952), 23-24

since there are hr digits of σ to base r in each digit of σ to base r^{hr}, and σ is simply normal to the base r^{hr}. The number of occurrences of A straddling the middle of blocks of length $2hr$ is $(v-1)r^{2hr+s}$. The frequency of these in σ, where the sequence of length $2hr$ starts in a place congruent to 1 (mod $2hr$), is

$$\frac{(v-1)r^{2hr+s}}{2hr\, r^{2hr}} = \frac{v-1}{2hr^{v+1}},$$

since there are $2hr$ digits of σ to base r to each digit of σ to base r^{2hr}.

Thus

$$f_2(A) = \frac{1}{r^v} - \frac{v-1}{hr^{v+1}} + \frac{v-1}{2hr^{v+1}}.$$

Similarly,

$$f_3(A) = f_2(A) + \frac{v-1}{4hr^{v+1}} = \frac{1}{r^v} - \frac{v-1}{hr^{v+1}} + \frac{v-1}{hr^{v+1}}\left[\frac{1}{2} + \frac{1}{4}\right]$$

and

$$f_n(A) = \frac{1}{r^v} - \frac{v-1}{hr^{v+1}} + \frac{v-1}{hr^{v+1}} \sum_{i=1}^{n-1} 1/2^i.$$

It follows that

$$\lim_{n\to\infty} f_n(A) = 1/r^v.$$

Accordingly, by the Niven-Zuckerman result [1], stating that a necessary and sufficient condition in order that a number σ be normal is that every fixed sequence of v digits occur in the expansion of σ with the frequency $1/r^v$, we see that σ is normal to the scale r.

References

1. Ivan Niven and H. S. Zuckerman, *On the definition of normal numbers*. Pacific J. Math. 1 (1951), 103-109.

2. S. S. Pillai, *On normal numbers*, Proceedings of the Indian Acad. Sci., Section A, 12 (1940), 179-184.

University of Oregon and
Naval Ordnance Test Station

QUASI-CONVEXITY AND THE LOWER SEMICONTINUITY OF MULTIPLE INTEGRALS

Charles B. Morrey, Jr.

1. Introduction. We are concerned in this paper with integrals of the form

$$I(z, D) = \int_D f\left[x, z^i(x), z^i_{x^\alpha}(x)\right] dx, \tag{1.1}$$

where

$$x = (x^1, \cdots, x^\nu), \quad z = (z^1, \cdots, z^N), \quad p = p^i_\alpha$$

$$(i = 1, \cdots, N; \ \alpha = 1, \cdots, \nu),$$

$f(x, z, p)$ is continuous in its arguments, and D is a bounded domain.

The object of the paper is to discuss necessary and sufficient conditions on the function f for the integral I to be lower semicontinuous with respect to various types of convergence of the vector functions z. Because of the success of the "direct methods" in the Calculus of Variations, many writers have shown that certain integrals are lower semicontinuous. However, the writer knows of no paper in which a *necessary* condition for lower semicontinuity was discussed, although such a condition is very easy to obtain (see Theorem 2.1).

In §2, a general condition called "quasi-convexity" (see Definition 2.2) on the behavior of f as a function of p is obtained which is both necessary and sufficient for the lower semicontinuity of I with respect to the type of convergence given in Definition 2.1. This condition is that any linear function furnish the absolute minimum to $I(z, D)$ among all Lipschitzian (see below) functions which coincide with it on D^*, D being any bounded domain and D^* its boundary; here, of course, we consider f to be a function of p only. Section 3 discusses cases involving more general types of convergence and gives an existence theorem. In §4, it is shown that if $f(p)$ is continuous and quasi-convex, then it satisfies a certain generalized Weierstrass condition which reduces to the ordinary one (for the case at hand) when f is of class C'; this is, in turn, seen to be equivalent to the Legendre-Hadamard condition (see (4.8)) (quasi-regularity in its general form) when f is of class C . In §5, a general sufficient

presented to the American Mathematical Society at the Summer meeting of 1949 in Boulder, Colorado, under the title "Quasi-convexity and the lower semicontinuity of double integrals".

Received February 19, 1951.

Pacific J. Math. 2 (1952), 25-53

condition for quasi-convexity is proved and the necessary condition of §4 is seen to be sufficient when f is either a quadratic form in the p^i_α or is the integrand of a parametric problem with $N = \nu + 1$. The view of Terpstra's negative result [5] that even the strong Legendre-Hadamard condition (> 0) does not necessarily imply the existence of an alternating form $C^{\alpha\beta}_{ij}\, p^i_\alpha\, p^j_\beta$ ($C^{\beta\alpha}_{ij} = -\, C^{\alpha\beta}_{ij}$, and so on) such that

$$(1.2) \qquad f(p) + C^{\alpha\beta}_{ij}\; p^i_\alpha\; p^j_\beta \;\equiv\; \left(a^{\alpha\beta}_{ij} + C^{\alpha\beta}_{ij}\right) p^i_\alpha\; p^j_\beta$$

is positive definite when $\nu > 2$, it would seem that there is still a wide gap in the general case between the necessary and sufficient conditions for quasi-convexity which the writer has obtained. In fact, after a great deal of experimentation, the writer is inclined to think that there is no condition of the type discussed, which involves f and only a finite number of its derivatives, and which is both necessary and sufficient for quasi-convexity in the general case.

In (1.2), we have used the usual tensor summation convention, and will continue to use it throughout the paper; unless otherwise specified, the Greek letters will run from 1 to ν and the Latin letters from 1 to N.

We shall denote the sum and difference of vectors of the various sorts (x, z, p, and so on) in the usual way. We shall define

$$|x| = (x^\alpha x^\alpha)^{1/2}, \quad |z| = \left(z^i z^i\right)^{1/2}, \quad |p| = \left(p^i_\alpha\; p^i_\alpha\right)^{1/2}.$$

If $\zeta(x)$ is a vector function with derivatives, $\pi(x)$ will denote the vector function $\pi^i_\alpha(x) = \zeta^i_{x^\alpha}(x)$; similar notations involving other letters will be introduced as the occasion demands.

All integrals are Lebesgue integrals, frequently of vector functions. It is sometimes desirable to consider the behavior of a function $z(x)$ with respect to a particular variable x^α or to the $\nu - 1$ variables $(x^1, \cdots, x^{\alpha-1}, x^{\alpha+1}, \cdots, x^\nu)$. In such a case, we write x'_α for $(x^1, \cdots, x^{\alpha-1}, x^{\alpha+1}, \cdots, x^\nu)$, (x'_α, x^α) for x and so on. It is also convenient to write the boundary integrals

$$\int_{D^*} A_\alpha\;(x)\; dx'_\alpha,$$

where each $A_\alpha(x)$ may be a vector $A^i_\alpha(x)$ and the boundary D^* of the domain is sufficiently regular; such an integral is to be regarded as a Lebesgue-Stieltjes integral with respect to the set function $x'_\alpha(e)$ on D^* chosen so that Green's theorem

$$\int_{D^*} \zeta^i\; dx'_\alpha = \int_D \zeta^i_{x^\alpha}\; dx$$

holds. The closure of a set E will be denoted by $\overline{E}$.

Ordinary functions *of class* $\mathfrak{P}_s$, $\mathfrak{P}_s'$, $\mathfrak{P}_s''$, and so on, $s \geqq 1$, have been discussed at length in the papers [1] and [2]; the extension to vector functions is trivial. We define the integrals $\overline{D}_s(z, G)$ and $D_s(z, G)$ by

$$\overline{D}_s(z, G) = \int_G |z(x)|^s \, dx + D_s(z, G), \quad D_s(z, G) = \int_G \left[z^i_{x\alpha}(x) \, z^i_{x\alpha}(x) \right]^{s/2} dx.$$

Each function z of class $\mathfrak{P}_s$ is equivalent to a function $\overline{z}$ defined uniquely almost everywhere as that number such that the Lebesgue derivative of the set function

$$\int_e |z(x) - \overline{z}(x_0)|^s \, dx$$

is zero at x_0; $\overline{z}$ is supposed to be defined at *every* point x_0 where such a number exists; $\overline{z}$ is of class $\mathfrak{P}_s'$ (see [1] and [2]) and is also of class $\mathfrak{P}_s'$ in any coordinate system related to the original by a regular Lipschitzian transformation (cf. [2], Theorem 6.3; the $\overline{z}$ there used has a slightly different definition from the present one but the present theorem has been proved for vectors z with values in a Riemannian manifold in [4], Lemma 2.3 and Theorem 2.5).

A function z is said to *satisfy a* (uniform) *Lipschitz condition with coefficient* M on a set S if and only if

$$|z(x_1) - z(x_2)| \leqq M|x_1 - x_2|, \quad x_1 \in S, \quad x_2 \in S.$$

A function is *Lipschitzian* if it satisfies a Lipschitz condition.

If $g(y)$, $y = (y^1, \cdots, y^n)$, is summable on a domain D, we define the h-average function g_h by

$$g_h(y) = (2h)^{-n} \int_{y-h}^{y+h} g(\eta) \, d\eta, \qquad h > 0;$$

if g is summable then g_h is continuous where defined; if g is continuous on D then g_h is of class C' and g_h tends uniformly to g on each bounded closed set interior to D; if g is of class $\mathfrak{P}_s$ on D then g_h tends strongly in $\mathfrak{P}_s$ to g on each domain G with $\overline{G} \subset D$ (see [1], Lemma 5.1).

A form

$$C^{\alpha_1, \cdots, \alpha_\mu}_{i_1, \cdots, i_\mu} \, \pi^{i_1}_{\alpha_1} \cdots \pi^{i_\mu}_{\alpha_\mu} \quad \left(\mu \leqq \nu, \ 1 \leqq \alpha_\gamma \leqq \nu, \ 1 \leqq i_\gamma \leqq N, \ \gamma = 1, \cdots, \mu\right),$$

is called *alternating* if and only if the C's satisfy the obvious symmetry requirements and also the antisymmetry condition that

$$C^{\beta_1, \cdots, \beta_\mu}_{i_1, \cdots, i_\mu} = \pm C^{\alpha_1, \cdots, \alpha_\mu}_{i_1, \cdots, i_\mu}$$

according as $(\beta_1, \cdots, \beta_\mu)$ is an even or odd permutation of the indices $(\alpha_1, \cdots, \alpha_\mu)$; if $\zeta(x)$ is a vector function, then

$$\mu!\, C^{\alpha_1, \cdots, \alpha_\mu}_{i_1, \cdots, i_\mu} \pi^{i_1}_{\alpha_1}(x) \cdots \pi^{i_\mu}_{\alpha_\mu}(x) = C^{\alpha_1, \cdots, \alpha_\mu}_{i_1, \cdots, i_\mu} \frac{\partial(\zeta^{i_1}, \cdots, \zeta^{i_\mu})}{\partial(x^{\alpha_1}, \cdots, x^{\alpha_\mu})},$$

the fractions on the right denoting Jacobians.

2. A necessary and sufficient condition for lower-semicontinuity. We begin with some definitions.

DEFINITION 2.1. For the purposes of this section, we say that the vector functions z_n tend to the vector function z on the domain D if and only if the z_n and z all satisfy a uniform Lipschitz condition on D, independent of n, and the z_n tend uniformly to z on D. We shall write $z_n \longrightarrow z$ to denote this type of convergence.

DEFINITION 2.2. A function $f(p^i_\alpha)$ is said to be *quasi-convex* if and only if

$$\int_D f[p + \pi(x)]\, dx \geqq f(p) \cdot m(D), \qquad \pi^i_\alpha(x) = \zeta^i_{x^\alpha}(x),$$

for each constant p, each domain D, and each vector function ζ which satisfies a uniform Lipschitz condition on D and vanishes on D^*.

We shall show in this section that the integral $I(z,D)$ is lower semicontinuous with respect to the type of convergence specified in Definition 2.1 on each bounded domain D if and only if $f(x,z,p)$ is quasi-convex in p for each fixed (x,z).

THEOREM 2.1. *Suppose $I(z,D)$ is lower semicontinuous with respect to the type of convergence indicated on every region D. Then f is quasi-convex in p for each fixed (x,z).*

Proof. Let x_0 be any point, R be the cell $x_0 \leqq x^i \leqq x^i_0 + h$, Q be the cell $0 \leqq x^i \leqq 1$, and ζ be any function of class C' and periodic in each x^i with period 1. Let z_0 be any function of class C' on R.

For each n, define $\zeta_n(x)$ on R by

$$\zeta_n^i(x) = n^{-1} h\, \zeta^i\, [nh^{-1}(x - x_0)].$$

Then

$$\zeta_{nx^\alpha}^i(x) = \zeta_{x^\alpha}^i\, [nh^{-1}(x - x_0)]$$

and

$$I(z_0 + \zeta_n, R) = \int_R f\left\{x, z_0^i(x) + \zeta_n^i(x),\ p_{0\gamma}^i(x) + \pi_\gamma^i\, [nh^{-1}(x - x_0)]\right\} dx$$

$$= \sum_\alpha \int_{R_\alpha} \Big(f\left\{x, z_0^i(x) + \zeta_n^i(x),\ p_{0\gamma}^i(x) + \pi_\gamma^i\, [nh^{-1}(x - x_0)]\right\}$$

$$- f\left\{x_\alpha, z_0^i(x_\alpha),\ p_{0\gamma}^i(x_\alpha) + \pi_\gamma^i\, [nh^{-1}(x - x_0)]\right\}\Big) dx$$

$$+ \sum_\alpha (n^{-1}h)^\nu \int_Q f\left\{x_\alpha, z_0^i(x_\alpha),\ p_{0\gamma}^i(x_\alpha) + \pi_\gamma^i(\xi)\right\} d\xi,$$

where

$$\alpha = (\alpha_1, \cdots, \alpha_\nu),\ R_\alpha = R_{\alpha_1, \cdots, \alpha_\nu},\ n^{-1}(\alpha_\beta - 1) \leqq x^\beta \leqq n^{-1}\alpha_\beta,$$

$$x_\alpha = \left(x_{\alpha_1}^\beta, \cdots, {}_{\alpha_\nu}\right),\ x_{\alpha_1, \cdots, \alpha_\nu}^\beta = n^{-1}(\alpha_\beta - 1),\ \beta = 1, \cdots, \nu.$$

As $n \to \infty$, we see, since f is uniformly continuous on any bounded part of space, $\zeta_n(x)$ tends uniformly to zero, and the π_γ^i are bounded, that

$$\lim_{n \to \infty} I(z_0 + \zeta_n, R) = \int_R \left[\int_Q f[x, z_0(x), p_0(x) + \pi(\xi)]\, d\xi\right] dx.$$

From the lower semicontinuity of I, we must have

$$\int_R \left\{\int_Q f[x, z_0(x), p_0(x) + \pi(\xi)]\, d\xi\right\} dx \geqq \int_R f[x, z_0(x), p_0(x)]\, dx.$$

Now, let x_0, z_0, and p_0 be any constant vectors. By letting

$$z_0(x) = z_0 + p_{0\alpha}(x^\alpha - x_0^\alpha),$$

dividing by h^ν and letting $h \to 0$, we obtain

$$\int_Q f[x_0, z_0, p_0 + \pi(\xi)]\, d\xi \geqq f(x_0, z_0, p_0).$$

By approximations, we can extend this to all ζ which satisfy a uniform Lipschitz condition over the whole space and are periodic of period 1 in each x^α.

Now, let D be a bounded domain and suppose ζ satisfies a uniform Lipschitz condition on D and vanishes on D^*. Let R be a hypercube of edge h, with edges parallel to the axes which contains D. Extend ζ to the whole space by first defining it to be zero on $\overline{R}-D$ and then extending it to be periodic of period h in each variable. Then a simple change of function and variable reduces R to Q and establishes the result.

LEMMA 2.1. *Suppose R is a cell with edges $(2h^1), \cdots, (2h^\nu)$ and center x_0. Let h be the smallest h^α. Suppose also that $0 < k < h$, that ζ^* satisfies a uniform Lipschitz condition with coefficient $M \geqq 1$ on R^*, and suppose*

$$|\zeta^*(x)| \leqq k, \; x \in R^*.$$

Then there is a function ζ on $\overline{R}$ which satisfies a Lipschitz condition with coefficient M on $\overline{R}$, coincides with ζ^* on R^*, and is zero except on a set of measure at most

$$m(R) \cdot [1 - (1 - h^{-1}k)^\nu].$$

Proof. Let R_1 be the cell with center at x_0 and edges $2(h^\alpha - k)$, $\alpha = 1, \cdots, \nu$. Then, since $h = \min h^\alpha$, we have

$$m(R_1) \geqq m(R) \cdot (1 - h^{-1}k)^\nu.$$

Define $\zeta_1 = 0$ on $\overline{R_1}$ and equal to ζ^* on R^*. Then

$$|\zeta_1(x_1) - \zeta_1(x_2)| \leqq |x_1 - x_2| \qquad \text{if } x_1 \in \overline{R}_1, \quad x_2 \in R^*.$$

Thus ζ_1 satisfies a uniform Lipschitz condition with coefficient M on $\overline{R}_1 \cup R^*$. By a well known theorem, there exists an extension of ζ_1 to $\overline{R}$ (the whole space in fact) which satisfies the same Lipschitz condition.

LEMMA 2.2. *Suppose the vectors $\zeta_n \longrightarrow 0$* (in our sense) *on $\overline{R}$ and suppose f is quasi-convex in p. Then if p_0 is a constant vector we have*

$$m(R)\, f(p_0) \leqq \liminf_{n \to \infty} \int_R f[p_0 + \pi_n(x)]\, dx.$$

Proof. For all sufficiently large n, we have $k_n < h$, and $k_n \longrightarrow 0$, k_n being the maximum of $|\zeta_n(x)|$ for $x \in R^*$. For each n for which $k_n < h$, let η_n be the function of the preceding lemma which coincides on R^* with ζ_n, and let $\omega_n = \zeta_n - \eta_n$. Then if each ζ_n satisfies a uniform Lipschitz condition with coefficient $M \geqq 1$ on $\overline{R}$, then each η_n and ω_n satisfies one with coefficient M and $2M$, respectively. Moreover, each derivative $\eta^i_{nx\alpha}$ is uniformly bounded and $\eta^i_{nx\alpha} \longrightarrow 0$ almost everywhere. Since f is uniformly continuous on any bounded portion of p-space, we see that

$$\lim_{n\to\infty} \int_R \left| f\left(p^i_{0\alpha} + \eta^i_{nx\alpha} + \omega^i_{nx\alpha}\right) - f\left(p^i_{0\alpha} + \omega^i_{nx\alpha}\right)\right| dx = 0.$$

But the result then follows, since, for each n, we have

$$\int_R f\left(p^i_{0\alpha} + \omega^i_{nx\alpha}\right) dx \geqq m(R)\, f(p_0),$$

because of the quasi-convexity of the function f.

THEOREM 2.2. *Suppose f is continuous in (x, z, p) for all (x, z, p) and is quasi-convex in p for each (x, z). Suppose also that $z_n \longrightarrow z_0$ on the bounded domain D. Then*

$$I(z_0, D) \leqq \liminf_{n\to\infty} I(z_n, D).$$

Proof. Let ϵ be any positive number. For each positive integer k, let D_k consist of all the hypercubes of edge 2^{-k} whose faces lie along hyperplanes $x^\alpha = 2^{-k} i^\alpha$ (each i^α an integer) which lie in D. Since all the points $[x, z_0(x), p_0(x)]$ and $[x, z_n(x), p_n(x)]$ for $x \in D$ lie in a bounded portion of (x, z, p) space, we may choose k_1 so large that

$$\text{(2.1)} \qquad \int_{D-D_{k_1}} |f(x, z_n, p_n)|\, dx < \epsilon/5, \quad \int_{D-D_{k_1}} |f(x, z_0, p_0)|\, dx < \epsilon/5$$

for all n.

Let the hypercubes of D_{k_1} be $R_1, \cdots, R_N$. For each $k \geqq k_1$, let R_{ki}, $i = 1, \cdots, N \cdot 2^{\nu(k-k_1)}$, be all the hypercubes of side 2^{-k} described above which lie in D_{k_1}. For each such k, define $x^*_k(x)$, $z^*_k(x)$, $p^*_k(x)$ on D_{k_1} by

$$x^*_k(x) = [m(R_{ki})]^{-1} \int_{R_{ki}} x\, dx, \quad z^*_k(x) = [m(R_{ki})]^{-1} \int_{R_{ki}} z_0(x)\, dx,$$

$$\text{(2.2)} \qquad p^*_k(x) = [m(R_{ki})]^{-1} \int_{R_{ki}} p_0(x)\, dx$$

$$r_k(x) = \left\{ |x^*_k(x) - x|^2 + |z^*_k(x) - z_0(x)|^2 + |p^*_k(x) - p_0(x)|^2 \right\}^{1/2},$$

where $x \in R_{ki}$. Let $\zeta_n(x) = z_n(x) - z_0(x)$, $\pi_n(x) = p_n(x) - p_0(x)$. Then, on D_{k_1},

$$\begin{aligned} & f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_0(x)] \\ & = \left\{ f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_n(x)] \right\} \\ & + \left\{ f[x, z_0(x), p_0(x) + \pi_n(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)] \right\} \\ & - \left\{ f[x, z_0(x), p_0(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)] \right\} \\ & + \left\{ f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)] - f[x_k^*(x), p_k^*(x), p_k^*(x)] \right\}. \end{aligned} \tag{2.3}$$

Now, all the arguments of f occurring in (2.3) for $x \in D_{k_1}$ lie in a bounded closed cell in (x, z, p)-space over which f is uniformly continuous. Let

$$\epsilon(\rho) = \max |f(x', z', p') - f(x'', z'', p'')|, \qquad \rho \geqq 0$$

for all (x', z', p') and (x'', z'', p'') in this cell with

$$|x' - x''|^2 + |z' - z''|^2 + |p' - p''|^2 \leqq \rho^2.$$

then $\epsilon(\rho)$ is continuous for $\rho \geqq 0$ with $\epsilon(0) = 0$. Then, for each n and each $k \geqq k_1$, we have

$$|f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_n(x)]| \leqq \epsilon(|z_n(x) - z_0(x)|),$$

$$|f[x, z_0(x), p_0(x) + \pi_n(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)]| \leqq \epsilon[r_k(x)],$$

$$|f[x, z_0(x), p_0(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)]| \leqq \epsilon[r_k(x)].$$

Now, the $r_k(x)$ are uniformly bounded on D_{k_1} and tend to zero almost everywhere on D_{k_1}. Hence we may choose a $k \geqq k_1$ so large that

$$\begin{aligned} & \int_{D_{k_1}} |f[x, z_0(x), p_0(x) + \pi_n(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)]|\, dx < \epsilon/5, \\ & \int_{D_{k_1}} |f[x, z_0(x), p_0(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)]|\, dx < \epsilon/5, \end{aligned} \tag{2.4}$$

for all n. Since z_n converges uniformly to z_0, there is an n_1 such that

$$\int_{D_{k_1}} |f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_n(x)]|\, dx < \epsilon/5, \quad n > n_1. \tag{2.5}$$

Finally, since $x_k^*(x)$, and so on, are constant on each R_{ki}, and f is quasi-convex, we conclude from the previous lemma that

$$\liminf_{n\to\infty} \int_{D_{k_1}} \left\{ f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)] \right\} dx \geqq 0.$$

Using (2.3)-(2.5) and the above inequality, we see that

$$\liminf_{n\to\infty} I(z_n, D) \geqq I(z_0, D) - \epsilon.$$

Since ϵ is any positive number, the result follows.

3. Lower semicontinuity and weak convergence in $\mathfrak{P}_s$ $(s \geqq 1)$. In this section, we discuss additional conditions which with the quasi-convexity of f in p are sufficient to guarantee the lower semicontinuity of $I(z, D)$ with respect to weak convergence in $\mathfrak{P}_s$ on D.

DEFINITION 3.1. Suppose ζ is of class $\mathfrak{P}_s$ on the bounded domain D and suppose R is a cell with $\overline{R} \subset D$. Then ζ is said to be *strongly of class* $\mathfrak{P}_s$ *on* R^* if and only if $\overline{\zeta}$ is of class $\mathfrak{P}_s$ in x_α' on each face $x^\alpha =$ const. of R^* and thére is a sequence ζ_n of class C' on $\overline{R}$ such that

$$\overline{D}_s(\zeta_n - \zeta, R) \longrightarrow 0, \quad \overline{D}_s(\zeta_n - \overline{\zeta}, R^*) \longrightarrow 0.$$

LEMMA 3.1. *Suppose ζ is of class $\mathfrak{P}_s$ $(s \geqq 1)$ on the bounded domain D. For each α, $1 \leqq \alpha \leqq \nu$, let (a^α, b^α) be the open interval projection of D on the x^α axis. Then there exist sets Z^α of measure zero such that if $R: c^\alpha \leqq x^\alpha \leqq d^\alpha$ $(\alpha = 1, \cdots, \nu)$ is any closed cell in D with*

$$c^\alpha \in (a^\alpha, b^\alpha) - Z^\alpha, \quad d^\alpha \in (a^\alpha, b^\alpha) - Z^\alpha \qquad (\alpha = 1, \cdots, \nu),$$

then ζ is strongly of class $\mathfrak{P}_s$ on R^*.

Proof. Let R' be any rational cell in D (that is, $R = [C, D]$ with C^α, D^α rational). In [1], Lemma 5.1, we have seen that if ζ is of class $\mathfrak{P}_s$ on D, then

$$\lim_{h\to 0} \overline{D}_s(\zeta_h - \zeta, R) = 0. \tag{3.1}$$

For each α, define

$$\phi_h^\alpha(x^\alpha, R') = \int_{C_\alpha'}^{D_\alpha'} \left\{ |\zeta_h - \overline{\zeta}|^s + \left[\sum_{\substack{\beta=1\\ \beta\neq\alpha}}^{\nu} |\zeta_{hx^\beta} - \overline{\zeta}_{x^\beta}|^2 \right]^{s/2} \right\} dx_\alpha'.$$

Since $\overline{\zeta}$ is obviously of class $\mathfrak{P}_s'$ in x_α' for almost all x^α on $[C^\alpha, D^\alpha]$, $\phi_h^\alpha(x^\alpha, R')$ is defined for almost all x^α and

$$\lim_{h \to 0} \int_{C^\alpha}^{D^\alpha} |\phi_h^\alpha(x^\alpha, R')| \, dx^\alpha = 0.$$

By arranging the rational cells R' in some order and choosing successive subsequences, we may choose (on account of (3.1)) a final sequence $h_n \longrightarrow 0$ such that $\phi_{h_n}^\alpha(x^\alpha, R') \longrightarrow 0$ and $\overline{\zeta}$ is of class $\mathfrak{P}_s'$ in x_α' on $[C_\alpha', D_\alpha']$ for each x^α not in a set $Z^\alpha(R')$ of measure zero $(\alpha = 1, \cdots, \nu)$. Now let

$$Z^\alpha = \bigcup Z^\alpha(R');$$

then

$$m(Z^\alpha) = 0 \qquad (\alpha = 1, \cdots, \nu).$$

Now suppose R is one of the cells described in the lemma. Then it lies in some rational cell R' and we may take $\zeta_n = \zeta_{h_n}$.

LEMMA 3.2. *Suppose R is a cell with edges $(2h^1), \cdots, (2h^\nu)$ and center x_0. Let*

$$h = \min_{1 \leq \alpha \leq \nu} h^\alpha, \quad K = h^{-1}(h^\alpha h^\alpha)^{1/2}.$$

Suppose also that $0 < k < h$, that ζ^* is of class $\mathfrak{P}_s$ on an open domain containing $\overline{R}$ in its interior, and that ζ^* is strongly of class $\mathfrak{P}_s$ on R^* with

$$\int_{R^*} |\zeta^*|^s \, dS \leq k^s, \quad D_s(\zeta^*, R^*) \leq M^s \qquad (s \geq 1).$$

Then there is a function ζ of class $\mathfrak{P}_s$ on R which coincides with ζ^* on R^*, is zero except on a set of measure

$$m(R) \cdot [1 - (1 - h^{-1}k)^\nu],$$

and satisfies

$$D_s(\zeta, R) \leq \tau_s h^{-1} k (1 + K^s M^s), \quad \tau_s = \begin{cases} 2^{s/2} & (s \leq 2), \\ 2^{s-1} & (s \geq 2). \end{cases}$$

Proof. For each $x \in \overline{R}$, $x \neq x_0$, let $x^*(x)$ be the intersection of the ray $\overrightarrow{x_0 x}$ with R^*, and for each $x \in \overline{R}$ define

$$r(x) = \begin{cases} 0 & (x = x_0), \\ |x^*(x) - x_0|^{-1} \cdot |x - x_0| & (x \neq x_0). \end{cases}$$

Let $\Pi_\alpha^{\pm}$ be the pyramid in $\overline{R}$ with vertex x_0 and base the face $F_\alpha^{\pm}$ where

$$x^\alpha = x_0^\alpha \pm h^\alpha.$$

On the pyramid Π_ν^+, introduce coordinates $\xi^1, \cdots, \xi^{\nu-1}, r$ by

$$x^\nu = x_0^\nu + rh^\nu,\ x^\gamma = x_0^\gamma + r\xi^\gamma \quad (0 \leqq r \leqq 1,\ \gamma = 1, \cdots, \nu - 1).$$

Then, if r and ξ^γ are considered as functions of x, we have

$$r(x) = r,\ x^*(x) = [\xi^1(x) + x_0^1, \cdots, \xi^{\nu-1}(x) + x_0^{\nu-1}, h^\nu + x_0^\nu].$$

Similar coordinate systems may be set up on each of the other $\Pi_\alpha^\pm$.

Define

$$\phi(r) = \begin{cases} 0 & (0 \leqq r \leqq 1 - kh^{-1}), \\ hk^{-1}(r - 1 + kh^{-1}) & (1 - kh^{-1} \leqq r \leqq 1). \end{cases}$$

Choose a sequence ζ_n^* satisfying the conditions of Definition 3.1; and for each n, define

$$\zeta_n(x) = \phi[r(x)] \cdot \zeta_n^*[x^*(x)].$$

Then each $\zeta_n(x)$ is of class D' on $\overline{R}$.

We now compute the derivatives of ζ_n on each pyramid $\Pi_\alpha^\pm$ taking Π_ν^+ as an example. Then

$$\zeta_{nx^\gamma} = r^{-1}\phi(r)\zeta^*_{n\xi^\gamma} \qquad (1 \leqq \gamma \leqq \nu - 1),$$

$$\zeta_{nx^\nu} = (h^\nu)^{-1}\phi'(r)\zeta_n^* - (h^\nu)^{-1} r\phi(r)\,\xi^\gamma\zeta^*_{n\xi^\gamma} \ (\gamma \text{ summed from 1 to } \nu - 1).$$

Then, since $r^{-1}\phi(r) \leqq 1$ and $\phi'(r) = k^{-1}h$ for $1 - hk^{-1} \leqq r \leqq 1$,

$$|\pi_n(x)|^2 \leqq \left(\zeta^{*i}_{n\xi^\gamma}\zeta^{*i}_{n\xi^\gamma}\right) + 2k^{-2}|\zeta_n^*|^2 + 2(h^\nu)^{-2}(\xi^\gamma\xi^\gamma)\left(\zeta^{*i}_{n\xi^\gamma}\zeta^{*i}_{n\xi^\gamma}\right)$$

$$\leqq 2\left[k^{-2}|\zeta_n^*|^2 + K^2\left(\zeta^{*i}_{n\xi^\gamma}\zeta^{*i}_{n\xi^\gamma}\right)\right] \qquad (n \text{ not summed}).$$

Using the inequality

$$(a^2 + b^2)^{s/2} \leqq \sigma_s(|a|^s + |b|^s), \quad \sigma_s \leqq \begin{cases} 1 & (s \leqq 2) \\ 2^{(s-2)/2} & (s \geqq 2), \end{cases}$$

we obtain

$$D_s(\zeta_n, \Pi_\nu^+) \leqq \tau_s \int_{1-kh^{-1}}^1 r^{\nu-1}\, dr \int_{F_\nu^+} \Big[k^{-s} |\zeta_n^*|^s + K^s \left(\zeta_{n\xi\gamma}^{*i}\, \zeta_{n\xi\gamma}^{*i}\right)^{s/2}\Big]\, ds$$

$$\leqq \tau_s\, h^{-1} k \left[k^{-s} \int_{F_\nu^+} |\zeta_n^*|^s\, ds + K^s D_s(\zeta_n^*, F_\nu^+)\right].$$

Also

$$\int_{\Pi_\nu^+} |\zeta_n|^s\, dx = \int_{1-kh^{-1}}^1 r^{\nu-1} \phi^s(r)\, dr \int_{F_\nu^+} |\zeta_n^*|^s\, dS$$

$$\leqq h^{-1} k \int_{F_\nu^+} |\zeta_n^*|^s\, dS.$$

Adding these results for all the $\Pi_\alpha^\pm$, we obtain the result for each n; and also $\overline{D}_s(\zeta_n, R)$ is uniformly bounded. Thus, we may extract a subsequence which tends weakly in $\mathfrak{P}_s$ to some function ζ of class $\mathfrak{P}_s$ on R. Since each $\zeta_n = \zeta_n^*$ on R^*, ζ_n^* tends strongly in L_s to $\bar{\zeta}^*$ on R^*, we see from [2], Theorem 8.5, that $\zeta = \bar{\zeta}^*$ on R^*. From the lower semicontinuity of D_s (see [2], Theorem 8.2), the result follows.

LEMMA 3.3. *Suppose f is quasi-convex and of class C' for all p, and suppose for all p that*

$$\sum_{i,\alpha} \left(f_{p_\alpha^i}\right)^2 \leqq K^2 \left(|p|^{s-1} + 1\right)^2 \qquad (s \geqq 1).$$

If p_0 is any constant vector, D is any bounded domain, and ζ is of class $\mathfrak{P}_s$ on D and vanishes on D^*, then $f[p_0 + \pi(x)]$ is summable over D and

$$\int_D f[p_0 + \pi(x)]\, dx \geqq m(D) \cdot f(p_0).$$

Proof. There exists a sequence of functions ζ_n, each of class C' on D and vanishing on and near D^*, such that $\overline{D}_s(\zeta_n - \zeta, D) \to 0$ (see [2], Definition 9.1). For each n and almost all x on D, we have

$$|f[p_0 + \pi_n(x)] - f[p_0 + \pi(x)]| =$$

$$\left| [\pi_{n\alpha}^i(x) - \pi_\alpha^i(x)] \int_0^1 f_{p_\alpha^i} [p_0 + (1-t)\pi(x) + t\pi_n(x)]\, dt \right|$$

$$\leqq |\pi_n(x) - \pi(x)| \cdot K \cdot \int_0^1 \left\{ |(1-t)\, p_0 + \pi(x) + tp_0 + \pi_n(x)|^{\;-} + 1 \right\} dt$$

(3.2)

$$\leqq K\, |\pi_n(x) - \pi(x)| \left\{ h_s\, |p_0 + \pi(x)|^{s-1} + h_s\, |p_0 + \pi_n(x)|^{s-1} + 1 \right\},$$

where

$$h_s = \begin{cases} s^{-1} & (1 \leqq s \leqq 2) \\ s^{-1}\, 2^{s-2} & (s \geqq 2). \end{cases}$$

Using the Hölder inequality, and so on, and the strong convergence in $\mathfrak{P}_s$, we see that

$$\lim_{n \to \infty} \int_D f[p_0 + \pi_n(x)]\, dx = \int_D f[p_0 + \pi(x)]\, dx.$$

Since f is quasi-convex, the result follows.

LEMMA 3.4. *Suppose that f satisfies the hypotheses of Lemma 3.3. Suppose also that each ζ_n is of class $\mathfrak{P}_s$ on a domain D and is strongly of class $\mathfrak{P}_s$ on R^*, $\overline{R} \subset D$, with*

$$\lim_{n \to \infty} \int_{R^*} |\zeta_n|^s\, dS = 0,\ D_s(\zeta_n, R^*) \leqq M^s,\ D_s(\zeta_n, R) \leqq M^s \quad (n = 1, 2 \cdots).$$

Then for each p_0, $f[p_0 + \pi_n(x)]$ is summable for all sufficiently large n, and

$$\liminf_{n \to \infty} \int_R f[p_0 + \pi_n(x)]\, dx \geqq m(R) \cdot f(p_0),\ \pi^i_{n\alpha}(x) = \zeta^i_{nx^\alpha}(x).$$

Proof. For each n, let

$$k_n = \left[\int_{R^*} |\zeta_n|^s\, dS \right]^{1/s},$$

and let K and h be the quantities of Lemma 3.2 for R. Since $k_n \to 0$, we have $k_n < h$ for all $n >$ some n_1. For each such n, let η_n be the function of Lemma 3.2 which coincides on R^* with ζ_n, and let

$$\chi_n = \zeta_n - \eta_n,\ \kappa^i_{n\alpha} = \eta^i_{nx^\alpha},\ \omega^i_{n\alpha} = \chi^i_{nx^\alpha}.$$

Then, since $\chi_n = 0$ on R^*, we have

$$\int_R f[\pi_0 + \omega_n(x)]\, dx \geqq m(R)\, f(p_0).$$

As in (3.2), we see that, for each n, and almost all x on D,

$$|f[p_0 + \omega_n(x) + \kappa_n(x)] - f[p_0 + \omega_n(x)]|$$

$$\leqq K \cdot |\kappa_n(x)| \cdot (h_s |p_0 + \omega_n(x) + \kappa_n(x)|^{s-1} + h_s |p_0 + \omega_n(x)|^{s-1} + 1)$$

$$\leqq K \cdot |\kappa_n(x)| \cdot [(1 + sh_s) h_s |p_0 + \pi_n(x)|^{s-1} + sh_s^2 |\kappa_n(x)|^{s-1} + 1].$$

Using the Hölder inequality, and so on, we see that

$$\lim_{n\to\infty} \int_R |f[p_0 + \pi_n(x)] - f[p_0 + \omega_n(x)]| \, dx = 0,$$

from which the result follows.

THEOREM 3.1. *Suppose f is of class C' in (x, z, p) and quasi-convex in p. Suppose also that there are numbers k and K, $K > 0$, such that*

(i) $f(x,z,p) \geqq k$, (iii) $f_{x^\alpha} f_{x^\alpha} \leqq K^2 (|p|^s + 1)^2$

(ii) $f_{p_\alpha^i} f_{p_\alpha^i} \leqq K^2 (|p|^{s-1} + 1)^2$, (iv) $f_{z^i} f_{z^i} \leqq K^2 (|p|^s + 1)^2$.

for all (x, z, p).

Suppose also that $z_n \longrightarrow z_0$ weakly in $\mathfrak{P}_s$ on the bounded domain D and that either

(a) *each z_n and z_0 are continuous on D and z_n converges uniformly to z_0 on each closed set interior to D, or*

(b) *the set functions $D_s(z_n, e)$ are uniformly absolutely continuous on each closed set interior to D.*

Then

$$I(z_0, D) \leqq \liminf_{n\to\infty} I(z_n, D).$$

REMARK. If $s = 1$, weak convergence in $\mathfrak{P}_s$ implies the hypothesis (b).

Proof. We note first that hypothesis (ii) implies

$$|f(x,z,p) - f(x,z,0)| = |p_\alpha^i \int_0^1 f_{p_\alpha^i}(x, z, tp_\alpha^i) \, dt| \tag{3.3}$$

$$\leqq |p| \cdot \int_0^1 [f_{p_\alpha^i} f_{p_\alpha^i}(x, z, tp_\alpha^i)]^{1/2} \, dt$$

$$\leq |p| \cdot \int_0^1 K(t^{s-1} p^{s-1} + 1)\, dt \leq K(s^{-1} |p|^s + |p|).$$

Also, hypotheses (iii) and (iv) similarly imply

$$|f(x, z, 0) - f(0, 0, 0)| \leq K(|x| + |z|). \tag{3.4}$$

Thus, for all (x, z, p), we have

$$|f(x, z, p)| \leq |f(0, 0, 0)| + K(|x| + |z| + s^{-1} |p|^s + |p|). \tag{3.5}$$

Therefore $I(z_0, D)$ and the $I(z_n, D)$ are uniformly bounded.

For each α $(1 \leq \alpha \leq \nu)$, let (a^α, b^α) be the open interval projection of D on the x^α axis and let Z_0^α and Z_n^α be the sets of Lemma 3.1 for z_0 and z_n. Also for each α, n, k, let $E_{n,k}^\alpha$ be the set of x^α in $(a^\alpha, b^\alpha) - Z_n^\alpha$, where

$$\overline{D}_s(\overline{z}_n, D_{x^\alpha}) \leq k,$$

D_{x^α} being the set of x'_α such that $(x'_\alpha, x^\alpha) \in D$. Suppose that $\overline{D}_s(z_n, D) \leq M$, some uniform bound existing because of the weak convergence. Let

$$Z_{n,k}^\alpha = (a^\alpha, b^\alpha) - E_{n,k}^\alpha .$$

Then

$$m(Z_{n,k}^\alpha) \leq Mk^{-1}, \quad m(E_{n,k}^\alpha) \geq (b^\alpha - a^\alpha) - Mk^{-1} .$$

For each α, let

$$E^\alpha = \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} E_{n,k}^\alpha, \quad \tilde{Z}_0^\alpha = (a^\alpha, b^\alpha) - E^\alpha \cup Z_0^\alpha \cup \bigcup_{n=1}^{\infty} Z_n^\alpha .$$

Then $m(\tilde{Z}_0^\alpha) = 0$. For each α, each natural number n, and each integer i, define $Z_{n,i}^\alpha$ as the set of all x^α such that $x^\alpha - i \cdot 2^{-n} \in \tilde{Z}_0^\alpha$, and define

$$Z^\alpha = \bigcup_{n,i} Z_{n,i}^\alpha .$$

Then $m(Z^\alpha) = 0$.

Now, choose a point x_0 such that x_0^α is not in Z^α $(\alpha = 1, \cdots, \nu)$. For each natural number k, let $\mathfrak{G}_k$ be the totality of hypercubes of side 2^{-k} bounded by hyperplanes of the form $x^\alpha = x_0^\alpha + i \cdot 2^{-k}$. None of the numbers $x_0^\alpha + i \cdot 2^{-k}$ is in Z^α and, moreover, $\overline{z}_0$ and each $\overline{z}_n$ is strongly of class $\mathfrak{P}_s$ on R^* with

$\overline{D}_s(\bar{z}_n, R)$ uniformly bounded for infinitely many values of n, R being any hypercube of any $\mathcal{G}_k$. Since the totality of these hypercubes is countable, we may choose a subsequence, still called z_n, such that $I(z_n, D)$ tends to the former lim inf, $\bar{z}_n \longrightarrow \bar{z}_0$ almost everywhere on D, and $\overline{D}_s(z_n, R^*)$ is uniformly bounded for R of any $\mathcal{G}_k$ in D. Since $z_n \longrightarrow z$ in $\mathfrak{P}_s$, we also have

$$\lim_{n \to \infty} \int_{R^*} |z_n - z_0|^s \, dS = 0$$

for each such R.

Now, we first consider the alternative (a). Let ϵ be any positive number. For each k, let D_k be the union of all the cells of $\mathcal{G}_k$ which are interior to D. Since f is bounded below and $I(z_0, D)$ is finite, we first choose k_1 so large that

$$\text{(3.6)} \qquad \begin{aligned} &I(z_n, D - D_{k_1}) > -\epsilon/5 \qquad (n = 1, 2, \cdots). \\ &I(z_0, D_{k_1}) > I(z_0, D) - \epsilon/5. \end{aligned}$$

For this k_1, let $R_1, \cdots, R_q$ be the cells of D_{k_1} and for each $k \geqq k_1$, let

$$R_{ki} \qquad \left(i = 1, \cdots, q \cdot 2^{\nu(k-k_1)}\right)$$

be the cells of $\mathcal{G}_k$ in D_{k_1} For each k, define $x_k^*(x)$, $z_k^*(x)$, and $p_k^*(x)$ on D_{k_1} by (2.9). Then, from (ii), (iii), and (iv), it follows that

$$\text{(3.7)} \qquad \begin{aligned} &|f[x, z_0(x), p_0(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)]| \\ &\leqq K(|p_0(x)|^s + 1) \cdot (|x - x_k^*(x)| + |z_0(x) - z_k^*(x)|) \\ &+ K(h_s |p_0(x)|^{s-1} + h_s |p_k^*(x)|^{s-1} + 1) \cdot |p_0(x) - p_k^*(x)|, \end{aligned}$$

where

$$h_s = \begin{cases} s^{-1} & (1 \leqq s \leqq 2), \\ s^{-1} \cdot 2^{s-2} & (s \geqq 2); \end{cases}$$

the method of proof is similar to that of (3.3). If we let

$$\zeta_n = z_n - z_0, \quad \pi_n = p_n - p_0,$$

we see similarly that

$$|f[x, z_0(x), p_0(x) + \pi_n(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)]|$$

$$\leqq K(|p_n(x)|^s + 1)\ (|x - x_k^*(x)| + |z_0(x) - z_k^*(x)|) \tag{3.8}$$

$$+ K(h_s\ |p_n(x)|^{s-1} + h_s\ |p_k^*(x) + \pi_n(x)|^{s-1} + 1) \cdot |p_0(x) - p_k^*(x)|;$$

$$|f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_n(x)]| \tag{3.9}$$

$$\leqq K(|p_n(x)|^s + 1) \cdot |z_n(x) - z_0(x)|.$$

Now, by the Hölder inequality on each R_{ki}, we see that

$$\int_{Dk_1} |p_k^*(x)|^s\ dx \leqq \int_{Dk_1} |p_0(x)|^s\ dx. \tag{3.10}$$

By applying the Minkowski inequality, we see that the integrals

$$\int_{Dk_1} |\pi_n(x)|^s\ dx, \quad \int_{Dk_1} |p_k^*(x) + \pi_n(x)|^s\ dx \tag{3.11}$$

are uniformly bounded. Finally,

$$\lim_{k \to \infty} \int_{Dk_1} |p_0(x) - p_k^*(x)|^s\ dx = 0. \tag{3.12}$$

Hence, using (3.7)-(3.12), we may choose a k so large that

$$\int_{Dk_1} |f[x, z_0(x), p_0(x)] - f[x_k^*(x), z_k^*(x), p_k^*(x)]|\ dx < \epsilon/5, \tag{3.13}$$

$$\int_{Dk_1} |f[x, z_0(x), p_n(x)] - f[x_k^*(x), z_k^*(x)\ p_k^*(x) + \pi_n(x)]|\ dx < \epsilon/5 \tag{3.14}$$

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

and then choose n_1 so large that

$$\int_{Dk_1} |f[x, z_n(x), p_n(x)] - f[x, z_0(x), p_n(x)]|\ dx < \epsilon/5, \qquad n > n_1. \tag{3.15}$$

Since $x_k^*(x)$, $z_k^*(x)$, $p_k^*(x)$ are constant on each R_{ki}, it follows from Lemma 3.4 that

$$\liminf_{n \to \infty} \int_{Dk_1} \overset{*}{f}[x_k^*(x), z_k^*(x), p_k^*(x) + \pi_n(x)]\ dx \tag{3.16}$$

$$\geqq \int_{D_{k_1}} f[x_k^*(x),\ z_k^*(x),\ p_k^*(x)]\ dx.$$

Using (3.6) and (3.13)-(3.16), we see that

$$\liminf_{n\to\infty} I(z_n, D) \geqq I(z_0, D) - \epsilon.$$

The result follows in this case.

We now consider the alternative (b). For each natural number q, we define

$$f_q(x,z,p) = [1 - a_q(x,z)]\ f(x,z,p) + k \cdot a_q(x,z),$$

$$a_q(x,z) = \begin{cases} 0 & (0 \leqq R \leqq q), \\ 3(R-q)^2 - 2(R-q)^3 & (q \leqq R \leqq q+1), \\ 1 & (R \geqq q+1), \quad R = (|x|^2 + |z|^2)^{1/2}. \end{cases}$$

Remembering (3.3)-(3.5), we see that each f_q satisfies hypotheses (i)-(iv) with the same k and some K_q. Moreover f_q is independent of (x,z) for $R \geqq q+1$, and also

$$f_q(x,z,p) \leqq f_{q+1}(x,z,p), \quad \lim_{q\to\infty} f_q(x,z,p) = f(x,z,p).$$

Thus it is sufficient to prove the lower semicontinuity for each q.

For a fixed q, we note that we may replace $|z_0(x) - z_k^*(x)|$ by $\phi_k(x)$ in (3.7) and (3.8) and $|z_n(x) - z_0(x)|$ by $\psi_n(x)$ in (3.9), where

$$\phi_k(x) = \min(|z_0(x) - z_k^*(x)|,\ 2q+2),$$

$$\psi_n(x) = \min(|z_n(x) - z_0(x)|,\ 2q+2).$$

From the uniform boundedness of the ϕ_k and ψ_n (q fixed), the uniform absolute continuity of the set function $D_s(z_n, e)$, and the facts that

$$\lim_{k\to\infty} \phi_k(x) = 0, \quad \lim_{n\to\infty} \psi_n(x) = 0$$

almost everywhere, it follows that the argument can be carried through as before for each fixed q.

THEOREM 3.2. *Suppose $s > \nu$ and suppose f satisfies the hypotheses of Theorem* 3.1 *with* (i) *replaced by*

$$\text{(i}'\text{)} \qquad f(x,z,p) \geqq m|p|^s + k \qquad (m > 0).$$

If z^* is any function of class $\mathfrak{P}_s$ on the bounded domain D, then there is a function z_0 of class $\mathfrak{P}_s$ which coincides with z^* on D^* and minimizes $I(z,D)$ among all such functions.

Proof. Let z_n be a minimizing sequence. It follows from (i$'$) that $D_s(z_n, D)$ is uniformly bounded. From [2], Theorem 9.4, it follows that $\overline{D}_s(z_n, D)$ is uniformly bounded. But then a subsequence, still called $\{z_n\}$, converges weakly in $\mathfrak{P}_s$ to some function z_0 of class $\mathfrak{P}_s$ which coincides with z^* on D^* by [2], Theorem 9.2. But, from [3], Chapter II, Theorem 2.1, it follows that the equivalent functions $\overline{z}_n$ and $\overline{z}_0$ are equicontinuous on closed sets interior to D. Hence z_n converges uniformly to $\overline{z}_0$ on each closed set interior to D. Hence, from the preceding theorem, z_0 is a desired solution.

More general theorems involving variable boundary values, similar to those in [3], Chapter III, §5, with $s > \nu$, can be proved.

4. Necessary conditions for quasi-convexity. In the two preceding sections, we have established the connection between quasi-convexity and lower semicontinuity. In this section, we shall establish some necessary conditions for quasi-convexity. In the next section, we establish some sufficient conditions which are also necessary when f has certain interesting special forms. Unfortunately, the writer is unable to establish conditions which are both necessary and sufficient in the general case.

LEMMA 4.1. *Suppose f is continuous, Q is the cell*

$$|x^\alpha| \leqq 1 \quad (\alpha = 1, \cdots, \nu), \quad \delta > 0,$$

and suppose

$$\int_Q f[p + \pi(x)]\, dx \geqq f(p) \cdot m(Q) \tag{4.1}$$

for every function ζ which satisfies a Lipschitz condition with coefficient $< \delta$ on $\overline{Q}$ and vanishes on Q^*. Then (4.1) *also holds with Q replaced by any bounded domain D.*

Proof. Suppose ζ satisfies the conditions on the bounded domain $\overline{D}$. Let $\overline{R}$ be a hypercube of side h which contains $\overline{D}$, and extend ζ to $\overline{R}$:

$$x_0^\alpha \leqq x^\alpha \leqq x_0^\alpha + h$$

by defining $\zeta = 0$ on $\overline{R} - \overline{D}$. Then ζ satisfies the conditions on $\overline{R}$, and

$$\zeta^*(x) = h^{-1}\,\zeta(x_0 + hx)$$

satisfies the conditions on $\overline{Q}$, and

$$\zeta^{*i}_{x^\alpha}(x) = \zeta^i_{x^\alpha}(x_0 + hx)\,.$$

DEFINITION 4.1. The function f is said to be *weakly quasi-convex* if with each p is associated a $\delta_p > 0$ such that (4.1) holds for all D and all ζ satisfying a Lipschitz condition with coefficient $< \delta_p$ and vanishing on D^*.

In other words, f is *weakly quasi-convex* if and only if each linear function furnishes a *weak relative minimum* among all Lipschitzian functions coinciding with it on the boundary, whereas f is *quasi-convex* if and only if any linear function furnishes the *absolute minimum* among all such functions. Thus we have the following result.

THEOREM 4.1. *If f is continuous and quasi-convex, it is weakly quasi-convex.*

We shall see that if f is weakly quasi-convex and continuous, then f satisfies a uniform Lipschitz condition on any bounded set in p-space and satisfies a *generalized Weierstrass condition* (see Theorem 4.3) which reduces to the *ordinary Weierstrass condition* if f is of class C' (see (4.7)) and is equivalent to the *Legendre-Hadamard condition* (see (4.8)) if f is of class C''.

LEMMA 4.2. *Suppose ϕ is continuous, and suppose corresponding to any point λ_0 in E_ν there is a $\delta > 0$ such that for any unit vector μ we have*

$$k\phi(\lambda_0 - h\mu) + h\phi(\lambda_0 + k\mu) \geqq (h + k)\,\phi(\lambda_0) \qquad (0 < h < \delta,\ 0 < k < \delta).$$

Then ϕ is convex in λ.

Proof. Let λ_0 be any point, and μ any point with $|\mu| = 1$. We shall show that

$$\psi(t) = \phi(\lambda_0 + \mu t)$$

is convex in t. From the hypothesis, it follows that for each t_0, there is a $\delta(t_0) > 0$ such that

$$\text{(4.2)}\quad k\psi(t_0 - h) + h\psi(t_0 + k) \geqq (h + k)\,\psi(t_0) \quad (0 \leqq h \leqq \delta,\quad 0 \leqq k \leqq \delta).$$

Now, suppose $t_1 < t_2$. Let

$$\chi(t) = \psi(t) - \psi(t_1) - \frac{t - t_1}{t_2 - t_1}\,[\psi(t_2) - \psi(t_1)].$$

Then $\chi(t)$ satisfies (4.2) and $\chi(t_1) = \chi(t_2) = 0$. Suppose $M = \max \chi(t)$ $(t_1 \leqq t \leqq t_2)$, and suppose $M > 0$. Let t_0 be the smallest value of t such that $\chi(t) = M$, and let the number $\delta(t_0)$ be chosen as above. Clearly $t_1 < t_0 < t_2$. Choose t_3 and t_4 with

$$|t_3 - t_0| < \delta, \ |t_4 - t_0| < \delta \quad (t_1 \leqq t_3 < t_0 < t_4 \leqq t_2).$$

Then $\chi(t_3) < M$, $\chi(t_4) \leqq M$, so that

$$(t_4 - t_0)\,\chi[t_0 - (t_0 - t_3)] + (t_0 - t_3)\,\chi[t_0 + (t_4 - t_0)] < (t_4 - t_3)\,\chi(t_0),$$

which contradicts the hypothesis. Thus $\chi(t) \leqq 0$, so that

$$\psi(t) \leqq \psi(t_1) + \frac{t - t_1}{t_2 - t_1}\,[\psi(t_2) - \psi(t_1)].$$

Since t_1 and t_2 were arbitrary with $t_1 < t_2$, the function ψ is convex in t. Thus ϕ is convex in λ.

THEOREM 4.2. *If f is weakly quasi-convex, then $f(p^i_\alpha + \lambda_\alpha \xi^i)$ is convex in λ for each fixed* p *and ξ.*

Proof. Let p^i_α, ξ^i and $\lambda_{0\alpha}$ be fixed and let μ_1 be any unit vector, and suppose $h > 0$, $k > 0$. Choose $\delta(p^i_\alpha, \xi^i, \lambda_{0\alpha}) > 0$ but so small that, for any bounded domain G,

$$\int_G f[p^i_\alpha + \lambda_{0\alpha}\xi^i + \zeta^i_{x^\alpha}(x)]\,dx \geqq m(G)\, f(p^i_\alpha + \lambda_{0\alpha}\xi^i) \tag{4.3}$$

for all ζ satisfying a Lipschitz condition of constant $< \delta$ on G and vanishing on G^*. Let $(\mu_1, \cdots, \mu_\nu)$ be a normal orthogonal set of unit vectors. If $\xi = 0$, the result is obvious. If $\xi \neq 0$, choose h and k with $0 < h\,|\xi| < \delta$, $0 < k\,|\xi| < \delta$, and let ρ be any number $> |\xi|/\delta$. Let $H = (1/\rho)\,k$, $K = (1/\rho)\,h$, and let R be the rectangular parallelepiped

$$-\rho H \leqq y^1 \leqq \rho K, \ |y^\beta| \leqq \rho \qquad (\beta = 2, \cdots, \nu)$$

where

$$y^\beta = x_\alpha \mu_\beta \ .$$

Let F_1^- be the face $y^1 = -\rho H$, F_1^+ be the face $y^1 = \rho K$, F_β^- be the face $y^\beta = -\rho$, F_β^+ be the face $y^\beta = \rho$, and let Π_β^- and Π_β^+ be the pyramids with vertex at the origin and base F_β^- and F_β^+, respectively. Let ζ be defined on $\overline{R}$ to be continuous on $\overline{R}$, zero on R^*, linear on each Π_β and Π_β^+, with $\zeta(0) = \xi$. Then

$$(4.4) \qquad \zeta^i_{x^\alpha} = \begin{cases} (\rho H)^{-1} \mu_{1\alpha} \xi^i = k\mu_{1\alpha} \xi^i & , \text{ on } \Pi_1^- \\ -(\rho K)^{-1} \mu_{1\alpha} \xi^i = -h\mu_{1\alpha} \xi^i, & \text{ on } \Pi_1^+ \\ \rho^{-1} \mu_{\beta\alpha} \xi^i & , \text{ on } \Pi_\beta^- \\ -\rho^{-1} \mu_{\beta\alpha} \xi^i & , \text{ on } \Pi_\beta^+ \end{cases}$$

Also

$$m(\Pi_1^-) = \nu^{-1} 2^{\nu-1} \rho^\nu H, \ m(\Pi_1^+) = \nu^{-1} 2^{\nu-1} \rho^\nu K, \ m(R) = 2^{\nu-1} \rho^\nu (H+K)$$

(4.5)

$$m(\Pi_\beta^-) = m(\Pi_\beta^+) = \nu^{-1} 2^{\nu-2} \rho^\nu (H+K) \qquad (\beta = 2, \cdots, \nu).$$

Then, by applying (4.3), (4.4), and (4.5), we obtain

$$\frac{1}{2\nu}\left\{ \frac{2k}{h+k} f[p^i_\alpha + (\lambda_{0\alpha} - h\mu_{1\alpha})\xi^i] + \frac{2h}{h+k} f[p^i_\alpha + (\lambda_{0\alpha} + k\mu_{1\alpha})\xi^i] \right.$$

$$\left. + \sum_{\beta=2}^{\nu} \left(f[p^i_\alpha + (\lambda_{0\alpha} - \rho^{-1}\mu_{\beta\alpha})\xi^i] + f\left[p^i_\alpha + (\lambda_{0\alpha} + \rho^{-1}\mu_{\beta\alpha})\right]\right)\right\}$$

$$\geqq f(p^i_\alpha + \lambda_{0\alpha}\xi^i).$$

Letting $\rho \longrightarrow \infty$, we obtain

$$kf[p^i_\alpha + (\lambda_{0\alpha} - h\mu_{1\alpha})\xi^i] + hf[p^i_\alpha + (\lambda_{0\alpha} + k\mu_{1\alpha})\xi^i] \geqq (h+k) f(p^i_\alpha + \lambda_{0\alpha}\xi^i)$$

From the preceding lemma, it follows that $f(p^i_\alpha + \lambda_\alpha \xi^i)$ is convex in λ for each ξ and p.

THEOREM 4.3. *Suppose f is continuous and convex in λ for p and ξ. Then f satisfies a uniform Lipschitz condition on each bounded closed set, and for each fixed p there exists a set of constants A_i^α such that*

$$f(p_\alpha^i + \lambda_\alpha \xi^i) \geqq f(p) + A_i^\alpha \lambda_\alpha \xi^i \tag{4.6}$$

for all λ and ξ. If f is of class C', (4.6) holds if and only if $A_i^\alpha = f_{p_\alpha^i}$, that is,

$$f(p_\alpha^i + \lambda_\alpha \xi^i) \geqq f(p) + f_{p_\alpha^i}(p) \lambda_\alpha \xi^i. \tag{4.7}$$

If f is of class C'', (4.7) holds for all p, λ, ξ if and only if

$$f_{p_\alpha^i p_\beta^j}(p) \lambda_\alpha \lambda_\beta \xi^i \xi^j \geqq 0 \tag{4.8}$$

for all λ, ξ, p.

Proof. Suppose, first, that f is of class C'. Let p and ξ be fixed. Then (4.7) follows from the convexity in λ. Moreover, since each unit vector e_α^i in the p-space is of the form $\lambda_\alpha \xi^i$, we see from the convexity in λ that

$$f(p) - f\left(p - e_\alpha^i\right) \leqq f_{p_\alpha^i}(p) \leqq f\left(p + e_\alpha^i\right) - f(p) \tag{4.9}$$

for all p. Thus the derivatives of f are uniformly bounded by these *differences in the values of f* on any bounded part of space. Moreover, in this case, if constants A_i^α satisfy (4.6), we must have

$$A_i^\alpha = f_{p_\alpha^i}(p).$$

Now, if f is of class C'', equation (4.8) with p replaced by $p_\alpha^i + \lambda_\alpha \xi^i$ is equivalent to the condition that f is convex in λ for each fixed p and ξ.

Finally, if f is continuous and has this stated convexity property, it is clear that the h-average function also does, and f_h is of class C'. By letting $h \to 0$, we see that f satisfies a uniform Lipschitz condition on any bounded closed set. Now, choose $h_n = n^{-1}$ and choose p fixed. From (4.9) and the uniform convergence of f_h to f on any bounded part of space, we conclude that the derivatives $f_{h_n p_\alpha^i}(p)$ are uniformly bounded. We may therefore choose a subsequence, still called h_n, such that

$$\lim_{n \to \infty} f_{h_n p_\alpha^i}(p) = A_i^\alpha.$$

Since (4.7) holds for all λ and ξ for each n, (4.6) holds in the limit.

5. **Sufficient conditions for quasi-convexity.** In this section we prove one general sufficient condition and then give conditions which are necessary and sufficient when f has certain interesting special forms.

LEMMA 5.1. *Suppose ζ satisfies a uniform Lipschitz condition on the closure $\overline{D}$ of the bounded domain D and suppose $\zeta = 0$ on D^*. If*

$$1 \leqq \mu \leqq \nu,\ 1 \leqq i_1, \cdots, i_\mu \leqq N,\ 1 \leqq \alpha_1 < \alpha_2 < \cdots < \alpha_\mu \leqq \nu,$$

then

$$\int_D \frac{\partial(\zeta^{i_1}, \cdots, \zeta^{i_\mu})}{\partial(x^{\alpha_1}, \cdots, x^{\alpha_\mu})}\, dx = 0.$$

Proof. Choose a large cell R containing $\overline{D}$ in its interior, and extend ζ by defining it to be zero outside $\overline{D}$. Then the second h-average function ζ_{hh} is of class C'' on $\overline{R}$ and vanishes on and near R^*. Since any integral of the above type formed for ζ_{hh} tends to that for ζ as $h \longrightarrow 0$, we need prove the theorem only for functions ζ of class C'' on cells $\overline{R}$.

As an example, take $i_\beta = \alpha_\beta = \beta$, $\beta = 1, \cdots, \mu$, $D = R$. Then

$$\begin{aligned}\int_R \frac{\partial(\zeta^1, \cdots, \zeta^\mu)}{\partial(x^1, \cdots, x^\mu)}\, dx &= \int_R \sum_{\alpha=1}^{\mu} (-1)^{\mu+\alpha} \zeta^\mu_{x^\alpha}\, Q\, dx \\ &= \int_{R^*} \zeta^\mu \sum_{\alpha=1}^{\mu} (-1)^{\mu+\alpha} Q\, dx'_\alpha \\ &\quad - \int_R (-1)^\mu \zeta^\mu \sum_{\alpha=1}^{\mu} (-1)^\alpha \frac{\partial}{\partial x^\alpha} Q\, dx,\end{aligned}$$

where

$$Q = \frac{\partial(\zeta^1, \cdots, \zeta^{\alpha-1}, \zeta^\alpha, \cdots, \zeta^{\mu-1})}{\partial(x^1, \cdots, x^{\alpha-1}, x^{\alpha+1}, \cdots, x^\mu)},$$

the last equality holding by Green's theorem. But the boundary integral vanishes since $\zeta = 0$ on R^*, and the integrand in the second integral vanishes on R (see [3], Chapter II, Lemma 1.1, for instance).

THEOREM 5.1. *A sufficient condition that f be quasi-convex is that for each p there exist alternating forms*

$$A_i^\alpha \pi_\alpha^i, \; A_{ij}^{\alpha\beta} \pi_\alpha^i \pi_\beta^j, \dots, A_{i_1,\dots,i_\nu}^{\alpha_1,\dots,\alpha_\nu} \pi_{\alpha_1}^{i_\nu} \cdots \pi_{\alpha_\nu}^{i_\nu}$$

such that for all π we have

$$f(p+\pi) \geqq f(p) + A_i^\alpha \pi_\alpha^i + \dots + A_{i_1,\dots,i_\nu}^{\alpha_1,\dots,\alpha_\nu} \pi_{\alpha_1}^{i_1} \cdots \pi_{\alpha_\nu}^{i_\nu}.$$

Proof. This is an immediate consequence of the preceding lemma.

THEOREM 5.2. *If the $a_{ij}^{\alpha\beta}$ are constants and*

$$f(p) = a_{ij}^{\alpha\beta} p_\alpha^i p_\beta^j, \tag{5.1}$$

a necessary and sufficient condition that f be quasi-convex is that

$$a_{ij}^{\alpha\beta} \lambda_\alpha \lambda_\beta \xi^i \xi^j \geqq 0 \tag{5.2}$$

for all λ and ξ.

Proof. If $\zeta = 0$ on D^*, we see from Lemma 5.1 that

$$\int_D f[p+\pi(x)]\,dx = f(p)\, m(D) + \int_D a_{ij}^{\alpha\beta} \pi_\alpha^i(x)\, \pi_\beta^j(x)\,dx.$$

But Van Hove [6] has shown that the condition (5.2) is necessary (this also follows from Theorem 4.3) and sufficient for the second integral to be $\geqq 0$ for all ζ of class D' on D which vanish on D^* (hence this is true also for all ζ of class $\mathfrak{P}_2$ on D and vanishing on D^*).

LEMMA 5.2. *Suppose*

$$\sum_{i,j=1}^{n} a_{ij} x^i y^j = 0$$

for all x and y for which

$$\sum_{i,j=1}^{n} b_{ij} x^i y^j = 0.$$

Then there is a constant K such that

$$a_{ij} = Kb_{ij} \quad (i,\ j = 1, \cdots, n).$$

Proof. We may introduce new variables ξ and η by

$$x = c\xi, \quad y = d\eta,$$

c and d being $n \times n$ nonsingular matrices. Let a and b be the matrices of the original forms and A and B those of the transformed forms. Then

$$A = c'ad, \quad B = c'bd \quad (c'_{ij} = c_{ji}).$$

We shall show that there is a scalar K such that $A = KB$. We may assume that

$$B_{ii} = 1 \quad (i = 1, \cdots, r); \quad B_{ij} = 0 \text{ otherwise}, \qquad r \leqq n,$$

unless $B = 0$ in which case $A = 0$ also and the theorem holds. By taking $\eta^s = 1$, $\eta^j = 0$ $(j \neq s,\ s = 1, \cdots, n)$ in turn we see that

$$A_{is} = 0 \ (i = 1, \cdots, n,\ s > r); \ A_{is} = 0 \quad (i \neq s,\ s = 1, \cdots, r,\ i = 1, \cdots, n).$$

Then, by choosing $1 \leqq s < t \leqq r$ and setting $\eta^s = \eta^t = 1$, $\eta^j = 0, j \neq s, j \neq t$, we have

$$(A_{is} + A_{it})\,\xi^i = 0 \qquad \text{for all } \xi \text{ with } \xi^s + \xi^t = 0.$$

Thus there exists a constant $K(s,t)$ such that

$$A_{ss} + A_{st} = K(s,t), \qquad A_{ts} + A_{tt} = K(s,t).$$

Hence

$$A_{11} = A_{22} = \cdots = A_{rr} = K,$$

so that $A = KB$.

THEOREM 5.3. *Suppose that $N = \nu + 1$ and*

$$f(p) = F(X_1, \cdots, X_{\nu+1}), \tag{5.3}$$

where F is positively homogeneous of the first degree in the X_i and

$$X_i = -\det M_i \quad (i = 1, \cdots, \nu), \quad X_{\nu+1} = \det M_{\nu+1},$$

$$M_{\nu+1} = \| p_\alpha^1, \cdots, p_\alpha^\nu \|, M_i = \| p_\alpha^1, \cdots, p_\alpha^{i-1}, p_\alpha^{\nu+1}, p_\alpha^{i+1}, \cdots, p_\alpha^\nu \|$$

$$(i = 1, \cdots, \nu).$$

Then f is quasi-convex in p if and only if F is convex in the X_i.

Proof. If F is convex in the X_i, it follows from Theorem 5.1 that f is quasi-convex in p.

Hence suppose f is given by (5.3) and is quasi-convex in p. If

$$\Delta X_k = X_k\left(p_\alpha^i + \lambda_\alpha \xi^i\right) - X_k\left(p_\alpha^i\right),$$

then

$$\Delta X_k = X_{kp_\alpha^i} \lambda_\alpha \xi^i. \tag{5.4}$$

Also, since

$$p_\beta^k X_k = 0 \quad (\beta = 1, \cdots, \nu),$$

we have

$$p_\beta^k X_{kp_\alpha^i} = -\delta_\beta^\alpha X_i. \tag{5.5}$$

Now, choose a set of X_i not all zero and choose any p such that

$$X_i(p) = X_i.$$

Since f is quasi-convex and hence weakly so, there are constants A_α^i such that

$$f\left(p_\alpha^i + \lambda_\alpha \xi^i\right) \geqq f(p) + A_i^\alpha \lambda_\alpha \xi^i.$$

Since f depends only on the X_i, we must have

$$A_i^\alpha \lambda_\alpha \xi^i \leqq 0 \text{ for all } \lambda,\ \xi \text{ with } X_{kp_\alpha^i} \lambda_\alpha \xi^i = 0 \quad (k = 1, \cdots, \nu + 1). \tag{5.6}$$

Obviously, then, the equality must hold in (5.6). Using (5.4) and (5.5), we see that

$$p_\beta^k \Delta X_k = -\lambda_\beta\left(X_i \xi^i\right) \qquad (\beta = 1, \cdots, \nu). \tag{5.7}$$

Hence, we must have

$$A_i^\alpha \lambda_\alpha \xi^i = 0 \tag{5.8}$$

for all λ, ξ for which

$$X_i \xi^i = 0 \text{ and } D_i^\alpha \lambda_\alpha \xi^i = 0, \quad D_i^\alpha = X_k X_{kp_\alpha^i}. \tag{5.9}$$

Now, since not all the X_i are zero, assume $X_k \neq 0$. Then

$$\sum_{i \neq k} \left(A_i^\alpha X_k - A_k^\alpha X_i\right) \lambda_\alpha \xi^i = 0 \tag{5.10}$$

for all λ, ξ for which

$$\sum_{i \neq k} \left(D_i^\alpha X_k - D_k^\alpha X_i\right) \lambda_\alpha \xi^i = 0. \tag{5.11}$$

From the preceding lemma, it follows that there is a constant K such that

$$A_i^\alpha X_k - A_k^\alpha X_i = K\left(D_i^\alpha X_k - D_k^\alpha X_i\right). \tag{5.12}$$

Hence

$$A_i^\alpha = KD_i^\alpha + L^\alpha X_i, \quad L^\alpha = X_k^{-1}\left(A_k^\alpha - KD_k^\alpha\right). \tag{5.13}$$

From (5.7) and (5.13) it follows that

$$A_i^\alpha \lambda_\alpha \xi^i = KD_i^\alpha \lambda_\alpha \xi^i + L^\alpha \lambda_\alpha X_i \xi^i = C^k \Delta X_k, \; C^k = \left(KX_k - L^\alpha p_\alpha^k\right) \tag{5.14}$$

Finally, if we are given any values of the ΔX_k, the quantities

$$h_i = p_i^k \Delta X_k \quad (i = 1, \cdots, \nu) \text{ and } h_{\nu+1} = X_k \Delta X_k$$

are determined and the ΔX_k are also uniquely determined by the h_i. Using (5.7), we may determine the λ_α in terms of the h_i $(i = 1, \cdots, \nu)$, and substitute them into

$$h_{\nu+1} = X_k \Delta X_k = D_i^\alpha \lambda_\alpha \xi^i,$$

and we merely have to choose the ξ^i to satisfy the equation

$$\left(D_i^\alpha h_\alpha + h_{\nu+1} X_i\right) \xi^i = 0 \text{ with } X_i \xi^i \neq 0;$$

this is always possible unless all the $D_i^\alpha h_\alpha = 0$. Thus, unless these linear relations in the ΔX_i hold, we have

$$F(X + \Delta X) = f\left(p_\alpha^i + \lambda_\alpha \xi^i\right) \geqq f(p) + A_i^\alpha \lambda_\alpha \xi^i = F(X) + C^k \Delta X_k. \tag{5.15}$$

The result follows in this case by continuity.

Finally, since F is homogeneous of the first degree, we see by taking

$$\Delta X = hX, \quad h > -1,$$

that

$$F[(1+h)X] = (1+h)F(X) \geqq F(X) + hC^k X_k,$$

or

$$h\left[F(X) - C^k X_k\right] \geqq 0, \quad h > -1.$$

Hence $F(X) = C^k X_k$. Then by setting $X = hX_0$, $X_0 \neq 0$, choosing the C^k for this X_0, and then letting $h \longrightarrow 0$, we see that (5.15) holds for some C^k even if $X = 0$.

References

1. J. W. Calkin, *Functions of several variables and absolute continuity, I,* Duke Math. J. 6 (1940), 170-186.

2. C. B. Morrey, Jr., *Functions of several variables and absolute continuity, II,* Duke Math. J. 6 (1940), 187-215.

3. ———, *Multiple integral problems in the calculus of variations and related topics,* Univ. California Publ. Math., new series, **1**, no. 1 (1943), 1-130.

4. ———, *The problem of Plateau on a Riemannian manifold,* Ann. of Math. **49** (1948), 807-851.

5. F. J. Terpstra, *Die Darstellung biquadratischer Formen als Summen von Quadraten mit Anwendung auf die Variationsrechnung,* Math. Ann. **116** (1938), 166-180.

6. L. Van Hove, *Sur l'extension de la condition de Legendre du Calcul des Variations aux integrals multiples a plusieurs fonctions inconnues,* Nederl. Akad. Wetensch. **50**, no. 1 (1947), 18-23.

University of California, Berkeley

SOME INEQUALITIES IN CERTAIN NONORIENTABLE RIEMANNIAN MANIFOLDS

P. M. PU

1. Introduction. As is well known, the projective plane and the Moebius strip are nonorientable manifolds of dimension two. After introducing a Riemannian metric on each of them, we obtain two 2-dimensional nonorientable Riemannian manifolds. For convenience of reference, let us denote them by $M^2_{p^2}$ and M^2_m, respectively. Each of these manifolds has an area A. Moreover, there exists a family of closed curves, which are not homotopic to zero, on each manifold; and hence the set of the lengths of all these closed curves in consideration has a positive greatest lower bound, a. The purpose of this paper is to investigate the relationship between these two geometrical constants, A and a. It is found that, in each case, there exists an inequality [1] connecting them, of the form

$$A \geq ka^2, \tag{1}$$

k being a constant depending only on the conformal character of the Riemannian manifold. To establish such inequalities and to determine the corresponding best possible constants are the two central problems in this investigation.

For the time being, the projective plane is used in the following realization: it is given as the unit sphere with identification of diametrically opposite points. We assume further that the metric on $M^2_{p^2}$ is given by

$$ds^2 = g(p)\ d\rho^2,$$

$d\rho^2$ being the line element of the unit sphere taken from the embedding Euclidean space; $g(p) \in C_\omega$, $g(p) > 0$ for any point p on the manifold. As for the Moebius strip, we assume that it is given by the strip

$$-\beta < y < \beta,$$

with identification given by the fundamental group

Received July 24, 1951. The author wishes to express his gratefulness to Professor Charles Loewner for his guidance and encouragement in the preparation of this paper which represents the essential contents of his doctor thesis at Syracuse University.

Pacific J. Math. 2 (1952), 55-71

$$x' = x + n(2\alpha), \qquad (n = 0, \pm 1, \pm 2, \cdots).$$
$$y' = (-1)^n y$$

We assume further that the metric on M_m^2 is given by

$$ds^2 = g(x,y)\ (dx^2 + dy^2),$$

where $g(x,y) \in C_\omega$ and $g(x,y) > 0$. We shall see later that these assumptions are admissible in our cases.

The main idea of the method for solving these problems is to reduce the general metric, $g(p)d\rho^2$, to a simple and special one, $\bar{g}d\rho^2$, for which the *equality* in (1) holds, by an averaging process over a certain continuous group space; this enables us to handle our problems more easily. Let A_g, a_g, $A_{\bar{g}}$, $a_{\bar{g}}$ be the geometrical constants defined in terms of the original metric and the simplified metric respectively. Fortunately, this averaging process provides us a means of comparison between A_g and $A_{\bar{g}}$ and between a_g and $a_{\bar{g}}$; namely, we have

$$(2) \qquad \begin{aligned} A_g &\geq A_{\bar{g}}, \\ a_g &\leq a_{\bar{g}}. \end{aligned}$$

A comparison of the equality yielded by the special metric mentioned above with the foregoing inequalities (2) gives us the desired result.

Take, for example, the manifold M_{p2}^2. Each rotation of a 2-sphere about its center in the ordinary space is actually a conformal mapping of M_{p2}^2 onto itself. All these rotations form a compact Lie group G. Averaging $[g(p)]^{1/2}$ over G by the Hurwitz integration,

$$\int_G [(g(p))^{1/2}]^\sigma\ \delta\sigma = h^{1/2}$$

where $\sigma \in G$, and where $\delta\sigma$ is the invariant volume element, we can easily show that h is a constant and that the simplified metric is an elliptic one; this produces the equality

$$(3) \qquad A_h = \frac{2}{\pi} a_h^2 .$$

A combination of (3) with the following inequalities corresponding to (2),

$$A_g \geq A_h ,$$

$$a_g \leq a_h ,$$

shows that we have, in general,

$$A \geq \frac{2}{\pi} a^2 .$$

The same method can be extended, with some restrictions, to the case of M^n_{pn}, that is, the Riemannian manifold whose underlying topological space is an n-dimensional projective space.

In the case of M^2_m, let the rectangle

$$R: \begin{cases} -\alpha \leq x < \alpha \\ -\beta < y < \beta \end{cases}$$

be its fundamental region, as will be explained in §3. There exists a one-parameter family of conformal mappings of M^2_m onto itself,

$$x' = x + c ,$$

$$y' = y ,$$

c being real mod (4α). Averaging $[g(x+c,y)]^{1/2}$ over the interval $[0, 4\alpha]$ by the formula

$$\frac{1}{4\alpha}\int_0^{4\alpha} [g(x+c,y)]^{1/2} \quad dc = [\bar{g}(y)]^{1/2}$$

we can see that $[\bar{g}(y)]^{1/2}$ is free of x and is an even function on account of the fact that the metric is invariant under the fundamental group Γ; that is,

$$g[x + n\,(2\alpha), (-1)^n y] = g(x,y).$$

A further consideration of the same problem with the metric $g(y)\,(dx^2 + dy^2)$, where $g(y)$ is positive and even, leads to a distinguished $g_0(y)$ such that $g_0(y)\,(dx^2 + dy^2)$ plays the same role as the elliptic metric in the case of M^2_{p2} or M^n_{pn}; that is, $g_0(y)\,(dx^2 + dy^2)$ leads to the equality in (1).

2. Riemannian manifold M^2_{p2}, whose underlying topological space is a projective plane P^2. To begin with, let us prove the following general lemma, which will often be used.

LEMMA 1. *Let M_i^2 $(i = 1, 2, \cdots, n)$ be a set of n 2-dimensional Riemannian manifolds smooth of order 1 such that each M_i^2, with the same underlying topological space T^2, has a metric of the form*

$$ds_i^2 = g_i(p)\,ds^2,$$

where $g_i(p) > 0$, $g_i(p) \in C_0$ for $p \in M_i^2$; and ds^2 is a Riemannian metric which can be defined locally by

$$ds^2 = \sum_{j,k=1}^{2} g_{jk}(u_1, u_2)\,du^j\,du^k, \qquad g_{jk}(u_1, u_2) \in C_0.$$

Let $\bar{g}_n(p)$ be defined by the formula

$$\bar{g}_n(p) = \left\{ \frac{[g_1(p)]^{1/2} + \cdots + [g_n(p)]^{1/2}}{n} \right\}^2.$$

If the sets of lengths $S_i = \int_C [g_i(p)]^{1/2}\,ds$ $(i = 1, 2, \cdots, n)$ of a family F of curves C on T^2 have the same nonnegative greatest lower bound,

$$a_{g_1} = a_{g_2} = \cdots = a_{g_n},$$

and the areas A_{g_i} of M_i^2 have the same value,

$$A_{g_1} = A_{g_2} = \cdots = A_{g_n},$$

then we have

$$\text{(i)} \qquad A_{\bar{g}_n} \le A_{g_1} = \cdots = A_{g_n}$$

and

$$\text{(ii)} \qquad a_{\bar{g}_n} \ge a_{g_1} = \cdots = a_{g_n}.$$

Proof. By the definition of area and that of $\bar{g}_n$, we have

$$A_{\bar{g}_n} = \iint \bar{g}_n(p)\,d\omega = \iint \frac{[(g_1)^{1/2} + \cdots + (g_n)^{1/2}]^2}{n^2}\,d\omega,$$

where $d\omega$ is the area element, which can be expressed locally by the formula

$$d\omega = \begin{vmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{vmatrix}^{1/2} du^1\, du^2 .$$

Making use of the inequality

$$(\alpha_1 + \cdots + \alpha_n)^2 \leq n(\alpha_1^2 + \cdots + \alpha_n^2),$$

we have

$$A_{\bar{g}_n} \leq \iint n \frac{(g_1 + \cdots + g_n)}{n^2}\, d\omega$$

$$\leq \frac{\iint g_1\, d\omega + \cdots + \iint g_n\, d\omega}{n}$$

$$\leq \frac{A_{g_1} + \cdots + A_{g_n}}{n} .$$

By hypothesis, it follows that

$$A_{\bar{g}_n} \leq A_{g_1} = \cdots = A_{g_n} ,$$

which is (i).

The proof of (ii) follows from the definitions of the concepts concerned,

$$\int_C \bar{g}_n^{\,1/2} ds = \int_C \frac{g_1^{\,1/2} + \cdots + g_n^{\,1/2}}{n}\, ds = \frac{\int_C g_1^{\,1/2}\, ds + \cdots + \int_C g_n^{\,1/2}\, ds}{n}$$

$$\geq \frac{a_{g_1} + \cdots + a_{g_n}}{n} = a_{g_1} = \cdots = a_{g_n} ,$$

the line integrals being extended along any curve C of the family F. Hence

$$a_{\bar{g}_n} = \underset{c \in F}{\text{g.l.b.}} \left(\int_C \bar{g}_n^{\,1/2}\, ds \right) \geq a_{g_1} = \cdots = a_{g_n} .$$

We shall now prove the following theorem, which characterizes the relationship between the two geometrical constants A and a in $M^2_{p^2}$.

THEOREM 1. *Let $M^2_{p^2}$ be the Riemannian manifold whose underlying topological space is a projective plane and whose metric is locally defined by*

$$ds^2 = \sum_{i,k=1}^{2} g_{ik}(u_1, u_2)\, du^i du^k, \qquad g_{ik}(u_1, u_2) \in C_0;$$

let A be its area, and a the greatest lower bound of the lengths of all the closed curves not homotopic to zero on $M^2_{p^2}$; then

$$A \geq \frac{2}{\pi} a^2 .$$

Moreover, $2/\pi$ is the best constant.

Proof. On account of the Weierstrass approximation theorem, it suffices to assume that the fundamental tensor $g_{ik}(u_1, u_2)$ is analytic. Then we can introduce, in the small, an isothermic coordinate system on $M^2_{p^2}$ so that the metric takes the form

$$g^*(v_1, v_2)\,(dv_1{}^2 + dv_2{}^2),$$

where $g^*(v_1, v_2) \in C$, $g^*(v_1, v_2) > 0$. We define the metric on the universal covering surface S^2 of the projective plane by a projection process. The Riemannian manifold $M^2{}_{S^2}$ thus obtained is actually a Riemannian surface. According to the uniformization theorem, we can map $M^2_{S^2}$ onto the unit 2-sphere manifold $M^2_{U^2}$, and can arrange it in such a way that two diametrically opposite points of U^2 correspond to the same point of $M^2_{p^2}$. The metric has then the form

$$ds^2 = g(p)\,d\rho^2. \qquad g(p) > 0,\ g(p) \in C_\omega \text{ for } p \in M^2_{U^2},$$

where $d\rho^2$ is the line element of the unit sphere U^2 taken from the embedding 3-dimensional Euclidean space.

We remark that the area A of $M^2_{p^2}$ is one half that of $M^2_{U^2}$.

Let us consider all the rotations σ of the unit sphere U^2 about its center. All these rotations form a compact Lie group G. Applying the process of averaging over a compact Lie group, in this case the Hurwitz integration [2; 3, p.188], we have

$$\int_G [(g(p))^{1/2}]^\sigma\, \delta\sigma = [h(p)]^{1/2}.$$

We shall show that $[h(p)]^{1/2}$ is invariant with respect to all the left translations

$$\tau: \sigma \to \sigma' = \tau\sigma, \quad \tau, \sigma, \sigma' \in G,$$

and hence is a constant. In fact, let τ be any element of G; then, by definition,

$$[(h(p))^{1/2}]^{\tau} = \int_G [(g(p))^{1/2}]^{\tau\sigma}\, \delta\sigma = \int_G [(g(p))^{1/2}]^{\tau\sigma}\, \delta\tau\sigma,$$

since $\delta\sigma$ is invariant under all left translations. Therefore,

$$[h(\tau p)]^{1/2} = \int_G [(g(p))^{1/2}]^{\lambda}\, d\lambda = [h(p)]^{1/2}, \qquad \lambda = \tau\sigma \in G.$$

As the group G is transitive, $h^{1/2}$ is a constant.

Using $hd\rho^2$ instead of $gd\rho^2$ as the metric on the unit sphere U^2, we obtain a manifold with the spherical geometry. Preserving the metric $hd\rho^2$, and identifying the diametrical points on U^2, we get a manifold ${}_hM^2{}_{p^2}$, with the elliptic geometry. The two geometrical constants A_h and a_h can actually be evaluated:

$$A_h = 2\pi h,$$

$$a_h = \pi h^{1/2}.$$

Hence

(4) $$A_h = \frac{2}{\pi} a_h{}^2.$$

It is clear that if $g(p)$ is subjected to a transformation σ of G, the resulting metric $g^{\sigma}(p)\, d\rho^2$ is such that

(5) $$a_{g^{\sigma}} = a_g$$

and

(6) $$A_{g^{\sigma}} = A_g.$$

By approximating integrals by suitable sums and using Lemma 1, we easily obtain

(7) $$a_h \geq a_g$$

and

$$(8) \qquad A_h \leq A_g .$$

Combining (4), (7), and (8), we have

$$A_g \geq A_h = \frac{2}{\pi} a_h^2 \geq \frac{2}{\pi} a_g^2 .$$

Dropping the unnecessary indices, we obtain the inequality

$$A \geq \frac{2}{\pi} a^2 .$$

That $2/\pi$ is the best constant is evident, since we already have shown that the equality sign actually is attained when the metric is elliptic.

A slight generalization of Theorem 1, referring to certain special Riemannian metrics on the n-dimensional projective space P_n, can be proved in a similar fashion, using Hölder's inequality

$$(a_1 b_1 + \cdots + a_m b_m) \leq (a_1^p + \cdots + a_m^p)^{1/p} (b_1^q + \cdots + b_m^q)^{1/q},$$

where $a_i, b_i \geq 0$ and $p, q > 1$ such that $1/p + 1/q = 1$. The generalized theorem reads as follows:

THEOREM 2. *Let $M^n_{p^n}$ be the Riemannian manifold whose underlying topological space is an n-dimensional projective space P^n, which we suppose represented by the unit n-sphere U^n of the $(n+1)$-dimensional Euclidean space with identification of diametrically opposite points* p *and* p_d, *and whose metric can be represented in the form*

$$ds^2 = g(p) d\rho^2 ,$$

where $g(p) > 0$, $g(p) \in C_0$, $g(p) = g(p_d)$ for $p \in M^n_{p^n}$, and $d\rho^2$ is the line-element of the n-sphere, taken from the embedding Euclidean space; let V be its volume, and a the greatest lower bound of the lengths of all the closed curves which are not homotopic to zero on $M^n_{p^n}$; then

$$V \geq \frac{\pi^{\frac{1-n}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} a^n = k_n a^n .$$

Further, the above k_n *is the best constant.*

The proof of Theorem 2 is left to the reader.

3. Riemannian manifold M_m^2, whose underlying topological space is a Moebius strip. Let M_m^2 be the Riemannian manifold whose underlying topological space is a Moebius strip and whose metric is locally defined by

$$ds^2 = \sum_{i,k=1}^{2} g_{ik}(u_1, u_2)\, du^i du^k, \qquad g_{ik} \in C_0 .$$

By Weierstrass' approximation theorem, it suffices to assume that $g_{ik} \in C_\omega$. After introduction of the isothermic coordinate system, the metric takes the form

$$ds^2 = g^*(u,v)(du^2 + dv^2),$$

where $g^*(u,v) \in C_\omega$ and $g^*(u,v) > 0$. We define the metric $d\bar{s}^2$ on the universal covering surface of the Moebius strip by a projection process: $ds^2 = d\bar{s}^2$; that is, the metric is invariant under the fundamental group of the Moebius strip. The covering manifold of M_m^2 thus obtained is actually a simply connected Riemann surface. According to the uniformization theorem, we can map it conformally onto a strip

$$S: \begin{cases} -\beta < y < \beta \\ -\infty < x < \infty \end{cases}$$

of the (x,y)-plane. The fundamental group Γ appears then in the form:

$$\sigma: \begin{cases} x' = x + n(2\alpha), \\ y' = (-1)^n y \end{cases} \qquad (n = 0, \pm 1, \pm 2, \cdots).$$

The given manifold M_m^2 is mapped isogonally onto the fundamental region

$$R: \begin{cases} -\alpha \le x < \alpha \\ -\beta < y < \beta \end{cases}$$

with a metric of the form

$$ds^2 = g(x,y)(dx^2 + dy^2),$$

where $g(x,y) > 0$ and $g(x,y) \in C_\omega$. Moreover,

$$g(x + n(2\alpha), (-1)^n y) = g(x,y).$$

We are now in a position to prove the following theorem, which connects the two geometrical constants A and a in M_m^2.

THEOREM 3. *Let M_m^2 be the Riemannian manifold whose underlying topological space is a Moebius strip and whose metric is locally defined by*

$$ds^2 = \sum_{i,k=1}^{2} g_{ik}(u_1, u_2)\, du^i\, du^k, \qquad g_{ik} \in C_0\,;$$

let A be its area, and a the greatest lower bound of the lengths of all the closed curves which are not homotopic to zero on M_m^2; then we have

$$A \geq \frac{2}{\pi} \cdot \frac{e^{\beta\pi/\alpha} - 1}{e^{\beta\pi/\alpha} + 1} \cdot a^2 = k_{\alpha\beta}\, a^2$$

where 2α and 2β are the Euclidean lengths of the sides of the fundamental region R of the Moebius strip. Moreover, the above constant $k_{\alpha\beta}$ is best for a given ratio β/α.

Proof. Let us consider the continuous group

$$H: \begin{cases} x' = x + c \\ y' = y\,, \end{cases}$$

c being real (mod 4α); H consists of conformal transformations of the Moebius strip onto itself. It is evident that every two points which are equivalent under Γ remain equivalent under Γ after being operated on by elements of H. Defining the mean value, $[\bar{g}(y)]^{1/2}$, of $[g(x+c,y)]^{1/2}$ by the formula

$$\frac{1}{4\alpha}\int_0^{4\alpha} [g(x+c,y)]^{1/2}\, dc = [\bar{g}(y)]^{1/2},$$

we can prove, by a method similar to that in the former cases, that

$$\begin{aligned} &\bar{A} \leq A\,, \\ (9)\qquad &\bar{a} \geq a\,, \end{aligned}$$

where $\bar{A}$ and $\bar{a}$ have the same meaning as A and a except that we use the metric $\bar{g}(y)(dx^2 + dy^2)$ instead of the general one. Moreover, from the invariance of $g(x, y)$ under the group Γ, it follows immediately that

$$\bar{g}(y) = \bar{g}(-y).$$

We now shall consider the same problem with the simple metric

$$ds^2 = g(y)(dx^2 + dy^2),$$

where $g(-y) = g(y) > 0$. Noting (9), we see that if the best inequality is found for such a $g(y)$, it is also found for all $g(x, y)$, and hence our problem is solved.

We are now going to determine a special positive, even, and for nonnegative y monotonically decreasing function $g(y)$ such that a family F^* of closed geodesic lines through the origin and not homotopic to zero on M_m^2 can be defined in terms of it.

Let us first establish a differential equation for such $g(y)$. Putting

$$ds = [g(y)(1 + x'^2)]^{1/2}\, dy \qquad \left(x' = \frac{dx}{dy}\right),$$

we know that the equation for the extremals is

$$\frac{d}{dy}\left\{ x'\left[\frac{g(y)}{1 + x'^2}\right]^{1/2}\right\} = 0.$$

Solving this equation, we have

$$x = \int_0^y \frac{c\,d\eta}{[g(\eta) - c^2]^{1/2}} + k.$$

Since the geodesics under consideration have to go through the origin, the constant k has to be zero, and hence the equation becomes

$$x = \int_0^y \frac{c\,d\eta}{[g(\eta) - c^2]^{1/2}}.$$

The condition that the geodesics of the family be closed and not homotopic to zero requires that

$$\left.\frac{dy}{dx}\right|_{(\alpha, \tau)} = 0$$

for $-\beta < \tau < \beta$; that is,

$$[g(\tau)]^{1/2} = c$$

for $-\beta < \tau < \beta$. Hence we have

$$\alpha = \int_0^{\tau} \left[\frac{g(\tau)}{g(\eta)-g(\tau)} \right]^{1/2} d\eta, \qquad 0 \leq \tau < \beta. \tag{10}$$

For simplicity, let us normalize $g(y)$ so that $g(0) = 1$. Since $g(y)$ is supposed to be monotonically decreasing for nonnegative y, we can put

$$1 - g(\eta) = t,$$

$$1 - g(\tau) = \omega,$$

and have

$$\omega - t = g(\eta) - g(\tau)$$

$$d\eta = - \frac{dt}{g'(\eta)}.$$

Then equation (10) takes the form

$$\frac{-\alpha}{[1-\omega]^{1/2}} = \int_0^{\omega} \frac{1}{g'(\eta)} \cdot \frac{dt}{[\omega-t]^{1/2}}. \tag{11}$$

This is an Abel integral equation. According to the formula (cf. [1, p.484]) for the solution of such an equation,

$$f(x) = \int_0^x \frac{y(t)\,dt}{(x-t)^{1/2}},$$

we have

$$y(t) = \frac{1}{\pi}\left[\frac{f(0)}{t^{1/2}} + \int_0^t \frac{f'(z)\,dz}{(t-z)^{1/2}} \right], \tag{12}$$

and, in our case,

$$\frac{1}{g'(\eta)} = - \frac{\alpha}{\pi}\left[\frac{1}{t^{1/2}} + \frac{1}{2}\int_0^t \frac{dz}{(1-z)^{3/2}(t-z)^{1/2}} \right]$$

$$= -\frac{\alpha}{\pi}\left[\frac{1}{t^{1/2}} + \frac{t^{1/2}}{1-t}\right] = -\frac{\alpha}{\pi}\cdot\frac{1}{(1-t)\,t^{1/2}} = -\frac{\alpha}{\pi}\cdot\frac{1}{g(\eta)[1-g(\eta)]^{1/2}}$$

Thus we have established the differential equation for $g(y)$,

$$\frac{dg}{dy} = -\frac{\pi}{\alpha}\, g(1-g)^{1/2}. \tag{13}$$

The general solution of (13) is found to be

$$\log\frac{1+(1-g)^{1/2}}{1-(1-g)^{1/2}} = \frac{\pi}{\alpha}\, y + k.$$

When $y = 0$, we have $[1-g(0)]^{1/2} = 0$ and hence $k = 0$. Therefore,

$$\log\frac{1+(1-g)^{1/2}}{1-(1-g)^{1/2}} = \frac{\pi}{\alpha}\, y.$$

An explicit expression for $g(y)$ is as follows:

$$g(y) = \frac{4e^{\pi y/\alpha}}{(1+e^{\pi y/\alpha})^2} \tag{14}$$

for $-\beta < y < \beta$.

From the explicit expression (14) for $g(y)$, every property of $g(y)$ we assumed at the beginning is verified. It is a positive, even, and monotonic decreasing function for $y \geq 0$. Moreover, $g(y) \to 0$ as $y \to \infty$ and $g'(0) = 0$. Such a $g(y)$, with those properties just mentioned and defining the family F^* of the closed geodesics through (0, 0) and not homotopic to zero, is distinguished. Let us denote it by $g_0(y)$; that is, $g_0(y)$ is defined by either (13) or (14).

We are now in a position to establish the inequality in question for a positive, even, analytic function $g(y)$. By the definition of a, we have

$$2\int_0^{\tau} [g(y)\cdot(1+x'^2)]^{1/2}\, dy = \varphi(\tau) \geq a$$

for a closed curve, not homotopic to zero, on M_m^2. By taking this curve as one of F^*, we obtain

$$(1 + x'^2)^{1/2} = \left[\frac{g_0(y)}{g_0(y) - g_0(\tau)} \right]^{1/2}.$$

Hence,

$$\int_0^\tau \left[\frac{g(y)\, g_0(y)}{g_0(y) - g_0(\tau)} \right]^{1/2} dy = \frac{\varphi(\tau)}{2} \geq \frac{a}{2}. \tag{15}$$

This equation can be put into a more suitable form by setting

$$1 - g_0(y) = t,$$

$$1 - g_0(\tau) = \omega.$$

Then we have

$$\omega - t = g_0(y) - g_0(\tau),$$

$$dy = - \frac{dt}{g_0'(y)},$$

$$\tau = g_0^{-1}(1 - \omega).$$

Equation (15) then takes the form

$$\int_0^\omega \left[\frac{g(y)\, g_0(y)}{\omega - t} \right]^{1/2} \frac{dt}{g_0'(y)} = \frac{\varphi}{2} [g_0^{-1}(1 - \omega)]. \tag{16}$$

In formula (12), we have

$$f(z) = - \frac{\varphi}{2} [g_0^{-1}(1 - z)]$$

$$f(0) = - \frac{\varphi(0)}{2}$$

$$f'(z) = \frac{1}{2} \cdot \frac{\varphi'(g_0^{-1}(1 - z))}{g_0'(g_0^{-1}(1 - z))}.$$

Solving (16), we obtain

$$\frac{[g(y)\, g_0(y)]^{1/2}}{g_0'(y)} = \frac{1}{\pi} \left[- \frac{\varphi(0)}{2\, t^{1/2}} + \frac{1}{2} \int_0^t \frac{\varphi'(g_0^{-1}(1 - z))}{g_0'(g_0^{-1}(1 - z))} \cdot \frac{dz}{(t - z)^{1/2}} \right].$$

Putting

$$g_0^{-1}(1 - z) = u,$$

$$g_0^{-1}(1 - t) = y,$$

we have

$$\frac{[g(y)\ g_0(y)]^{1/2}}{g_0'(y)} = -\ \frac{\varphi(0)}{2\pi[1 - g_0(y)]^{1/2}} - \frac{1}{2\pi}\int_0^y \frac{\varphi'(u)\,du}{[g_0(u) - g_0(y)]^{1/2}},$$

or

$$[g(y)\,g_0(y)]^{1/2} = -\ \frac{\varphi(0)\,g_0'(y)}{2\pi[1 - g_0(y)]^{1/2}} - \frac{1}{2\pi}\ g_0'(y)\int_0^y \frac{\varphi'(u)\,du}{[g_0(u) - g_0(y)]^{1/2}}.$$

Integrating from 0 to β, we get

$$\begin{aligned}
\int_0^\beta [g(y)\,g_0(y)]^{1/2}\,dy &= -\ \frac{\varphi(0)}{2\pi}\int_0^\beta \frac{g_0'(y)}{[1 - g_0(y)]^{1/2}}\,dy \\
&\quad - \frac{1}{2\pi}\int_0^\beta g_0'(y)\left\{\int_0^y \frac{\varphi'(u)\,du}{[g_0(u) - g_0(y)]^{1/2}}\right\}dy \\
&= \frac{\varphi(0)}{\pi}[1 - g_0(\beta)]^{1/2} \\
&\quad - \frac{1}{2\pi}\int_0^\beta \varphi'(u)\left\{\int_u^\beta \frac{g'(y)\,dy}{[g_0(u) - g_0(y)]^{1/2}}\right\}du \\
&= \frac{\varphi(0)}{\pi}[1 - g_0(\beta)]^{1/2} \\
&\quad + \frac{1}{\pi}\int_0^\beta \varphi'(u)[g_0(u) - g_0(\beta)]^{1/2}\,du \\
&= -\ \frac{1}{\pi}\int_0^\beta \varphi(u)\left\{\frac{d}{du}[g_0(u) - g_0(\beta)]^{1/2}\right\}du
\end{aligned}$$

$$= \frac{1}{\pi} \int_0^{\beta} \varphi(u) \left\{ - \frac{d}{du} [g_0(u) - g_0(\beta)]^{1/2} \right\} du.$$

We remark that $-d[g_0(u) - g_0(\beta)]^{1/2}/du \geq 0$. Squaring and applying Schwarz's inequality, we have

$$\left\{ \int_0^{\beta} [g(y)\ g_0(y)]^{1/2}\, dy \right\}^2$$

$$= \frac{1}{\pi^2} \left\{ \int_0^{\beta} \varphi(u) \left(- \frac{d}{du} [g_0(u) - g_0(\beta)]^{1/2} \right) du \right\}^2,$$

$$\int_0^{\beta} g(y)\, dy \cdot \int_0^{\beta} g_0(y)\, dy \geq \frac{a^2}{\pi^2} [1 - g_0(\beta)].$$

The equality sign holds when $g = g_0$ due to a converse part of the theorem on Schwarz's inequality. Then we have

$$\int_0^{\beta} g(y)\, dy \geq \frac{a^2}{\pi^2} \cdot \frac{1 - g_0(\beta)}{\int_0^{\beta} g_0(y)\, dy}. \tag{17}$$

From (14), we can easily compute

$$1 - g_0(\beta) = \left(\frac{e^{\pi\beta/\alpha} - 1}{e^{\pi\beta/\alpha} + 1} \right)^2 = \frac{\pi^2}{4} k_{\alpha\beta}^2 \tag{18}$$

$$\int_0^{\beta} g_0(y) dy = \frac{2\alpha}{\pi} \cdot \frac{e^{\pi\beta/\alpha} - 1}{e^{\pi\beta/\alpha} + 1} = k_{\alpha\beta}\, \alpha. \tag{19}$$

Combining (17), (18), and (19), we obtain

$$\int_0^{\beta} g(y)\, dy \geq k_{\alpha\beta} \frac{a^2}{4\alpha}. \tag{20}$$

We remark that the equality sign in (20) holds for $g = g_0$. In fact, in this case, it can easily be proved that $a = 2\alpha$ by using the Weierstrass theory of extremal fields. From (19), we have

$$\int_0^\beta g_0(y)\,dy = \alpha\, k_{\alpha\beta} = k_{\alpha\beta}\frac{a^2}{4\alpha}. \tag{21}$$

Combining the definition of area and (20), we obtain

$$A = 4\alpha \int_0^\beta g(y)\,dy \geq k_{\alpha\beta}\, a^2 .$$

By (21), we know that

$$A = k_{\alpha\beta}\, a^2$$

for $g = g_0$. This shows that our $k_{\alpha\beta}$ is the best constant.

3. **Added in proof.** In a course on Riemannian Geometry given at Syracuse University in 1949, Professor C. Loewner proved the inequality $A \geq 3^{1/2}\, a^2/2$ for the case of M_t^2, the Riemannian manifold whose underlying topological space torus. The present investigation originates from this idea and has a similar method of treatment.

References

1. Ph. Frank und R. v. Mises, *Die Differential- und Integralgleichungen der Mechanik und Physik*, Erster Teil, New York, 1943.

2. A. Hurwitz, *Über die Erzeugung der Invarianten durch Integration*, Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. Math.-Phys.-Chem. Abt. 1897, 71-90.

3. H. Weyl, *Classical Groups, Their Invariants and Representations*, Princeton University Press, 1939.

SYRACUSE UNIVERSITY AND UNIVERSITY OF CALIFORNIA, BERKELEY

ON THE BARYCENTRIC HOMOMORPHISM IN A SINGULAR COMPLEX

PAUL V. REICHELDERFER

INTRODUCTION

0.1. Radó has introduced and studied the following approach to singular homology theory (see [2; 3; 4] for details). With a general topological space X associate a complex $R = R(X)$ in the following manner. For integers $p \geq 0$, let $v_0, \cdots, v_p$ be a sequence of $p+1$ points in Hilbert space E_∞, which are not required to be distinct or linearly independent, and let $|v_0, \cdots, v_p|$ denote their convex hull. Suppose that T is a continuous mapping from $|v_0, \cdots, v_p|$ into X. Then the sequence $v_0, \cdots, v_p$ jointly with T determines a p-cell in R, which is denoted by $(v_0, \cdots, v_p, T)^R$. The free Abelian group C_p^R generated by the p-cells in R is termed the group of integral p-chains in R. For integers $p < 0$, C_p^R is defined to be the group consisting of the zero element alone. The boundary operator $\partial_p^R: C_p^R \to C_{p-1}^R$ is defined, in the usual manner, as the trivial homomorphism if $p \leq 0$, and by the relation

$$\partial_p^R (v_0, \cdots, v_p, T)^R = \sum_{i=0}^{p} (-1)^p (v_0, \cdots, \hat{v}_i, \cdots, v_p, T)^R$$

if $p > 0$. Since $\partial_{p-1}^R \partial_p^R = 0$, one introduces the subgroup Z_p^R of p-cycles in C_p^R and the subgroup B_p^R of p-boundaries in C_p^R in the customary way, and defines the quotient group of Z_p^R with respect to B_p^R to be the homology group H_p^R.

0.2. The approach to singular homology theory pursued by Radó differs from other approaches in that absolutely no identifications are made. Thus two p-cells $(v_0', \cdots, v_p', T')^R$ and $(v_0'', \cdots, v_p'', T'')^R$ are equal only if they are identical; that is, if $v_i' = v_i''$ for $i = 0, \cdots, p$ and $T' \equiv T''$ on $|v_0', \cdots, v_p'| = |v_0'', \cdots, v_p''|$. In [3; 4], Radó introduces a technique for making identifications in a general Mayer complex and applies his procedure to study identifications in R, particularly those which yield homology groups isomorphic to the H_p^R. It is a primary purpose of the present paper to pursue the matter further in

Received January 24, 1951.

Pacific J. Math. 2 (1952), 73-97

order to establish stronger results than those obtained by Radó.

The identification scheme of Radó for the complex R is briefly described in §0.3 below; the reader should consult [3, §1] or [4, §5] for details.

0.3. Let $\{G_p\}$ be a collection of subgroups G_p of the group C_p^R of integral p-chains in R such that $\partial_p^R\, G_p \subset G_{p-1}$ for every integer p; such a system is termed an *identifier* for R. Let C_p^m be the quotient group of C_p^R with respect to G_p, and denote that element of C_p^m to which a chain c_p^R in C_p^R belongs by $\{c_p^R\}$. The restriction on the groups G_p clearly implies that the element $\{\partial_p^R\, c_p^R\}$ in C_{p-1}^m is independent of the choice of the representative c_p^R of the element $\{c_p^R\}$ in C_p^m; thus one may define homomorphisms ∂_p^m: $C_p^m \longrightarrow C_{p-1}^m$ by the formula $\partial_p^m \{c_p^R\} = \{\partial_p^R\, c_p^R\}$. The resulting system of groups C_p^m together with the operator ∂_p^m constitutes a Mayer complex m with homology groups H_p^m. Define a natural homomorphism π_p: $C_p^R \longrightarrow C_p^m$ by the formula $\pi_p\, c_p^R = \{c_p^R\}$. It is readily verified that π_p is a chain mapping; hence it induces homomorphisms π_{*p}: $H_p^R \longrightarrow H_p^m$. If for every integer p these homomorphisms are isomorphisms onto, then the identifier $\{G_p\}$ is termed unessential for R. Radó notes that a necessary and sufficient condition in order that an identifier G_p be unessential for R is that every cycle z_p^R in G_p should be the boundary of some chain c_{p+1}^R in G_{p+1}. (See [3, §§1.3, 1.4, 1.5] or [4, §5].)

0.4. One of the principal results in this paper may now be described. Let β_p^R: $C_p^R \longrightarrow C_p^R$ be the barycentric homomorphism in R (see [3, §3.1] or [4, §6]; also §1.3), and denote by $N(\beta_p^R)$ the nucleus of this homomorphism for every integer p.

THEOREM. *The system of nuclei $N(\beta_p^R)$ of the barycentric homomorphisms in in R constitutes an unessential identifier for R* (see §3.2).

This result is combined with those of Radó in [3] to obtain stronger theorems concerning identifiers than any previously obtained. Since further definitions are necessary before these results can be described, the reader is requested to consult §3 for their statements.

0.5. In the process of proving the theorem above, various results of independent interest have been attained. The reader is referred especially to §§1.6, 1.7, 1.10, 2.2 for theorems which show the structural description of the barycentric homomorphism and of the barycentric homotopy operator.

I. Further Relations in the Auxiliary Complex K

1.1. As in Radó [3;4], the auxiliary complex K is the "formal complex", in the sense of [1], for the set E_∞ of points in Hilbert space. For integers $p \geq 0$, p-cells in K are ordered sequences $(v_0, \cdots, v_p)$ of $p+1$ points in E_∞, which are not required to be distinct or linearly independent. These p-cells are taken as the base for a free Abelian group C_p, which is termed the group of finite integral p-chains in K. For $p < 0$, the group C_p is defined to be the group composed of the zero element alone. (See [3, §2.1] or [4, §6].)

1.2. In K the following known homomorphisms will be used. (See [3, §2.2] or [4, §6].)

(i) For integers j, p such that $0 \leq j \leq p$, $p > 0$, the homomorphism

$$j_p\colon\ C_p \longrightarrow C_{p-1}$$

is defined by the relation $j_p(v_0, \cdots, v_p) = (-1)^j (v_0, \cdots, \hat{v}_j, \cdots, v_p)$, where the symbol ^ is placed over the point v_j to indicate that v_j is to be deleted. For $j = p = 0$, j_p is defined to be the trivial homomorphism. A homomorphism differing from this one only by the absence of the factor $(-1)^j$ has been used by Radó in [2, §2.6]. The definition given above has been chosen because it permits simplifications in later definitions and formulas.

(ii) For integers $p > 0$, the *boundary operator*

$$\partial_p\colon\ C_p \longrightarrow C_{p-1}$$

is defined by the formula

$$\partial_p (v_0, \cdots, v_p) = \sum_{j=0}^{p} (-1)^j (v_0, \cdots, \hat{v}_j, \cdots, v_p).$$

For integers $p \leq 0$, ∂_p is defined to be the trivial homomorphism.

(iii) For integers $p \geq 0$ and an arbitrary point v in E_∞, the *cone homomorphism* $h_p^v\colon C_p \longrightarrow C_{p+1}$ is defined by the relation

$$h_p^v (v_0, \cdots, v_p) = (-1)^{p+1} (v_0, \cdots, v_p, v).$$

For integers $p < 0$, h_p^v is defined to be the trivial homomorphism.

(iv) For integers j, p such that $0 \leq j \leq p-1$, the *transposition homomorphism* $t_{pj}\colon C_p \longrightarrow C_p$ is defined by the relation

$$t_{pj}(v_0, \cdots, v_j, v_{j+1}, \cdots, v_p) = (v_0, \cdots, v_{j+1}, v_j, \cdots, v_p).$$

Observe that $t_{pj}(v_0, \cdots, v_p) = (v_0, \cdots, v_p)$ if and only if $v_j = v_{j+1}$.

(v) The *barycentric homomorphism* $\beta_p : C_p \longrightarrow C_p$ is defined as follows. For integers $p < 0$, β_p is the trivial homomorphism; for $p = 0$, $\beta_0 = 1$; and for $p > 0$, β_p is defined by the recursion formula

$$\beta_p(v_0, \cdots, v_p) = h^b_{p-1}\, \beta_{p-1}\, \partial_p(v_0, \cdots v_p),$$

where b is the barycenter of the points $v_0, \cdots, v_p$.

(vi) The barycentric homotopy operator ρ_p used by Radó [1; 3, §2.2 (iv); 4, §6] will not be used in this paper. In its stead, a modification ρ_{*p} is presently introduced, which has a simpler form, satisfies all the important identities which hold for the ρ_p, and has useful properties not possessed by ρ_p. The *modified barycentric homotopy operator*

$$\rho_{*p}\colon C_p \longrightarrow C_{p+1}$$

is defined as follows. For integers $p < 0$, ρ_{*p} is the trivial homomorphism; for $p = 0$, ρ_{*p} is defined by the relation

$$\rho_{*0}(v_0) = -h_0^{v_0}(v_0) = (v_0, v_0);$$

and for $p > 0$, ρ_{*p} is defined by the recursion formula

$$\rho_{*p}(v_0, \cdots, v_p) = -h^b_p[1 + \rho_{*p-1}\, \partial_p]\, (v_0, \cdots, v_p),$$

where b is the barycenter of the points $v_0, \cdots, v_p$.

1.3. Amongst the preceding homomorphisms the following identities hold (see [2, §2; 3, §2.3]):

$$\partial_p = \sum_{j=0}^{p} j_p \qquad (p \geq 0);$$

$$\partial_{p+1}\, h^v_p + h^v_{p-1}\, \partial_p = 1 \qquad (p > 0);$$

$$\partial_p\, \beta_p = \beta_{p-1}\, \partial_p \qquad (-\infty < p < +\infty);$$

$$\beta_p\, t_{pj} = -\beta_p \qquad (0 \leq j \leq p-1);$$

$$\partial_{p+1}\, \rho_{*p} + \rho_{*p-1}\, \partial_p = \beta_p - 1 \qquad (0 \leq p < +\infty).$$

Of these identities, only the last is new; it may be established by an inductive reasoning similar to that used to prove the corresponding identity for the conventional barycentric homotopy operator ρ_p.

1.4. For integers k, p such that $0 \leq k \leq p$, the homomorphism

$$k_{*p} : C_p \longrightarrow C_p$$

is defined by the relation

$$k_{*p}(v_0, \cdots, v_p) = (-1)^{p+k}(v_0, \cdots, \hat{v}_k, \cdots, v_p, v_k),$$

and the homomorphism

$$\gamma_p : C_p \longrightarrow C_p$$

is defined by the formula $\gamma_p = \sum_{k=0}^{p} k_{*p}$. Obviously one has the identities

$$k_{*p}(v_0, \cdots, v_p) = -k_{p+1}\, h_p^{v_k}\,(v_0, \cdots, v_p),\ p \geq 0,$$

$$k_{*p}(v_0, \cdots, v_p) = h_{p-1}^{v_k}\, k_p(v_0, \cdots, v_p),\ p > 0.$$

Now the reader will easily verify the relations

$$j_p\, k_{*p} = \begin{cases} (k-1)_{*p-1}\, j_p, & 0 \leq j < k \leq p, \\ k_{*p-1}(j+1)_p, & 0 \leq k \leq j < p, \\ k_p, & 0 \leq k \leq j = p; \end{cases}$$

$$k_{*p-1}\, j_p = \begin{cases} (j-1)_p\, k_{*p}, & 0 \leq k < j \leq p, \\ j_p\,(k+1)_{*p}, & 0 \leq j \leq k < p. \end{cases}$$

From these relations the following identity is readily established:

$$\gamma_{p-1}\, \partial_p = \partial_p\, (\gamma_p - 1).$$

Using the identity, the reader will easily prove the following result.

LEMMA. *If $P(x)$ be any polynomial having integral coefficients, then*

$$P(\gamma_{p-1})\, \partial_p = \partial_p\, P(\gamma_p - 1).$$

Explicitly, if $P(x) = \sum_{i=0}^{m} a_i x^i$, where the a_i are integers, then

$$\sum_{i=0}^{m} a_i \, \gamma_{p-1}^{i} \, \partial_p = \sum_{i=0}^{m} a_i \, \partial_p \left[\gamma_p^{i} - i \, \gamma_p^{i-1} + \cdots + (-1)^i\right].$$

where γ_p^i means that the homomorphism γ_p is to be repeated i times.

1.5. For integers k,p such that $0 \leq k \leq p$, the homomorphism

$$b_{pk}\colon C_p \longrightarrow C_{p+1}$$

is defined by the relation

$$b_{pk}(v_0, \cdots, v_p) = (-1)^k \left[v_0, \cdots, v_k, b(v_0, \cdots, v_k),\right.$$
$$\left. b(v_0, \cdots, v_k, v_{k+1}), \cdots, b(v_0, \cdots, v_k, \cdots, v_p)\right],$$

where $b(v_0, \cdots, v_q)$ is the barycenter of the points $v_0, \cdots, v_q$. Verification of the following simple relations is left to the reader:

$$-h_p^{b(v_0, \cdots, v_p)} (v_0, \cdots, v_p) = b_{pp}(v_0, \cdots, v_p);$$

$$-h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} (v_0, \cdots, v_{p-1}) = b_{pk} h_{p-1}^{v_p} (v_0, \cdots, v_{p-1})$$
$$(0 \leq k \leq p-1);$$

$$-h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} \, j_p(v_0, \cdots, v_p) = b_{pk} \, j_{*p}(v_0, \cdots, v_p)$$
$$(0 \leq k \leq p-1,\ 0 \leq j \leq p);$$

$$-h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} \, \partial_p(v_0, \cdots, v_p) = b_{pk} \, \gamma_p (v_0, \cdots, v_p)$$
$$(0 \leq k \leq p-1);$$

$$\begin{cases} -h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} \, \partial_p \, \gamma_p^{i-1} \, j_{*p}(v_0, \cdots, v_p) = b_{pk} \, \gamma_p^{i} \, j_{*p}(v_0, \cdots, v_p) \\ \qquad (0 \leq k \leq p-1,\ 0 \leq j \leq p,\ 1 \leq i) \\ -h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} \, \partial_p \, \gamma_p^{i} (v_0, \cdots, v_p) = b_{pk} \, \gamma_p^{i+1} (v_0, \cdots, v_p) \\ \qquad (0 \leq k \leq p-1,\ 0 \leq i). \end{cases}$$

If $P(x)$ be any polynomial having integral coefficients, then, for $0 \leq k \leq p-1$, we have

$$-h_p^{b(v_0, \cdots, v_p)} b_{p-1\,k} \, \partial_p \, P(\gamma_p) \, (v_0, \cdots, v_p) = b_{pk} \, \gamma_p \, P(\gamma_p) \, (v_0, \cdots, v_p).$$

1.6. For the homomorphisms β_p and ρ_{*p} the following structural descriptions are now obtained.

THEOREM. *The following relations hold:*

$$\rho_{*0} = b_{00},$$
$$\rho_{*p} = b_{pp} + \sum_{j=1}^{p} b_{pp-j}\, \gamma_p \cdots (\gamma_p - j + 1) \qquad (p > 0).$$

Proof. It is sufficient to verify these formulas for a given p-cell $(v_0, \cdots, v_p)$. For $p = 0$, the formula $\rho_{*0}(v_0) = b_{00}(v_0)$ is obvious from the definitions. So assume that

$$\rho_{*p-1} = b_{p-1\,p-1} + \sum_{j=1}^{p-1} b_{p-1\,p-1-j}\, \gamma_{p-1} \cdots (\gamma_{p-1} - j + 1) \quad (p \geq 1).$$

Using §1.2, §1.4, §1.5, and this assumption, and letting $b = b(v_0, \cdots, v_p)$, one obtains

$$\rho_{*p}(v_0, \cdots, v_0)$$
$$= -h_p^b(v_0, \cdots, v_p) - h_p^b\, \rho_{*p-1}\, \partial_p(v_0, \cdots, v_p)$$
$$= b_{pp}(v_0, \cdots, v_p) - h_p^b\, b_{p-1\,p-1}\, \partial_p(v_0, \cdots, v_p)$$
$$-\sum_{j=1}^{p-1} h_p^b\, b_{p-1\,p-1-j}\, \gamma_{p-1} \cdots (\gamma_{p-1} - j + 1)\, \partial_p(v_0, \cdots, v_p)$$
$$= b_{pp}(v_0, \cdots v_p) + b_{p\,p-1}\, \gamma_p(v_0, \cdots, v_p)$$
$$-\sum_{j=1}^{p-1} h_p^b\, b_{p-1\,p-1-j}\partial_p\, (\gamma_p - 1) \cdots (\gamma_p - j)\, (v_0, \cdots, v_p)$$
$$= b_{pp}(v_0, \cdots, v_p) + b_{pp-1}\, \gamma_p(v_0, \cdots, v_p)$$
$$+\sum_{j=2}^{p} b_{p\,p-j}\, \gamma_p(\gamma_p - 1) \cdots (\gamma_p - j + 1)\, (v_0, \cdots, v_p)$$
$$= b_{pp} + \sum_{j=1}^{p} b_{p\,p-j}\, \gamma_p \cdots (\gamma_p - j + 1)\, (v_0, \cdots, v_p).$$

So the proof is complete by induction.

1.7. THEOREM. *The following relations hold*:

$$\beta_0 = 0_1 \; b_{00},$$

$$\beta_p = 0_{p+1} \; b_{p0} \; \gamma_p (\gamma_p - 1) \cdots (\gamma_p - p + 1), \; p > 0.$$

The proof is similar to that for the theorm in the preceding section.

1.8. From these formulas for β_p and ρ_{*p} and the identities in §1.3, many further interesting relations may be obtained. For example, it is easy to establish the following results:

$$\beta_p = [\partial_{p+1} - (p+1)_{p+1}] \; \rho_{*p} \qquad (p \geq 0);$$

$$\partial_p = -p_p \; \rho_{*p-1} \; \partial_p \qquad (p \geq 0);$$

$$\beta_p = (p+1)_{p+1} \; (p+2)_{p+2} \; \rho_{*p+1} \; \rho_{*p} \qquad (p \geq 0).$$

These relations are not needed for the present purposes; they may be studied on a later occasion.

In order to clarify the structural descriptions for β_p and ρ_{*p} given in §§1.6, 1.7, it is convenient to introduce another homomorphism.

1.9. For integers $p \geq 0$, let $i_0, \cdots, i_p$ be any rearrangement of the sequence $0, \cdots, p$, and put $\epsilon_{i_0 \cdots i_p}$ equal to $+1$ or to -1 according as $i_0, \cdots, i_p$ is obtained from $0, \cdots, p$ by an even or by an odd number of transpositions. With each rearrangement one associates a homomorphism

$$\tau_p \colon C_p \longrightarrow C_p$$

defined by the formula

$$\tau_p (v_0, \cdots, v_p) = \epsilon_{i_0 \cdots i_p} (v_{i_0}, \cdots, v_{i_p}).$$

Sometimes, for clarity, the more explicit notation $\tau_p (i_0, \cdots, i_p)$ is used for this homomorphism. For integers j such that $0 \leq j \leq p$, denote by T_{pj} the class of all $\tau_p (i_0, \cdots, i_p)$ for which $i_0 < \cdots < i_j$ — that is, for which $i_0, \cdots, i_j$ are in natural order. Obviously T_{pp} consists of just one element, namely $\tau_p (0, \cdots, p) = 1$; and T_{p0} consists of the τ_p obtained by all possible rearrangements of $0, \cdots, p$. Moreover, $T_{p\,j-1} \supset T_{pj}$ for $1 \leq j \leq p$. Clearly the number of elements in the class T_{pj} is $(p+1) \, p \cdots (j+2)$ for $0 \leq j \leq p-1$. For each integer j in $0 \leq j \leq p$, define a homomorphism

$$P_{pj}\colon C_p \longrightarrow C_p$$

by the formula

$$P_{pj} = \sum \tau_p \qquad (\tau_p \in T_{pj}).$$

Observe that $P_{pp} = 1$. The reader will readily verify these identities:

$$k_{*p}\, P_{pj} = P_{pj},\ 0 \leq j < k \leq p;$$

$$\sum_{k=0}^{j} k_{*p}\, P_{pj} = P_{p\,j-1},\ 0 < j \leq p.$$

From these identities, the following result is established.

LEMMA. *The following relations hold*:

$$P_{pp} = 1,$$

$$P_{p\,p-j} = \gamma_p(\gamma_p - 1)\cdots(\gamma_p - j + 1),\ 1 \leq j \leq p.$$

Proof. That $P_{pp} = 1$ was noted above. From the second relation above it follows that

$$P_{p\,p-1} = \sum_{k=0}^{p} k_{*p}\, P_{pp} = \gamma_p\, P_{pp} = \gamma_p,$$

so the general formula is established for $j = 1$. Now suppose that

$$P_{p\,p-j+1} = \gamma_p(\gamma_p - 1)\cdots(\gamma_p - j + 2) \qquad (2 \leq j \leq p).$$

Using the preceding identities, one finds

$$\gamma_p\, P_{p\,p-j+1} = \sum_{k=0}^{p} k_{*p}\, P_{p\,p-j+1}$$

$$= \sum_{k=0}^{p-j+1} k_{*p}\, P_{p\,p-j+1} + \sum_{k=p-j+2}^{p} k_{*p}\, P_{p\,p-j+1}$$

$$= P_{p\,p-j} + (j-1)\, P_{p\,p-j+1};$$

$$P_{p\,p-j} = (\gamma_p - j + 1)\, P_{p\,p-j+1} = \gamma_p(\gamma_p - 1)\cdots(\gamma_p - j + 1).$$

Thus the lemma is established.

1.10. Combining the results of the preceding lemma with those in the theoems in §§1.6, 1.7, one obtains the following description for the homomorphisms β_p and ρ_{*p}.

THEOREM. *The following relations hold*:

$$\beta_p = 0_{p+1}\, b_{po}\, P_{po} = \sum_{\tau_p \in T_{p0}} 0_{p+1}\, b_{po}\, \tau_p \qquad (p \geq 0);$$

$$\rho_{*p} = \sum_{k=0}^{p} b_{pk}\, P_{pk} = \sum_{k=0}^{p} \sum_{\tau_p \in T_{pk}} b_{pk}\, \tau_p \qquad (p \geq 0).$$

1.11. Let $v_0, \cdots, v_p$ $(p \geq 0)$ be any sequence of $p+1$ points in E_∞. In §§1.2, 1.4, 1.5, 1.9, homomorphisms j_p, t_{pj}, k_{*p}, b_{pk}, τ_p, have been introduced which, when applied in any appropriate combination h_p to the special chain $(v_0, \cdots, v_p)$, yield a special chain either of the form $+(y_0, \cdots, y_q)$ or of the form $-(y_0, \cdots, y_q)$. In the sequel, $[h_p(v_0, \cdots, v_p)]$ is defined to be the p-cell $(y_0, \cdots, y_q)$, and $|h_p(v_0, \cdots, v_p)|$ denotes its convex hull $|y_0, \cdots, y_q|$. For example,

$$[0_{p+1}\, b_{po}\, \tau_p(i_0, \cdots, i_p)(v_0, \cdots, v_p)]$$
$$= (b(v_{i_0}),\ b(v_{i_0}, v_{i_1}), \cdots, b(v_{i_0}, v_{i_1}, \cdots, v_{i_p})).$$

If for two sequences of points $u_0, \cdots, u_p$ and $v_0, \cdots, v_p$ it is true that

$$(b(u_0),\ b(u_0, u_1), \cdots, b(u_0, u_1, \cdots, u_p))$$
$$= (b(v_0),\ b(v_0, v_1), \cdots, b(v_0, v_1, \cdots, v_p))$$

then clearly $u_j = v_j$ for $0 \leq j \leq p$. From the remarks in §1.9 and the preceding theorem, one thus obtains the following result.

LEMMA. *If the points* $v_0, \cdots, v_p$ $(p \geq 0)$ *are distinct, then the chain* $\beta_p(v_0, \cdots, v_p)$ *contains* $(p+1)!$ *terms; that is, for distinct elements* τ_p' *and* τ_p'' *in* T_{po}, *we have*

$$[0_{p+1}\, b_{po}\, \tau_p'(v_0, \cdots, v_p)] \neq [0_{p+1}\, b_{po}\, \tau_p''(v_0, \cdots, v_p)].$$

1.12. LEMMA. *Let* $v_0, \cdots, v_p$ $(p \geq 0)$ *be any set of* $p+1$ *points in* E_∞, *not necessarily distinct or linearly independent. A necessary and sufficient condition that a point v belong to the convex hull of the points*

(i) $b(v_0),\ b(v_0, v_1), \cdots, b(v_0, v_1, \cdots, v_p)$

is that it possess a representation of the form

$$\text{(ii)}\quad v = \sum_{j=0}^{p} \mu_j v_j \qquad \left(\sum_{j=0}^{p} \mu_j = 1,\ \mu_0 \geq \mu_1 \geq \cdots \geq \mu_p \geq 0 \right).$$

Proof. If v belongs to the convex hull of the points (i), then it has a representation of the form

$$\text{(iii)}\quad v = \sum_{i=0}^{p} \lambda_i\, b(v_0, \cdots, v_i) \qquad \left(\sum_{i=0}^{p} \lambda_i = 1,\ 0 \leq \lambda_i,\ 0 \leq i \leq p \right).$$

Thus

$$v = \sum_{i=0}^{p} \lambda_i \sum_{j=0}^{i} \frac{v_j}{i+1} = \sum_{j=0}^{p} \sum_{i=j}^{p} \frac{\lambda_i}{i+1} v_j,$$

which gives a representation of form (ii) for v. Conversely, if v has a representation of form (ii), put $\lambda_i = (i+1)\ (\mu_i - \mu_{i+1})$ for $0 \leq i \leq p-1$, $\lambda_p = (p+1)\ \mu_p$. It follows at once that v has a representation of form (iii), and hence belongs to the convex hull of the set of points (i).

1.13. For integers $p \geq 0$, if $u_0, \cdots, u_p$ is any sequence of $p+1$ points in E_∞, then $|u_0, \cdots, u_p|$ will denote its convex hull. Let k be any integer such that $0 \leq k \leq p$, and consider the sequence of $p+2$ points

(i) $u_0, \cdots u_k,\ b(u_0, \cdots u_k), \cdots, b(u_0, \cdots, u_k, \cdots, u_p)$,

that is (see §1.5), the sequence of points occurring in $b_{pk}(u_0, \cdots, u_p)$. Let

(ii) $w_0, \cdots, w_{p+1}$

be any rearrangement of the sequence of points (i). Designate by $x_0 = w_{h_0} = u_{i_0}$ the first u_i $(0 \leq i \leq k)$ occurring in the sequence (ii). In general, let $x_l = w_{h_l} = u_{i_l}$ $(0 \leq l \leq k)$ be the $(l+1)$st u_i $(0 \leq i \leq k)$ occurring in the sequence (ii), and put $x_l = u_l$ for $k+1 \leq l \leq p$ in case $k < p$. Now clearly $x_0, \cdots, x_p$ is a rearrangement of the sequence $u_0, \cdots, u_p$ in which the last $p-k$ elements are unaltered; the sequence (i) is a rearrangement of the sequence

(iii) $x_0, \cdots, x_k, b(x_0, \cdots, x_k), \cdots, b(x_0, \cdots, x_k, \cdots, x_p)$

in which the last $p + 1 - k$ elements are unaltered; and the sequence (ii) is a rearrangement of the sequence (iii) in which the points $x_0, \cdots, x_k$ appear in the same order as in (iii); that is, $x_l = w_{h_l}$ for $0 \leq l \leq k$, where $0 \leq h_0 < h_1 < \cdots < h_k \leq p$. Now let q be any integer such that $0 \leq q \leq p + 1$. It will be shown that

$$\text{(iv)} \quad b(w_0, \cdots, w_q) \in |b(x_0), b(x_0, x_1), \cdots, b(x_0, x_1, \cdots, x_p)| \qquad (0 \leq q \leq p+1).$$

Case $q = 0$. Then $b(w_0) = w_0$. If w_0 is one of the u_i $(0 \leq i \leq k)$, it follows by the choice above that $h_0 = 0$ and $w_0 = x_0 = b(x_0)$. If w_0 is not one of the u_i $(0 \leq i \leq k)$, there must be a $l \geq k$ such that $w_0 = b(u_0, \cdots, u_k, \cdots, u_l) = b(x_0, \cdots, x_k, \cdots, x_l)$. Thus relation (iv) is established when $q = 0$.

General case. By a rearrangement, the points $w_0, \cdots, w_q$ may be ordered into two sets

$$w_{h_0} = x_0, \cdots, w_{h_l} = x_l \qquad (0 \leq l \leq k,\ 0 \leq h_0 < \cdots < h_l \leq p),$$

$$\begin{cases} w_{h_{l+1}} = b(u_0, \cdots, u_k, \cdots, u_{i_{l+1}}) = b(x_0, \cdots, x_{i_{l+1}}) \\ w_{h_{l+2}} = b(u_0, \cdots, u_k, \cdots, u_{i_{l+2}}) = b(x_0, \cdots, x_{i_{l+2}}) \\ \qquad \cdots \\ w_{h_q} = b(u_0, \cdots, u_k, \cdots, u_{i_q}) = b(x_0, \cdots, x_{i_q}) \\ \qquad\qquad (k \leq i_{l+1} < i_{l+2} < \cdots < i_q \leq p). \end{cases}$$

The special cases which arise when one of these sets is missing are left to the reader. Now clearly

$$b(w_0, \cdots, w_q) = b(w_{h_0}, \cdots, w_{h_q})$$

$$= \sum_{j=0}^{l} \frac{1}{q+1}\left[1 + \sum_{h=l+1}^{q} \frac{1}{i_h+1}\right] x_j + \sum_{j=l+1}^{i_{l+1}} \frac{1}{q+1} \sum_{h=l+1}^{q} \frac{1}{i_h+1} x_j$$

$$+ \sum_{j=i_{l+1}+1}^{i_{l+2}} \frac{1}{q+1} \sum_{h=l+2}^{q} \frac{1}{i_h+1} x_j + \cdots + \sum_{j=i_{q-1}+1}^{i_q} \frac{1}{q+1}\frac{1}{i_q+1} x_j.$$

In view of this equation and of the lemma in §1.12, the relation (iv) now follows.

1.14. From the facts presented above, the following result is presently established.

LEMMA. *Let* $v_0, \cdots, v_p$ $(p \leq 0)$ *be any sequence of* $p+1$ *points in* E_∞. *Fix* $\tau_{p+1} \in T_{p+1\,0}\,(0 \leq k \leq p)$, $\tau_p \in T_{pk}$ (see §1.9). *Then there exists a* $\tau_p' \in T_{p0}$ *such that* (see §1.11).

$$|0_{p+2}\, b_{p+1\,0}\, \tau_{p+1}\, b_{pk}\, \tau_p(v_0, \cdots, v_p)| \subset |0_{p+1}\, b_{p0}\, \tau_p'(v_0, \cdots, v_p)|.$$

Proof. Evidently $[\tau_p(v_0, \cdots, v_p)] = (v_{i_0}, \cdots, v_{i_p})$, where $i_0, \cdots, i_p$ is a rearrangement of $0, \cdots, p$ such that $i_0 < \cdots < i_k$. Put $u_j = v_{i_j}$ for $0 \leq j \leq p$, so that $[\tau_p(v_0, \cdots, v_p)] = (u_0, \cdots, u_p)$. Then

$$[b_{pk}\, \tau_p(v_0, \cdots, v_p)]$$
$$= (u_0, \cdots, u_k, b(u_0, \cdots u_k), \cdots, b(u_0, \cdots, u_k, \cdots, u_p)),$$

and $[\tau_{p+1}\, b_{pk}\, \tau_p(v_0, \cdots, v_p)] = (w_0, \cdots, w_{p+1})$, where $w_0, \cdots, w_{p+1}$ is a rearrangement of

$$u_0, \cdots, u_k, b(u_0, \cdots, u_k), \cdots, b(u_0, \cdots, u_k, \cdots, u_p).$$

Finally,

$$[0_{p+2}\, b_{p+1\,0}\, \tau_{p+1}\, b_{pk}\, \tau_p(v_0, \cdots, v_p)]$$
$$= [b(w_0), b(w_0, w_1), \cdots, b(w_0, w_1, \cdots, w_{p+1})].$$

The reasoning of §1.13 shows that there is a rearrangement $x_0, \cdots, x_p$ of $u_0, \cdots, u_p$, and hence of $v_0, \cdots, v_p$, such that

$$|0_{p+2}\, b_{p+1\,0}\, \tau_{p+1}\, b_{pk}\, \tau_p(v_0, \cdots, v_p)|$$
$$\subset |b(x_0), b(x_0, x_1), \cdots, b(x_0, x_1, \cdots, x_p)|.$$

Let τ_p' be that element of T_{p0} such that $[\tau_p'(v_0, \cdots, v_p)] = (x_0, \cdots, x_p)$. Since

$$[0_{p+1}\, b_{p0}\, \tau_p'(v_0, \cdots, v_p)] = (b(x_0), b(x_0, x_1), \cdots, b(x_1, x_1, \cdots, x_p)),$$

the lemma is established.

1.15. If c_p is a p-chain in K, and A is a convex subset in E_∞, then the inclusion $c_p \subset A$ will mean that either $c_p = 0 \in C_p$ or else

$$c_p = \sum_{j=1}^{n} m_j (v_{0j}, \cdots, v_{pj}),$$

where the m_j are nonzero integers and $|v_{0j}, \cdots, v_{pj}| \subset A$ for $1 \leq j \leq n$. One readily verifies the following inclusions (see [3, §2.4]):

$$j_p (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (0 \leq j \leq p),$$

$$\partial_p (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (p \geq 0),$$

$$\beta_p (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (p \geq 0),$$

$$\rho_{*p} (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (p \geq 0),$$

$$t_{pj} (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (0 \leq j \leq p-1),$$

$$k_{*p} (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (0 \leq k \leq p),$$

$$\gamma_p (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (p \geq 0),$$

$$b_{pk} (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (0 \leq k \leq p),$$

$$\tau_p (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (\tau_p \in T_{po}),$$

$$P_{pj} (v_0, \cdots, v_p) \subset |v_0, \cdots, v_p| \qquad (0 \leq j \leq p).$$

II. Relations in the Complex R = R(X).

2.1. If A is a convex subset of E_∞, then for integers $p \geq 0$, C_p^A denotes that subgroup of C_p generated by those p-cells $(v_0, \cdots, v_p)$ for which $|v_0, \cdots, v_p| \subset A$; for $p < 0$, we have $C_p^A = 0 \in C_p$ (see §1.1). Suppose $T: A \longrightarrow X$ is a continuous mapping (see §0.1). For integers $p \geq 0$ define a homomorphism

$$T_p: C_p^A \longrightarrow C_p^R$$

by the relation $T_p (v_0, \cdots, v_p) = (v_0, \cdots, v_p, T)^R$ for $(v_0, \cdots, v_p) \in C_p^A$. For $p < 0$, let T_p be the trivial homomorphism. For chains c_p in C_p^A the notation $T_p c_p = (c_p, T)^R$ is used. In terms of this notation one finds the relation (see §0.1): $\partial_p^R (c_p, T)^R = (\partial_p c_p, T)^R$.

Now suppose that, for certain integers p,

$$h_p: C_p \longrightarrow C_q$$

is a homomorphism from the group C_p of p-chains into the group C_q of q-chains in K with the property that for all p-cells $(v_0, \cdots, v_p)$ in K one has

$$h_p(v_0, \cdots, v_p) \subset |v_0, \cdots, v_p|.$$

Then clearly one may define for these integers p a homomorphism

$$h_p^R\colon C_p^R \longrightarrow C_q^R$$

by the formula $h_p^R(v_0, \cdots, v_p, T)^R = (h_p(v_0, \cdots, v_p), T)^R$ in case $p \geq 0$, and one may make h_p^R the trivial homomorphism if $p < 0$. In view of the inclusions in §1.15, one observes that this definition creates the following homomorphisms in R (see [3, §3.1]):

$$j_p^R\colon C_p^R \longrightarrow C_{p-1}^R \quad (0 \leq j \leq p);$$

$$\beta_p^R\colon C_p^R \longrightarrow C_p^R \quad (-\infty < p < +\infty); \qquad \gamma_p^R\colon C_p^R \longrightarrow C_p^R \quad (p \geq 0);$$

$$\rho_{*p}^R\colon C_p^R \longrightarrow C_{p+1}^R \quad (-\infty < p < +\infty); \qquad b_{pk}^R\colon C_p^R \longrightarrow C_{p+1}^R \quad (0 \leq k \leq p);$$

$$t_{pj}^R\colon C_p^R \longrightarrow C_{p-1}^R \quad (0 \leq j \leq p-1); \qquad \tau_p^R\colon C_p^R \longrightarrow C_p^R \quad (\tau_p \in T_{p_0});$$

$$p_{pj}^R\colon C_p^R \longrightarrow C_p^R \quad (0 \leq j \leq p).$$

2.2. From the relations in §1.3, one derives the following (see [3, §3.1]):

$$\partial_p^R\, \beta_p^R = \beta_{p-1}^R\, \partial_p^R \qquad (-\infty < p < +\infty);$$

$$\beta_p^R\, t_{pj}^R = -\beta_p^R \qquad (0 \leq j \leq p-1);$$

$$\partial_{p+1}^R\, \rho_{*p}^R + \rho_{*p-1}^R\, \partial_p^R = \beta_p^R - 1 \qquad (0 \leq p < +\infty).$$

The theorems in §§1.6, 1,7 give rise to these formulas for β_p^R and ρ_{*p}^R:

$$\rho_{*0}^R = b_{00}^R,$$

$$\rho_{*p}^R = b_{pp}^R + \sum_{j=1}^{p} b_{p\,p-j}^R\, \gamma_p^R, \cdots, (\gamma_p^R - j + 1) \quad (p > 0);$$

$$\beta_0^R = 0_p^R\, b_{00}^R;$$

$$\beta_p^R = 0_{p+1}^R\, b_{p0}^R\, \gamma_p^R(\gamma_p^R - 1), \cdots, (\gamma_p^R - p + 1) \quad (p > 0).$$

From the theorem in §1.10, one obtains the following description for β_p^R and ρ_{*p}^R.

THEOREM. *The following relations hold*:

$$\beta_p^R = \mathfrak{d}_{p+1}^R \, \mathfrak{b}_{po}^R \, P_{po}^R = \sum_{\tau_p \in T_{po}} \mathfrak{d}_{p+1}^R \, \mathfrak{b}_{po}^R \, \tau_p^R \quad (p \geq 0);$$

$$\rho_{*p}^R = \sum_{k=0}^{p} \mathfrak{b}_{pk}^R \, P_{pk}^R = \sum_{k=0}^{p} \sum_{\tau_p \in T_{pk}} \mathfrak{b}_{pk}^R \, \tau_p^R \quad (p \geq 0).$$

2.3. The writer is indebted to T. Radó for suggestions which led to the results presently presented in §§2.3-2.7, 2.9, 2.10, 2,12. The new facts contributed by this paper are contained in §§2.8, 2.11, 2.13. For integers $p \geq 1$, any chain of the form $(1 + t_{pj}^R)\,(v_0, \cdots, v_p, T)^R$ $(0 \leq j \leq p-1)$ is termed an *elementary t-chain* in R (see [3, §3.2] or [4, §7]), and the subgroup of C_p^R generated by these elementary t-chains is denoted by T_p^R. For $p < 1$, T_p^R is defined to be the subgroup of C_p^R composed of the zero element alone.

LEMMA. *If* $c_p^R \in T_p^R$, *then*

(i) $\partial_p^R \, c_p^R \in T_{p-1}^R$,

(ii) $\beta_p^R \, c_p^R = 0$,

(iii) $\rho_{*p}^R \, c_p^R \in T_{p+1}^R$.

This lemma differs from that in Radó [3, §3.2], only by the fact that the barycentric homotopy operator ρ_p^R has been replaced by the modified operator ρ_{*p}^R (see §1.2). It may be established by the same reasoning as that employed by Radó.

2.4. For integers $p \geq 1$, any chain of the form

$$(v_0, \cdots, v_j, v_{j+1}, \cdots, v_p, T)^R$$

with $v_j = v_{j+1}$ for some j such that $0 \leq j \leq p-1$ is called an elementary d-chain in R (see [3, §3.3] or [4, §7]), and the subgroup of C_p^R generated by these elementary d-chains is denoted by D_p^R. For $p < 1$, D_p^R is defined to be that subgroup of C_p^R composed of the zero element alone.

LEMMA. *If* $c_p^R \in D_p^R$, *then*

(i) $\partial_p^R \, c_p^R \in D_{p-1}^R$,

(ii) $\beta_p^R \, c_p^R = 0$,

(iii) $\rho^R_{*p}\ c^R_p \in D^R_{p+1}$.

This is the lemma in [3, §3.3], except that the modified barycentric homotopy operator ρ^R_{*p} is used in place of ρ^R_p; it is proved in the same way.

2.5. LEMMA. *Let $(v_0, \cdots, v_p, T)^R$ be any p-cell in R ($p \geq 1$). Suppose that the sequence $w_0, \cdots, w_p$ is obtainable from the sequence $v_0, \cdots, v_p$ by n transpositions. Then there is an element t^R_p in T^R_p such that*

$$(v_0, \cdots, v_p, T)^R = (-1)^n (w_0, \cdots, w_p, T)^R + t^R_p.$$

Proof. By assumption there exist $n + 1$ sequences $v_{0j}, \cdots, v_{pj}$ for $0 \leq j \leq n$ where $v_{i0} = v_i$ and $v_{in} = w_i$ for $0 \leq i \leq p$ such that

$$(v_{0j}, \cdots, v_{pj}, T)^R = t^R_{pi_j} (v_{0\,j-1}, \cdots, v_{p\,j-1}, T)^R$$

for some integer i_j satisfying $0 \leq i_j \leq p - 1$, $1 \leq j \leq n$. Clearly

$$(v_0, \cdots, v_p, T)^R = (-1)^n (w_0, \cdots, w_p, T)^R$$
$$+ \sum_{j=1}^{n} (-1)^{j-1} (1 + t^R_{pi_j}) (v_{0\,j-1}, \cdots, v_{p\,j-1}, T)^R,$$

and the lemma is established.

2.6. LEMMA. *Let $(v_0, \cdots, v_p, T)^R$ be any p-cell in R ($p \geq 1$), for which $v_i = v_k$ for some i,k such that $0 \leq i \leq k \leq p$. Then there are elements t^R_p in T^R_p and d^R_p in D^R_p such that*

$$(v_0, \cdots, v_p, T)^R = t^R_p + d^R_p.$$

Moreover, $2(v_0, \cdots, v_p, T)^R$ is in T^R_p.

Proof. Since the sequence $v_0, \cdots, v_{i-1}, v_k, v_i, \cdots, v_{k-1}, v_{k+1}, \cdots, v_p$ is obtained from $v_0, \cdots, v_i, \cdots, v_k, \cdots, v_p$ by $k - i$ transpositions, and $v_i = v_k$ by assumption, if follows that

$$(-1)^{k-i} (v_0, \cdots, v_{i-1}, v_k, v_i, \cdots, v_{k-1}, v_{k+1}, \cdots, v_p)^R$$

is an element d^R_p of D^R_p. Moreover, from the lemma in §2.5 it follows that there is an element t^R_p in T^R_p such that $(v_0, \cdots, v_p, T)^R = d^R_p + t^R_p$, and the first part of the lemma is proven. Now the sequence $v_0, \cdots, v_k, \cdots, v_i, \cdots, v_p$ is obtained from $v_0, \cdots, v_i, \cdots, v_k, \cdots, v_p$ by $2(k - i) - 1$ transpositions.

Again, from the lemma in §2.5 it follows that there is an element t_p^R in T_p^R such that

$$(v_0, \cdots, v_i, \cdots, v_k, \cdots, v_p,\ T)^R = -(v_0, \cdots, v_k, \cdots, v_i, \cdots, v_p,\ T)^R + t_p^R.$$

Since $v_i = v_k$, one obtains $2(v_0, \cdots, v_p,\ T)^R = t_p^R$; and the second part of the lemma is demonstrated.

2.7. For integers $p \geq 0$, a chain c_p^R is termed an elementary n-chain in R if it has the form

$$c_p^R = \sum_{r=1}^{n} m_r(v_0, \cdots, v_p,\ T_r)^R,$$

where

(i) for $1 \leq r \leq n$, the m_r are nonzero integers;

(ii) for $1 \leq r_1 < r_2 \leq n$, the transformations T_{r_1} and T_{r_2} are not identical on $|v_0, \cdots, v_p|$;

(iii) the points $v_0, \cdots, v_p$ are distinct. The p-cell $(v_0, \cdots, v_p)$ in K (see §1.11) is called the base for c_p^R, and the notation $c_p^R = c_p^R(v_0, \cdots, v_p)$ is used when it is desirable to display the base.

2.8. LEMMA. *Suppose that c_p^R is an elementary n-chain in R for which $\beta_p^R c_p^R = 0$. Then $\beta_{p+1}^R \rho_{*p}^R c_p^R = 0$.*

Proof. With the notation of §2.7, one finds (see §§2.1, 2.2).

$$\text{(i)}\quad \beta_p^R c_p^R = \sum_{\tau_p \in T_{p0}} \sum_{r=1}^{n} m_r(0_{p+1}\, b_{p0}\, \tau_p(v_0, \cdots, v_p),\ T_r)^R = 0;$$

$$\text{(ii)}\quad \beta_{p+1}^R \rho_{*p}^R c_p^R = \sum_{\tau_{p+1} \in T_{p+1\,0}} \sum_{k=0}^{p} \sum_{\tau_p \in T_{pk}} \sum_{r=1}^{n} m_r(0_{p+2}\, b_{p+1\,0}\, \tau_{p+1}\, b_{pk}\, \tau_p(v_0, \cdots, v_p),\ T_r)^R.$$

In view of §2.7 (iii), and §1.11, it follows from (i) that for each $\tau_p' \in T_{p0}$, one has

$$\text{(iii)}\quad \sum_{r=1}^{n} m_r(0_{p+1}\, b_{p0}\, \tau_p'(v_0, \cdots, v_p),\ T_r)^R = 0 \qquad (\tau_p' \in T_{p0}),$$

Fix

$$\tau_{p+1} \in T_{p+10}, \ \tau_p \in T_{pk} \qquad (0 \leq k \leq p).$$

From the lemma in §1.14 follows the existence of a $\tau_p' \in T_{p0}$ such that

(iv) $|0_{p+2}\ b_{p+10}\ \tau_{p+1}\ b_{pk}\ \tau_p(v_0, \cdots, v_p)|$

$$\subset |0_{p+1}\ b_{p0}\ \tau_p'(v_0, \cdots, v_p)|.$$

From (iii) and (iv) one concludes that for each

$$\tau_{p+1} \in T_{p+10}, \ \tau_p \in T_{pk} \qquad (0 \leq k \leq p),$$

we have

(v) $$\sum_{r=1}^{n} m_r (0_{p+2}\ b_{p+10}\ \tau_{p+1}\ b_{pk}\ \tau_p(v_0, \cdots, v_p), T)^R = 0.$$

In view of (ii) and (v) the lemma is now established.

2.9. For integers $p \geq 0$, the class N_p^R is defined to be that subset of C_p^R composed of the chain $0 \in C_p^R$ and of all c_p^R having a representation of the form

$$c_p^R = \sum_{s=1}^{n} c_{ps}^R (v_{0s}, \cdots, v_{ps})$$

where

(i) for $1 \leq s \leq n$ the $c_{ps}^R(v_{0s}, \cdots, v_{ps})$ are elementary n-chains (see 2.7);

(ii) for $1 \leq s_1 < s_2 \leq n$, the point sets $v_{0s_1}, \cdots, v_{ps_1}$ and $v_{0s_2}, \cdots, v_{ps_2}$ are distinct. For $p < 0$, the class N_p^R consists of the chain $0 \in C_p^R$ alone. Each of the elementary n-chains $c_{ps}^R(v_{0s}, \cdots, v_{ps})$ $(1 \leq s \leq n)$, is termed a n-composant of c_p^R. Observe that the sets N_p^R are *not* generally subgroups of C_p^R.

2.10. LEMMA. *Let*

$$c_p^R = \sum_{s=1}^{n} c_{ps}^R (v_{0s}, \cdots, v_{ps})$$

be any nonzero element in N_p^R. A necessary and sufficient condition in order that

$\beta_p^R c_p^R = 0$ *is that* $\beta_p^R c_{ps}^R = 0$ *for every n-composant* c_{ps}^R $(1 \leq s \leq n)$.

Proof. Trivially the condition suffices. It is presently shown to be necessary. With explicit notations (see §§2.7, 2.9),

$$\beta_p^R c_p^R = \sum_{s=1}^{n} \beta_p^R c_{ps}^R = \sum_{s=1}^{n} \sum_{r=1}^{n_s} m_{rs} (\beta_p (v_{0s}, \cdots, v_{ps}) T_{rs})^R$$

$$= \sum_{s=1}^{n} \sum_{r=1}^{n_s} \sum_{\tau_p \in T_{p0}} m_{rs} (0_{p+1}\, b_{p0}\, \tau_p (v_{0s}, \cdots, v_{ps}), T_{rs})^R = 0.$$

In view of §2.9 (ii) and of the remarks in §1.11, it is clear (see §0.2) that, for $1 \leq s \leq n$ we have

$$\beta_p^R c_{ps}^R = \sum_{r=1}^{n_s} \sum_{\tau_p \in T_{p0}} m_{rs} (0_{p+1}\, b_{p0}\, \tau_p (v_{0s}, \cdots, v_{ps}), T_{rs})^R = 0$$

and hence the assertion in the lemma is verified.

2.11. LEMMA. *Let* c_p^R *be any element in* N_p^R *for which* $\beta_p^R c_p^R = 0$. *Then*

$$\beta_{p+1}^R \rho_{*p}^R c_p^R = 0.$$

This result is an immediate consequence of the lemmas in §§2.8, 2.10.

2.12. LEMMA. *Every chain* c_p^R *has a representation of the form* (see §§2.3, 2.4, 2.9)

$$c_p^R = t_p^R + d_p^R + n_p^R \qquad (t_p^R \in T_p^R,\ d_p^R \in D_p^R,\ n_p^R \in N_p^R).$$

Generally this representation is not unique.

Proof. The nonuniqueness of the representation will be evident from the proof of its existence which follows. For chains $c_p^R = 0 \in C_p^R$, the result is trivial, so assume that $c_p^R \neq 0$. Then c_p^R has a unique representation of the form

$$\text{(i)}\quad c_p^R = \sum_{j=1}^{n} m_j (v_{0j}, \cdots, v_{pj}, T_j)^R,$$

where the m_j are nonzero integers and the p-cells $(v_{0j_1}, \cdots, v_{pj_1}, T_{j_1})^R$ and $(v_{0j_2}, \cdots, v_{pj_2}, T_{j_2})^R$ are distinct for $1 \leq j_1 \leq j_2 \leq n$. The proof is made by an induction on n. If $n = 1$, then $c_p^R = m_1(v_{01}, \cdots, v_{p1}, T_1)^R$. If, for some integers i, k such that $0 \leq i < k \leq p$, one finds $v_{i1} = v_{k1}$, then the fact that c_p^R has a representation of the prescribed form follows from the lemma in §2.6. On the other hand, if all the $v_{01}, \cdots, v_{p1}$ are distinct, then c_p^R is an elementary n-chain (see §2.7). Thus the lemma is established in case $n = 1$. Suppose that the lemma is true for all chains c_p^R having a representation of the form (i) with at most $n = N - 1$ terms $(N > 1)$. For chains c_p^R whose representations (i) have N terms it is convenient to consider several cases.

Case 1. Assume there is some term in the representation (i) of c_p^R — without loss of generality one may assume it to be the first — for which there are integers i, k such that $0 \leq i < k \leq p$ and $v_{i1} = v_{k1}$. By the lemma in §2.6 there are elements t_{p1}^R in T_p^R and d_{p1}^R in D_p^R such that

$$m_1(v_{01}, \cdots, v_{p1}, T_1)^R = t_{p1}^R + d_{p1}^R.$$

By assumption there are elements t_{p2}^R in T_p^R, d_{p2}^R in D_p^R, and n_p^R in N_p^R such that

$$\sum_{j=2}^{N} m_j(v_{0j}, \cdots, v_{pj}, T_j)^R = t_{p2}^R + d_{p2}^R + n_p^R.$$

Thus

$$c_p^R = (t_{p1}^R + t_{p2}^R) + (d_{p1}^R + d_{p2}^R) + n_p^R,$$

and since T_p^R and D_p^R are subgroups of C_p^R, the existence of a representation of the prescribed form for c_p^R follows in Case 1.

Case 2. Assume that for each j $(1 \leq j \leq N)$ the $v_{0j}, \cdots, v_{pj}$ are distinct. By rearranging terms one may obtain from (i) a representation of the form

$$\text{(ii)}\quad c_p^R = \sum_{s=1}^{m} \sum_{r=1}^{n_s} m_{rs}(v_{0s}, \cdots, v_{ps}, T_{rs})^R, \quad \sum_{s=1}^{m} n_s = N,$$

satisfying these conditions: none of the m_{rs} is zero; for the same s $(1 \leq s \leq m)$, $1 \leq r_1 < r_2 \leq n_s$, the mappings $T_{r_1 s}$ and $T_{r_2 s}$ are not identical on $|v_{0s}, \cdots, v_{ps}|$; for $1 \leq s_1 < s_2 \leq m$, the p-cells $(v_{0s_1}, \cdots, v_{ps_1})$ and $(v_{0s_2}, \cdots, v_{ps_2})$ are distinct in K (see §1.1). Now for each s $(1 \leq s \leq m)$ clearly each of the chains

$$c_{ps}^R = \sum_{j=1}^{n_s} m_{rs}(v_{0s}, \cdots, v_{ps}, T_{rs})^R$$

is an elementary n-chain in R (see §2.7). The proof is carried forth by an inductive reasoning on m. If $m = 1$ then c_p^R is an elementary n-chain in R, and the representation (ii) already has the prescribed form. So assume that c_p^R, whose representation (i) has at most N terms, has a representation of the prescribed form whenever its representation (ii) has at most $m = M - 1$ terms ($M > 1$). Suppose now that C_p^R is a chain whose representation (i) has N terms while its representation (ii) has M terms

$$\sum_{s=1}^{M} n_s = N$$

Subcase 2.1. Assume that for $1 \leq s_1 < s_2 \leq M$ the point sets $v_{0s_1}, \cdots, v_{ps_1}$ and $v_{0s_2}, \cdots, v_{ps_2}$ are distinct. From §2.9 it is clear that c_p^R is itself an element in N_p^R and representation (ii) has the prescribed form.

Subcase 2.2. Assume that there are distinct integers s — with no loss of generality one may assume these to be $s = 1$ and $s = 2$ — such that the sets $v_{01}, \cdots, v_{p1}$ and $v_{02}, \cdots, v_{p2}$ are the same. It follows that the sequence $v_{02}, \cdots, v_{p2}$ is obtainable from $v_{01}, \cdots, v_{p1}$ by a positive number l of transpositions. Hence by the lemma in §2.5 there exists for each r in $1 \leq r \leq n_1$ an element t_{pr}^R in T_p^R such that

$$(v_{01}, \cdots, v_{p1}, T_{r1})^R = (-1)^l (v_{02}, \cdots, v_{p2}, T_{r1})^R + t_{pr}^R \qquad (1 \leq r \leq n_1).$$

Since T_p^R is a subgroup of C_p^R, the chain

$$\sum_{r=1}^{n_1} m_{r1} t_{pr}^R$$

is an element t_{p*}^R in T_p^R. Consequently,

$$c_p^R = t_{p*}^R + \left[\sum_{r=1}^{n_1} (-1)^l m_{r1} (v_{02}, \cdots, v_{p2}, T_{r1})^R \right.$$
$$\left. + \sum_{s=2}^{M} \sum_{r=1}^{n_s} m_{rs} (v_{0s}, \cdots, v_{ps}, T_{rs})^R \right].$$

Clearly the terms in square brackets may be rearranged into the form (ii) with an integer $m \leq M - 1$, and their representation in form (i) has an integer $n \leq N$. By the inductive assumption there are elements $t_{p\#}^R$ in T_p^R, d_p^R, in D_p^R and n_p^R in

N_p^R such that $c_p^R = (t_{p*}^R + t_{p\#}^R) + d_p^R + n_p^R$, and the existence of a representation of the prescribed form for c_p^R now follows in Case 2. Indeed, it is obvious in this case that $d_p^R = 0 \in C_p^R$. So the lemma is completely established.

2.13. LEMMA. *If c_p^R is any chain in C_p^R for which $\beta_p^R\, c_p^R = 0$, then*

$$\beta_{p+1}^R\, \rho_{*p}^R\, c_p^R = 0.$$

The proof follows at once from the lemmas in §§2.3, 2.4, 2.11, 2.12.

RESULTS

3.1. In [3, §4.1] (see also [4, §8]) Radó has established a lemma from which one derives the following statement by replacing the barycentric homotopy operator ρ_p^R by the modified barycentric homotopy operator ρ_{*p}^R (see §§1.2, 2.1).

LEMMA. *Let $\{G_p\}$ be an identifier for R* (see §0.3) *such that the following conditions hold:*

(i) *$c_p^R \in G_p$ implies that $\beta_p^R\, c_p^R = 0$;*

(ii) *$c_p^R \in G_p$ implies that $\rho_{*p}^R\, c_p^R \in G_{p+1}$.*

Then $\{G_p\}$ is unessential.

3.2. For each integer p let $N(\beta_p^R)$ be the nucleus of the homomorphism $\beta_p^R\colon C_p^R \longrightarrow C_p^R$ (see §2.1). Since β_p^R is a chain mapping (see §2.2) it is clear that the nuclei $N(\beta_p^R)$ constitute an identifier for R (see §0.3). Now in view of the lemma in §2.13, conditions (i) and (ii) of the lemma above are clearly fulfilled for the identifier $\{N(\beta_p^R)\}$, and furthermore, this choice of an identifier yields the maximum amount of information that may be obtained from that lemma. Thus the $\{N(\beta_p^R)\}$ constitute an unessential identifier for R, and one of the main results is now established (see §0.4). It is summarized in the following statement.

THEOREM. *The system of nuclei $N(\beta_p^R)$ of the barycentric homomorphisms $\beta_p^R\colon C_p^R \longrightarrow C_p^R$ constitutes an unessential identifier for R.*

3.3. In order to compare this result with those in Radó [3; 4], first observe that it follows from the lemmas in §§2.3, 2.4 that

$$N(\beta_p^R) \supset T_p^R + D_p^R \qquad (-\infty < p < +\infty).$$

Moreover, since C_p^R is a free group, it is clear that the division hull of $N(\beta_p^R)$

must be identical with the group $N(\beta_p^R)$. Thus the group $N(\beta_p^R)$ also contains the the division hull of the group $T_p^R + D_p^R$ for all integers p. An example is now given to show that the group $N(\beta_p^R)$ generally contains more.

3.4. Denote by d_0, d_1, d the points $(1, 0, 0, \cdots)$, $(0, 1, 0, 0, \cdots)$, $(1/2, 1/2, 0, 0, \cdots)$ respectively, let X be Euclidean x-space, and define transformations by the following relations:

$$T_1\colon x = v_0 - 1/2 \qquad (v \in |d_0, d_1|);$$

$$T_2\colon x = \begin{cases} 0 & (v \in |d_0, d\ |); \\ v_0 - 1/2 & (v \in |d\ , d_1|); \end{cases}$$

$$T_3\colon x = \begin{cases} v_0 - 1/2 & (v \in |d_0, d\ |); \\ 0 & (v \in |d\ , d_1|); \end{cases}$$

$$T_4\colon x = 0 \qquad (v \in |d_0, d_1|).$$

Clearly

$$c_1^R = (d_0, d_1, T_1)^R - (d_0, d_1, T_2)^R - (d_0, d_1, T_3)^R + (d_0, d_1, T_4)^R$$

belongs to C_1^R and $\beta_1^R\, c_1^R = 0$. Moreover, c_1^R is an elementary n-chain (see §2.7). An elementary reasoning shows that it cannot belong to the division hull for the group $T_1^R + D_1^R$.

3.5. In order to describe the largest unessential identifier for R obtained by Radó, a further definition is needed. For integers $p \geq 0$, let $(v_0, \cdots, v_p, T)^R$ be any p-cell in R (see §0.1). Let $w_0, \cdots, w_p$ be any set sequence of $p + 1$ linearly independent points in E_∞. Then there is a linear mapping

$$\alpha\colon |w_0, \cdots, w_p| \longrightarrow |v_0, \cdots, v_p$$

such that $\alpha(w_i) = v_i$ for $0 \leq i \leq p$. The p-chain

$$c_p^R = (v_0, \cdots, v_p, T)^R - (w_0, \cdots, w_p, T\alpha)^R$$

is termed an *elementary a-chain* in R (see [3, §3.4]), and the subgroup of C_p^R generated by the elementary a-chains is denoted by A_p^R. For $p < 0$, A_p^R consists of the zero element alone. In [3, §3.4] Radó has a simple characterization for the group A_p^R which he uses to define the group in [4, §7].

3.6. For each integer p, put $\Gamma_p^R = A_p^R + D_p^R + T_p^R$ (see §§2.3, 2.4, 3.5), and let $\hat{\Gamma}_p^R$ denote the division hull of Γ_p^R. Then Radó shows that $\{\hat{\Gamma}_p^R\}$ is an unessential identifier in R (see [3, §4.7] or [4, §9]), and this is his best result. If one sets $\Delta_p^R = A_p^R + N(\beta_p^R)$ (see §3.2) and lets $\hat{\Delta}_p^R$ denote the division hull of Δ_p^R, then clearly $\Delta_p^R \supset \Gamma_p^R$, and hence $\hat{\Delta}_p^R \supset \hat{\Gamma}_p^R$. If one modifies the reasoning of Radó in [3, §4] by replacing the barycentric homotopy operator ρ_p^R by the modified barycentric homotopy operator ρ_{*p}^R (see §2.1), one finds that $\hat{\Delta}_p^R$ is an unessential identifier for R. Thus one obtains the following result.

THEOREM. *If $\hat{\Delta}_p^R$ is the division hull of the group $A_p^R + N(\beta_p^R)$ then the system $\{\hat{\Delta}_p^R\}$ is an unessential identifier for R.*

REFERENCES

1. S. Eilenberg and N. E. Steenrod, *Foundations of Algebraic Topology* (unpublished).

2. T. Radó, *On identifications in singular homology theory*, Rivista Mat. Univ. Parma, 2 (1951), 3-18.

3. ———, *An approach to singular homology theory*, Pacific J. Math. **1** (1951), 265-290.

4. ———, *On the foundations of singular homology theory*, to appear in Mem. Accad. Sci. Ist. Bologna, Cl. Sci. Fis.

OHIO STATE UNIVERSITY

INCIDENCE RELATIONS IN MULTICOHERENT SPACES III

A. H. STONE

1. Introduction.

1.1. PRELIMINARIES. The present paper is concerned with relations between systems of sets and their frontiers in a locally connected space S of given degree of multicoherence, $r(S)$. The results are generalizations of those derived in [4] for the unicoherent case $[r(S) = 0]$, and of those in [5] for the case of two sets; the methods are those used in [5] and [6]. First we apply the "analytic" method (cf. [1; 2; 6]) to obtain a general "addition theorem" for arbitrary sets with "nearly disjoint" frontiers (Theorem 1), which is shown to be "best possible" (Theorem 2), and to derive also relations between arbitrary systems of sets and their frontiers (Theorems 3 and 4). Next (§4) we consider a function of sets which measures (roughly speaking) the amount of disconnectedness of the frontiers of the components of the complementary set, and, after deriving some of its properties, use it to extend the Phragmén-Brouwer theorem to arbitrary sets (Theorem 6), and to obtain some related results. A modified "addition theorem" is then established (Theorem 9) which includes both Theorem 1 and Theorem 6 as special cases. Finally, we consider the incidences of sets with disjoint frontiers and subject to further restrictions (for example, that the sets be connected and have connected complements), showing that many problems of this type can be reduced to purely combinatorial problems in graph-theory.

1.2. NOTATIONS. We shall be concerned throughout with subsets of a fixed *nonempty, connected, locally connected, completely normal*[1] T_1 *space*, S. The notations are, in general, the same as in [4; 5; 6]; but the following items are repeated for the convenience of the reader.

The number of components, less one, of a set E, is denoted by $b_0(E)$; thus $b_0(0) = -1$. If the number of components of E is infinite, we write $b_0(E) = \infty$, without distinction as to cardinality. The *degree of multicoherence* of S is defined by $r(S) = \sup b_0(A \cap B)$, where A and B are closed connected sets such that $A \cup B = S$. It is known [5] that "closed" can be replaced by "open" here.

If $A_1, A_2, \cdots, A_n$ are any n sets (that is, subsets of S), and J is any nonempty collection of distinct suffixes

[1] As was remarked in [5, §6.6(3)], there would be no difficulty in reformulating the theorems so as to apply if complete normality were weakened to normality.

Received January 10, 1951.
Pacific J. Math. 2 (1952), 99-126

$$j_1, j_2, \cdots, j_k \qquad (1 \leq j_1 < j_2 < \cdots < j_k \leq n),$$

we write A_J as an abbreviation for $A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}$, and write

$$\bigcup \{ A_J \mid |J| = k \} \text{ as } X_k(A_1, A_2, \cdots, A_n),$$

or simply as X_k. Thus

$$\bigcup A_j = X_1 \supset X_2 \supset \cdots \supset X_n = \bigcap A_j .$$

For convenience, we introduce the conventions $X_0 = S$ and $X_k = 0$ if $k > n$. We write

(1) $$h(A_1, A_2, \cdots, A_n) = \sum b_0(X_k) - \sum b_0(A_k) \qquad (1 \leq k \leq n),$$

with the convention that in interpreting an equality or inequality involving $h(A_1, \cdots, A_n)$ in which $\sum b_0(A_k) = \infty$, we first transpose all negative terms. If the sets A_j are all closed, or all open, or more generally have separated differences[2], it is known [4, Th. 6b] that $h(A_1, \cdots, A_n) \geq 0$.

Again, following Eilenberg [1], we consider (continuous) mappings f of subsets of S into the circle S^1 of complex numbers of unit modulus, and write "$f \sim 1$ on X" to mean that there exists a real (continuous) function ϕ on X such that $f(x) = \exp[i\phi(x)]$ when $x \in X$. Mappings $f_1, f_2, \cdots, f_m$ of X in S^1 are *independent* on X if the only (positive or negative) integers $p_1, p_2, \cdots, p_m$, for which the product (in the sense of complex numbers)

$$f_1^{p_1} f_2^{p_2} \cdots f_m^{p_m} \sim 1 \text{ on } X,$$

are $p_1 = p_2 = \cdots = p_m = 0$. If $A_1, A_2, \cdots, A_n$ are closed sets whose union is X, the greatest number of mappings f of X in S^1 which are independent on X and such that $f \sim 1$ on each A_j (or ∞ if there is no such greatest number) is denoted by $p(A_1, A_2, \cdots, A_n)$. For fixed X and n, we write

(2) $$r_n(X) = \sup p(A_1, \cdots, A_n),$$

the supremum being taken over all systems of n closed sets $A_1, \cdots, A_n$ whose union is X. Clearly $0 = r_1(X) \leq r_2(X) \leq \cdots$; it is known [1] that

$$\sup_n r_n(X) = b_1(X)$$

and [1; 6] that $r_2(S) = r(S)$.

---

[2] That is, $A_j - A_k$ and $A_k - A_j$ are separated ($1 \leq j < k \leq n$). (Two sets are "separated" if neither meets the closure of the other.)

1.3. SOME LEMMAS. We shall require the following lemmas, some of which are known; the proofs of the rest are easy.

(1) If $A_1, A_2, \cdots, A_n$ have separated differences, then

(i) $\bigcup \mathrm{Fr}(A_j) = \bigcup \mathrm{Fr}(X_j)$;

(ii) $A_j \cap A_k$ and $\mathrm{Co}(A_j \cup A_k)$ are separated $(1 \leq j < k \leq n)$ if and only if $\mathrm{Cl}(X_k) \subset X_j$ and $X_k \subset \mathrm{Int}(X_j)$;

(iii) $\mathrm{Fr}(A_j \cap A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ $(1 \leq j < k \leq n)$ if and only if $X_1, X_2, \cdots, X_n$ have disjoint[3] frontiers; that is, $\mathrm{Cl}(X_k) \subset \mathrm{Int}(X_j)$;

(iv) $A_1, A_2, \cdots, A_n$ are of finite incidence[4] if and only if

$$\sum b_0(X_j) < \infty.$$

(2) If A_1 and A_2 are both open, or both closed, then $A_1 - A_2$ and $A_2 - A_1$ are separated; and further, $A_1 \cap A_2$ and $\mathrm{Co}(A_1 \cap A_2)$ are separated if and only if $\mathrm{Fr}(A_1 \cap A_2) \cap \mathrm{Fr}(A_1 \cup A_2) = 0$. If A_1 and A_2 are open, this condition is equivalent to $\mathrm{Fr}(A_1) \cap \mathrm{Fr}(A_2) \cap \mathrm{Fr}(A_1 \cap A_2) = 0$.

(3) "Approximation lemma." If $A_j - A_k$ and $A_k - A_j$ are separated, and also $A_j \cap A_k$ and $\mathrm{Co}(A_j \cup A_k)$ are separated $(1 \leq j < k \leq n)$, then, given any open sets $W(J) \supset A_J$ (where J runs over all nonempty sets of suffixes between 1 and n), there exist open sets $A_j^* \supset A_j$ such that, for any open sets B_j satisfying $A_j \subset B_j \subset A_j^*$, we have $B_J \subset W(J)$ and

$$\mathrm{Fr}(B_j) \cap \mathrm{Fr}(B_k) \cap \mathrm{Fr}(B_j \cap B_k) = 0 \quad (1 \leq j < k \leq n).$$

If further $\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ $(j \neq k)$, the sets A_j^* can be chosen so that the sets B_j have disjoint frontiers.

(If $n = 2$, this reduces to [5, Ths. 7 and 7a]; the general case follows by a straightforward induction over n.)

(4) If $A_1, A_2, \cdots, A_n$ are closed sets of finite incidence, then

$$p(A_1, A_2, \cdots, A_n) \leq h(A_1, A_2, \cdots, A_n);$$

if further no three of the sets A_j have a common point (for example, if $n = 2$), then $p = h$. (Cf. [6, §2.6].)

[3] Throughout this paper, "disjoint" means "pairwise disjoint".

[4] That is, the sets A_j and all their intersections A_J have only finitely many components.

(5) If f maps X in S^1, and X is a finite union of disjoint closed sets on each of which $f \sim 1$, then $f \sim 1$ on X. (Trivial.)

(6) If f maps S in S^1, and $f \sim 1$ on a closed set $A \subset S$, then there exists an open set $V \supset A$ such that $f \sim 1$ on V. (Cf. [1, p. 157; 6, §2.2 (2)].)

(7) If f maps A in S^1, and $f \sim 1$ on $\mathrm{Fr}(A)$, then f may be extended to a mapping f^* of S in S^1 such that $f^* \sim 1$ on $\mathrm{Cl}(S - A)$.

For $f = \exp(i\phi)$ on $\mathrm{Fr}(A)$; since $\mathrm{Cl}(S - A)$ is normal, ϕ can be extended to a continuous real function ϕ^* on $\mathrm{Cl}(S - A)$; define $f^* = \exp(i\phi^*)$ on $\mathrm{Cl}(S - A)$, and $f^* = f$ elsewhere.

(8) *If $A_1, A_2, \cdots, A_n$ are n closed sets, and $1 \leq m \leq n$, then*

$$p(A_1, A_2, \cdots, A_n) \leq r_m(X_1) + r_{n+1-m}(X_m) \leq r_m(X_1) + b_1(X_m).$$

For consider N mappings $f_1, \cdots, f_N$ of X_1 $(= \bigcup A_j)$ in S^1 which are independent on X_1 and satisfy $f_k \sim 1$ on A_j $(1 \leq k \leq N,\ 1 \leq j \leq n)$. We must prove

$$N \leq r_m(X_1) + r_{n+1-m}(X_m).$$

Let s be the greatest number of mappings f_k which are independent on X_m; since $X_m \subset A_1 \cup A_2 \cup \cdots \cup A_{n+1-m}$, clearly $s \leq r_{n+1-m}(X_m)$. We may suppose that the mappings f_k are independent on X_m for $N - s < k \leq N$, and then have, for each $k \leq N - s$, a relation of the form

$$g_k \equiv f_k^{p_k} \prod_{t > N-s} f_t^{q_{kt}} \sim 1$$

on X_m, where the exponents p_t, q_{kt} are integers not all zero, so that clearly $p_k \neq 0$. It readily follows that the mappings g_k $(1 \leq k \leq N - s)$ of X_1 in S^1 are independent on X_1, and they clearly satisfy $g_k \sim 1$ on each A_j. Further, from (6) above, there exists an open set $V_m \supset X_m$ such that each $g_k \sim 1$ on $\mathrm{Cl}(V_m)$. Now $X_{m-1} - V_m$ is a finite union of disjoint closed sets of the form $A_J - V_m$ (where $|J| = k - 1$), on each of which each $g_k \sim 1$: hence, by (5), $g_k \sim 1$ on $X_{m-1} - V_m$, so that there exists an open set $V_{m-1} \supset X_{m-1} - V_m$ such that each $g_k \sim 1$ on $\mathrm{Cl}(V_{m-1})$. Proceeding in this way, we obtain open sets

$$V_\lambda \supset X_\lambda - (V_{\lambda+1} \cup V_{\lambda+2} \cup \cdots \cup V_m) \quad (1 \leq \lambda \leq m)$$

such that each $g_k \sim 1$ on $\mathrm{Cl}(V_\lambda)$. Since $\bigcup \mathrm{Cl}(V_\lambda) \supset X_1$, the number $N - s$ of mappings g_k is at most $p(\overline{V}_1, \overline{V}_2, \cdots, \overline{V}_m) \leq r_m(X_1)$, and the result follows.

As corollaries, we have:

(9) If, in the proof of (8), each of the mappings $f_k \sim 1$ on X_m, then $N \leq r_m(X_1)$.

(10) *If no* $m+1$ *of the sets* A_j *in* (8) *can have a common point* (for example, if $m = n$), *then* $p(A_1, A_2, \cdots, A_n) \leq r_m(X_1)$.

For in this case, X_m falls into disjoint closed sets A_J, each contained in a single A_j; hence, from (5), each $f_k \sim 1$ on X_m.

2. An additional theorem.

2.1. INTRODUCTION. The last result, 1.3 (10), combined with 1.3 (4), gives another proof of the fact [6, Ths. 3 and 4a] that if $A_1, A_2, \cdots, A_n$ are closed sets which cover S, and no three of them have a common point, then

$$h(A_1, \cdots, A_n) \leq r(S).$$

In the present section we shall obtain a considerable extension of this property (Theorem 1), and show that it is the "best possible" of its kind, incidentally obtaining a new characterization of $r(S)$ (Theorem 2).

2.2. THEOREM 1. *Let* $A_1, A_2, \cdots, A_n$ *be any subsets of* S *having separated differences and such that* $A_j \cap A_k$ *and* $\mathrm{Co}(A_j \cup A_k)$ *are separated whenever* $j \neq k$.[5] *Suppose that no point belongs to* A_j *for more than m distinct values of j, where* $2 \leq m \leq n$.[6] *Then*

$$0 \leq h(A_1, A_2, \cdots, A_n) \leq (m-1)r(S).$$

Proof. Clearly we may assume that $r(S)$ and $b_0(A_j)$ are finite ($1 \leq j \leq n$); from [5, Th. 9], the sets A_j are then of finite incidence. Further, it will suffice to prove the theorem under the additional assumptions that the sets A_j are *closed* and have *disjoint frontiers*. For if the theorem is known in this case, the method of "approximation" extends it first [applying the second part of 1.3 (3) to the sets $\mathrm{Co}(A_j)$] to the case in which the sets A_j are open and satisfy

[5]These hypotheses are implied by: (a) the sets $\mathrm{Fr}(A_j)$ are disjoint, or (b) $A_1, \cdots, A_n$ are all open, or all closed, and $\mathrm{Fr}(A_j \cap A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ whenever $j \neq k$, or (c) $A_1, \cdots, A_n$ are all closed and $\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ ($j \neq k$), or dually, and thus also by: (d) $A_1, \cdots, A_n$ are closed and cover S, and no three of them have a common point. A slight relaxation of the hypotheses on the sets A_j is possible; see 2.3 (3) below.

[6]The case $m = 1$ is trivial. If equality holds in the conclusion of Theorem 1, and both sides are finite, then the sets A_j must in fact satisfy stronger frontier conditions; see 5.6 below.

$$\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cap A_k) = 0 \qquad (j \neq k),$$

and thence [by the first part of 1.3(3)] to the general case; we omit the details, since the argument is a straightforward generalization of that in [5, §§7.4 and 7.5] (cf. also [6, §4.4]).

We write $X_s(A_1, A_2, \cdots, A_t)$ as X_s^t $(1 \leq s \leq t \leq n)$, and introduce the conventions $X_s^t = S$ if $1 \leq s \leq n < t$, or if $0 = s < t$, and $X_s^t = 0$ if $s > t$. Now (all the numbers involved being finite here) one readily verifies that

$$(1) \quad h(A_1, A_2, \cdots, A_n) = h(A_1, A_2, \cdots, A_{n-1}) + \sum h(A_n \cap X_{s-1}^{n-1}, X_s^{n-1}) \quad (1 \leq s \leq n-1),$$

and repeated application of this identity gives

$$(2) \quad h(A_1, A_2, \cdots, A_n) = \sum\nolimits_1 + \sum\nolimits_2 + \cdots + \sum\nolimits_{n-1},$$

where

$$\sum\nolimits_s = \sum\nolimits_t h(A_{t+1} \cap X_{s-1}^t, X_s^t) \qquad (s \leq t \leq n-1).$$

We first show that

$$(3) \quad \sum\nolimits_s \leq r(S) \qquad (1 \leq s \leq n-1).$$

For, from 1.3(4), we have

$$\sum\nolimits_s = \sum\nolimits_t p(A_{t+1} \cap X_{s-1}^t, X_s^t) \qquad (s \leq t \leq n-1).$$

Let f_{tj}, where $j = 1, 2, \cdots, n_t$, be mappings of

$$X_s^{t+1} = (A_{t+1} \cap X_{s-1}^t) \cup X_s^t$$

in the unit circle such that

(i) $f_{tj} \sim 1$ on $A_{t+1} \cap X_{s-1}^t$,

(ii) $f_{tj} \sim 1$ on X_s^t,

(iii) for fixed t, these mappings are independent on X_s^{t+1}.

To prove (3), it suffices to show that the total number $\sum n_t$

$(s \leq t \leq n-1)$ of these mappings is at most $r(S)$.

We have

$$\mathrm{Fr}[\mathrm{Cl}(S - X_s^{t+1})] \subset \mathrm{Fr}(X_s^{t+1}) \subset \mathrm{Fr}(A_{t+1} \cap X_{s-1}^t) \cup \mathrm{Fr}(X_s^t),$$

a union of two closed sets which are easily seen [from 1.3(1)] to be disjoint. Hence [from 1.3(5) and 1.3(7)] $f_{tj} \sim 1$ on $\mathrm{Fr}[\mathrm{Cl}(S - X_s^{t+1})]$, and so f_{tj} can be extended to a mapping, which we still denote by f_{tj}, of S in the unit circle, in such a way that

(iv) $f_{tj} \sim 1$ on $\mathrm{Cl}(S - X_s^{t+1})$.

We assert that the extended mappings f_{tj} are *all* independent on S. For suppose not; then, for each t, there exists a mapping of the form

$$g_t = \Pi_j f_{tj}^{p_{tj}} \qquad (1 \leq j \leq n_t).$$

where the exponents p_{tj} are positive or negative integers, not all zero for all t, such that

(4) $g_s\, g_{s+1} \cdots g_{n-1} \sim 1$ on S.

From (ii), we have $g_t \sim 1$ on X_s^t; and so, if $t > s$, we have $g_t \sim 1$ on X_s^{s+1}. Thus (4) gives $g_s \sim 1$ on X_s^{s+1}; hence, from (iii), it follows that $g_s = 1$, and all the exponents p_{sj} are zero. A similar argument, with s replaced by $s+1$, then proves $g_{s+1} = 1$, and so on; finally all the exponents p_{tj} must be zero, giving the desired contradiction.

Now write

$$E_k = \mathrm{Cl}(X_s^{s+k-1} - X_s^{s+k-2}), \qquad k = 1, 2, \cdots, n+2-s;$$

thus the sets E_k are closed and cover S, and it is easy to see that no three of them have a common point. We shall show:

(5) $f_{tj} \sim 1$ on E_k.

In fact, if $k \leq t+1-s$, then $E_k \subset X_s^{s+k-1} \subset X_s^t$; if $k = t+2-s$, then $E_k \subset A_{t+1} \cap X_{s-1}^t$; and if $k \geq t+3-s$, then $E_k \subset \mathrm{Cl}(S - X_s^{t+1})$; thus in each case (5) follows from (ii), (i), or (iv).

Thus the total number of mappings f_{tj} is at most

$$p(E_1, E_2, \cdots, E_{n+2-s});$$

but, by 1.3(10), this number is at most $r_2(\cup E_k) = r(S)$; thus (3) is established.

Now we further have $\sum_s = 0$ if $s \geq m$, since the sets $A_{t+1} \cap X^t_{s-1}$ and X^t_s are then disjoint (for $X_{s+1} = 0$). Thus the theorem follows from (2) and (3).

2.3. Corollaries and Remarks. We make the following observations.

(1) *For any two sets A, B, satisfying the hypotheses of Theorem* 1, *we have*

$$b_0(A) + b_0(B) \leq b_0(A \cup B) + b_0(A \cap B) \leq b_0(A) + b_0(B) + r(S).$$

(this generalizes [5, Th. 9].)

(2) *For any set E, we have*

$$b_0(\mathrm{Fr}(E)) \leq b_0(\overline{E}) + b_0(\mathrm{Cl}(\mathrm{Co}(E)) + r(S).$$

(this generalizes [4, §6.5].)

(3) *In Theorem* 1, *the hypothesis that $A_j \cap A_k$ and* $\mathrm{Co}(A_j \cup A_k)$ *be separated* ($j \neq k$) *may be omitted for each pair j, k for which $A_j \subset A_k$*; that is, it may be replaced by: For each j, k ($1 \leq j, k \leq n$), either $A_j \subset A_k$, or $A_j \supset A_k$, or $A_j \cap A_k$ and $\mathrm{Co}(A_j \cup A_k)$ are separated. This is proved by noting that a more careful application of the approximation argument will still lead to closed sets with disjoint frontiers.

(4) Other results may be derived by observing that, under suitable conditions on the sets $A_1, \cdots, A_n$, further sums $\sum_s$ in 2.2(1) above will vanish. For example, Theorem 1 can be slightly sharpened as follows:

If $A_1, \cdots, A_n$ satisfy the hypotheses of Theorem 1 (*as weakened in* (3) *above*), *and if they can be renumbered so that*

$$A_{\lambda+1} \subset A_{\lambda+2} \subset \cdots \subset A_n,$$

then

$$h(A_1, \cdots, A_n) \leq \min(\lambda, m-1)\, r(S).$$

For the approximation argument enables us to assume, as before, that the sets $A_1, \cdots, A_n$ are closed and have disjoint frontiers. In 2.2(2) we easily verify that now $X^t_s \subset A_{t+1} \cap X^t_{s-1}$ whenever

$$\lambda + 1 \leq s \leq t \leq n - 1;$$

hence $\sum_s = 0$ whenever $s > \lambda$.

(5) A further slight sharpening of Theorem 1 is implied by the following result.

If the sets $A_1, \cdots, A_n$ have separated differences, and if A_n (say) is either disjoint from, or contains, or is contained in, each other set, then

$$h(A_1, \cdots, A_{n-1}, A_n) = h(A_1, \cdots, A_{n-1}).$$

We may assume that A_n is disjoint from $A_1, \cdots, A_k$, contains

$$A_{k+1}, \cdots, A_l,$$

and is contained in $A_{l+1}, \cdots, A_{n-1}$ (where $0 \leq k \leq l \leq n - 1$). It is easy to see that we may take $A_1, \cdots, A_n$ to be of finite incidence, and then, by 2.2(1), have only to prove that

$$h(A_n \cap X_{s-1}^{n-1}, X_s^{n-1}) = 0 \qquad (1 \leq s \leq n - 1).$$

If $s < n - l$, then $A_n \subset X_s^{n-1}$, and the result is trivial. If $s \geq n - l$, write

$$Y_p = X_p(A_1, \cdots, A_k, A_{l+1}, \cdots, A_{n-1})$$

and

$$Z_q = Z_q(A_{k+1}, \cdots, A_l, A_{l+1}, \cdots, A_{n-1});$$

it is easily verified that $X_s^{n-1} \mp Y_s \cup Z_s$ and that $Y_s \subset \mathrm{Co}(A_n)$ and $Z_s \subset A_n$, from which again the result follows.

(6) Finally, as a corollary from (4), we have the following extension of (1):

If $B_1, \cdots, B_p, C_1, \cdots, C_q$ are arbitrary sets such that $B_j - C_k$ and $C_k - B_j$ are separated, and $B_j \cap C_k$ and $\mathrm{Co}(B_j \cup C_k)$ are separated, whenever $1 \leq j \leq p$, $1 \leq k \leq q$, then

$$h(B_1, \cdots, B_p) + h(C_1, \cdots, C_q) \leq h(B_1, \cdots, B_p, C_1, \cdots, C_q)$$

$$\leq h(B_1, \cdots, B_p) + h(C_1, \cdots, C_q) + \min(p, q, m - 1)\, r(S),$$

where m is the greatest number of the $p + q$ sets $B_1, \cdots, C_q$ which have a common point.

This follows on application of (4) and Theorem 1 to the $p + q$ sets $X_j(B_1, \cdots, B_p)$, $X_k(C_1, \cdots, C_q)$.

2.4. CONVERSE. The converse of Theorem 1 holds in the following rather strong form, which represents an extension to any number of sets of the defining property of $r(S)$.

THEOREM 2. *Let integers m, n be given, where* $2 \leq m \leq n$. *Let* $A_1, \cdots, A_n$ *be any n closed connected sets, no* $m + 1$ *of which have a common point, such that* $\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) = 0$ *whenever* $j \neq k$, *and such that* $A_j \cup A_k = S$ *whenever* $1 \leq j < k \leq m$.[7] *Then*

$$\sup b_0(X_m) = (m - 1)\, r(S) + n - m.$$

In this statement, the word "closed" may be replaced by "open".

To show that

(1) $b_0(X_m) \leq (m - 1)\, r(S) + n - m,$

we clearly may assume $X_m \neq 0$; then $b_0(X_s) \geq 0$ if $s \leq m$, and

$$b_0(X_s) = -1$$

for $m < s \leq n$, so that (1) is a trivial consequence of Theorem 1.

To complete the proof, let N be any integer such that

$$0 \leq N \leq (m - 1)\, r(S).$$

We first construct m closed connected sets $B_1, B_2, \cdots, B_m$, such that

(2) $B_j \cup B_k = S \ (1 \leq j < k \leq m)$ and $b_0(\bigcap B_j) \geq N$.

If $r(S) = \infty$, this is trivial (take all but two of the sets B_j to be S), so we may assume $r(S) < \infty$. From [6, §4.1], there exists a finite covering of S by closed connected sets $E_1, E_2, \cdots, E_M$, no three of which have a common point, whose nerve G satisfies $r(G) = r(S) = r$, say, and such that G is arbitrarily often "dispersed"; this implies [6, §3.4(7)] that G is obtainable from a graph H by subdividing each arc l_λ of H which belongs to a simple closed curve in H, into at least $2m + 2$ subarcs by extra vertices of order 2. We can select[8] r such (disjoint, open)

[7]Note that we do not require *every* two sets A_j, A_k to cover S. In fact, if n nonempty closed sets are such that every two of them cover S, then trivially all of them have a common point.

[8]See, for example, the argument proving [6, §4.1(3)].

arcs l_λ in H, say $l_1, l_2, \cdots, l_r$, whose removal does not disconnect H; let l_λ (where $1 \leq \lambda \leq r$) contain the consecutive vertices $p_{\lambda,0}, p_{\lambda,1}, p_{\lambda,2}, \cdots, p_{\lambda,2m}$ of order 2 in G. Denote by $E_{\lambda,j}$ the set E_k which corresponds to $p_{\lambda,j}$; thus, if $1 \leq \lambda \leq r$ and $1 \leq j \leq 2m-1$, each $E_{\lambda,j}$ meets two and only two other sets E_k, namely $E_{\lambda,j-1}$ and $E_{\lambda,j+1}$. Define B_q, where $1 \leq q \leq m$, to be the union of all the sets E_k *except*

$$E_{1,2q-1}, E_{2,2q-1}, \cdots, E_{r,2q-1}.$$

Then B_q is closed, and is easily seen to be connected (cf. [6, Th. 1]). Further, since $\operatorname{Co}(B_q) \subset \bigcup_\lambda E_{\lambda,2q-1}$, we have $\operatorname{Co}(B_q) \cap \operatorname{Co}(B_s) = 0$ if $q \neq s$, so that $B_q \cup B_s = S$. On the other hand, let D be the union of those sets E_k which are not of the form $E_{\lambda,j}$ $(1 \leq \lambda \leq r,\ \ 1 \leq j \leq 2m-1)$; then

$$\bigcap B_q \subset D \cup \bigcup E_{\lambda,2h} \qquad (1 \leq \lambda \leq r,\ 1 \leq h \leq m-1),$$

a union of $1 + (m-1)r$ disjoint closed sets, each of which it meets; thus $b_0(\bigcap B_q) \geq (m-1)r \geq N$.

There exist (cf. 1.3(3) and [6, §6.1]) connected open sets $C_q \supset B_q$ whose closures A_q have the same incidences as the sets B_q; then

$$\operatorname{Fr}(A_j) \cap \operatorname{Fr}(A_k) \subset \operatorname{Fr}(C_j) \cap \operatorname{Fr}(C_k) \subset \operatorname{Co}(C_j \cup C_k) = 0$$

whenever $j \neq k$, and moreover we have $A_j \cup A_k = S$ $(1 \leq j < k \leq m)$ and $b_0(\bigcap A_j) \geq N$.

If $n = m$, the theorem is thus established. If $n > m$, we note that the open set $\operatorname{Int}[X_{m-1}(A_1, \cdots, A_m)] - X_m(A_1, \cdots, A_m)$ is nonempty, from 1.3(1), and take $A_{m+1}, \cdots, A_n$ to be $n - m$ distinct points in it; clearly

$$b_0[X_m(A_1, \cdots, A_m, \cdots, A_n)] \geq N + n - m,$$

and the proof is complete.

The modifications required to produce *open* sets A_j with similar properties are obvious.

3. Index inequalities for arbitrary sets.

3.1. AN INEQUALITY. Let $E_1, E_2, \cdots, E_n$ be arbitrary subsets of S. As in [4, §7], we write

$$A_j = \mathrm{Cl}(E_j),\ B_j = \mathrm{Cl}(S - E_j),\ P_j = X_j(A_1, \cdots, A_n),\ Q_j = X_j(B_1, \cdots, B_n).$$

An argument entirely analogous to that in [4, §7], based on 2.3 (1) and (2), gives:

THEOREM 3. *We have*

$$\begin{aligned} h[\mathrm{Fr}(E_1), \cdots, \mathrm{Fr}(E_n)] - nr(S) &\leq h(A_1, \cdots, A_n) + h(B_1, \cdots, B_n) \\ &\quad + h(P_1 \cap Q_n, P_2 \cap Q_{n-1}, \cdots, P_n \cap Q_1) \\ &\leq h[\mathrm{Fr}(E_1), \cdots, \mathrm{Fr}(E_n)] + nr(S). \end{aligned}$$

COROLLARY. We have

$$h(\bar{E}_1, \bar{E}_2, \cdots, \bar{E}_n) \leq h[\mathrm{Fr}(E_1), \mathrm{Fr}(E_2), \cdots, \mathrm{Fr}(E_n)] + nr(S).$$

3.2. THE CASE $m = 2$. It is easy to see that the inequalities in Theorem 3 are "best possible"; however, Theorem 1 suggests that in the Corollary the term $nr(S)$ could be replaced by $(n-1)\,r(S)$, or more generally by $(m-1)\,r(S)$, where no $m+1$ of the sets $\mathrm{Cl}(E_j)$ have a common point. I have been able to prove this only in the case $m = 2$:

THEOREM 4. *If $E_1, E_2, \cdots, E_n$ are arbitrary subsets of S, no three of whose closures have a common point, then*

$$h(\bar{E}_1, \bar{E}_2, \cdots, \bar{E}_n) \leq h[\mathrm{Fr}(E_1), \mathrm{Fr}(E_2), \cdots, \mathrm{Fr}(E_n)] + r(S).$$

Proof. We can assume that $r(S)$ is finite, and that the systems of sets $[\mathrm{Cl}(E_1), \cdots, \mathrm{Cl}(E_n)]$ and $[\mathrm{Fr}(E_1), \cdots, \mathrm{Fr}(E_n)]$ are both of finite incidence, since otherwise (in view of the convention regarding infinite terms in the h-function; see 1.2) Theorem 4 asserts no more than Theorem 3, Corollary. Hence, in view of 1.3(4), Theorem 4 will follow [if we take $A_j = \mathrm{Cl}(E_j)$ and $F_j = \mathrm{Fr}(E_j)$] from:

THEOREM 4a. *Let $A_1, A_2, \cdots, A_n, F_1, F_2, \cdots, F_n$ be any closed sets such that $A_j \supset F_j$ and $\bigcup F_j \supset \bigcup \mathrm{Fr}(A_j)$. Then*

$$p(A_1, A_2, \cdots, A_n) \leq p(F_1, F_2, \cdots, F_n) + r(S).$$

3.3. PROOF OF THEOREM 4a. Let $f_1, f_2, \cdots, f_N$ be N independent mappings of $\bigcup A_j$ in the unit circle such that each $f_k \sim 1$ on each A_j; we must prove that $N \leq p(F_1, \cdots, F_n) + r(S)$. Let s be the greatest number of mappings f_t which are independent on $\bigcup F_j$: clearly $s \leq p(F_1, \cdots, F_n)$. We may suppose that the mappings f_k are independent on $\cup F_j$ for $N-s+1 \leq k \leq N$, and then have, for each $t \leq N-s$, a relation (say)

$$g_t \equiv f_t^{p_t} \prod_k f_k^{q_{kt}} \sim 1$$

on $\cup F_j$, where $N-s+1 \leq k \leq N$. Thus g_t is a mapping of $\cup A_j$ in S^1 which ~ 1 on each A_j; and, since clearly $p_t \neq 0$, the mappings g_t $(1 \leq t \leq N-s)$ are independent on $\cup A_j$.

Write $C_0 = \mathrm{Cl}(S - \cup A_j)$; then $\mathrm{Fr}(C_0) \subset \cup F_j$, so that, from 1.3(7), each g_t may be extended to a mapping (still denoted by g_t) of S in S^1 such that $g_t \sim 1$ on C_0. Now define $C_1 = A_1$, $C_j = \mathrm{Cl}[A_j - (A_1 \cup A_2 \cup \cdots \cup A_{j-1})]$ $(2 \leq j \leq n)$; then the sets $C_0, C_1, \cdots, C_n$ are closed and cover S, and each $g_t \sim 1$ on each C_j. Let $Z = \cup(C_j \cap C_k)$, where $0 \leq j < k \leq n$; then $Z \subset \cup \mathrm{Fr}(A_j) \subset \cup F_j$, so that each $g_t \sim 1$ on Z. From 1.3(9), the number $N-s$ of mappings g_t is at most $r(S)$, and the theorem follows.

3.4. REMARK. We remark that no inequality similar to Theorem 4, but in the reverse direction, can hold in general. For example, take S to be the plane, and let A be a circular disc and B an inscribed convex polygon plus its interior; then A, B are closed and connected, and $h(A, B) = 0$, but $h[\mathrm{Fr}(A), \mathrm{Fr}(B)]$ can be arbitrarily large.

4. Frontiers of complementary components.

4.1. DEFINITION. For any $A \subset S$, let $\{C_\lambda\}$ be the components of the *complement* of A, and write

(1) $c(A) = \sum b_0(\mathrm{Fr}(C_\lambda))$,

with the usual convention that a vacuous sum is zero. [Thus $c(S) = 0$, $c(0) = -1$.] From [5, Th. 4] we have

(2) $c(A) + b_0[\mathrm{Cl}(S-A)] \geq b_0[\mathrm{Fr}(A)]$,

and (a weaker statement unless $b_0[\mathrm{Cl}(S-A)]$ is infinite)

(3) $c(A) \geq b_0(\overline{A})$.

If A is open, we evidently have equality in (2). (Note that (3) contains the well-known fact that, if $\overline{A}$ is not connected, at least one component of $\mathrm{Co}(A)$ has a disconnected frontier.)

4.2 LEMMA. *Let C be a component of $S - A$, and let U be an open set containing* $\mathrm{Fr}(C)$. *Then there exists an open set $V \supset \overline{A}$ such that $\overline{V} \cap \overline{C} \subset U$.*

This follows from [6, §6.1] applied to the sets $\bar{A}$, $\bar{C}$; a direct proof is also easy.

4.3 THEOREM 5. *If $c(A) \geq n$, then there exists an open set $A^* \supset \bar{A}$ such that, for each set B satisfying $A \subset B \subset A^*$, we have $c(B) \geq n$.*

For if $c(A) \geq n$, then there exist finitely many components, say C_1, C_2, $\cdots$, C_m, of $\mathrm{Co}(A)$, such that $b_0[\mathrm{Fr}(C_j)] \geq n_j$ where $\sum n_j \geq n$ $(1 \leq j \leq m)$. Thus, for each j, $\mathrm{Fr}(C_j)$ is a union of $n_j + 1$ disjoint closed nonempty sets F_{jk} $(1 \leq k \leq n_j + 1)$, and there exist open sets $U_{jk} \supset F_{jk}$ such that $\mathrm{Cl}(U_{jk}) \cap \mathrm{Cl}(U_{jl}) = 0$ $(j \neq l)$. Let $U_j = \bigcup_k U_{jk}$, an open set containing $\mathrm{Fr}(C_j)$; from the lemma in 4.2, there exists an open set $V_j \supset \bar{A}$ such that $\mathrm{Cl}(V_j) \cap \mathrm{Cl}(C_j) \subset U_j$. Take $A^* = \bigcap_j V_j$, and suppose that B is any set satisfying $A \subset B \subset A^*$. Then, since $\bigcup_k F_{jk} \subset \bar{B} \cap \bar{C}_j \subset \bigcup_k U_{jk}$, we have $b_0(\bar{B} \cap \bar{C}_j) \geq n_j$. Now let $\{D_{j\mu}\}$ be those components of $\mathrm{Co}(B)$ which are contained in C_j, and write $E_j = \bigcup_\mu D_{j\mu}$. One readily verifies that $\mathrm{Fr}(E_j) \subset \bar{B} \cap \bar{C}_j \subset \mathrm{Cl}(S - E_j)$, and that $E_j \cup (\bar{B} \cap \bar{C}_j) = \bar{C}_j$; hence, from [5, Th. 4], $\sum_\mu b_0(\mathrm{Fr}(D_{j\mu})) \geq b_0(\bar{B} \cap \bar{C}_j) \geq n_j$, so that $c(B) \geq \sum_{j,\mu} b_0(\mathrm{Fr}(D_{j\mu})) \geq \sum n_j \geq n$.

COROLLARY. *We have $c(A) \leq c(\bar{A})$.*

4.4. EXTENSION OF THE PHRAGMÉN-BROUWER THEOREM. This theorem, as extended in [5, Th. 5], can now be extended still further.

THEOREM 6. *For any set A, we have $c(A) \leq b_0(\bar{A}) + r(S)$.*

The proof is almost identical with that for the case in which $\bar{A}$ is connected, in [5, §4.2]; the difference arises from the fact that the sets L, M there constructed need not here be connected. But we may assume without loss that $b_0(\bar{A}) < \infty$, and have $b_0(L) \leq b_0(\bar{A})$ and $b_0(M) \leq b_0(\bar{A})$; hence, from 2.3 (1), we have $b_0(L \cap M) < 2b_0(\bar{A}) + r(S)$. Since $b_0(\bar{A}) + 1$ of the components of $L \cap M$ now arise from $\bar{A}$, the argument can be concluded in the same way as before.

COROLLARY 1. *If $r(S)$ is finite, and A is any subset of S such that $\bar{A}$ has only a finite number of components, then all but at most a finite number of the components of $S - A$ have connected frontiers.*

COROLLARY 2. *If S is unicoherent, then $c(A) = b_0(\bar{A})$; and, conversely, this equality is characteristic of unicoherence.*

(This follows from 4.1 (2) and [5, Th. 5].)

4.5. ANOTHER EXTENSION. It has been shown in [5, Th. 5] that, conversely, Theorem 6 serves to characterize $r(S)$, even when restricted to the case in which A is closed (or open) and connected. However, Theorem 6 can be restated in a slightly different though equally natural way, in which the converse question is more difficult.

THEOREM 6a. *For any set A, we have*

$$\text{(i)}\quad b_0(\operatorname{Fr}(A)) \leq c(A) + b_0(\operatorname{Cl}(S-A)) \leq c(\bar{A}) + b_0(\operatorname{Cl}(S-A)) \leq b_0(\operatorname{Fr}(A)) + r(S).$$

Conversely, if for some fixed (finite) n we have

$$\text{(ii)}\quad c(A) \leq b_0(\operatorname{Fr}(A)) + n$$

whenever A is nowhere dense, and if further

(iii) *S is metrizable,* or *$r(S)$ is finite,*

then $r(S) \leq n$.

The first inequality in (i) is a restatement of 4.1 (2), the second follows from Theorem 5, Corollary, and the third from Theorem 6 applied to $\bar{A}$, in view of the fact [4, §6.2] that $b_0(\bar{A}) + b_0(\operatorname{Cl}(S-A)) \leq b_0(\operatorname{Fr}(A))$. For the converse, suppose that (ii) holds, but that $r(S) > n$. From [5, Th. 5a), there exists a closed connected set A' such that $S - A'$ has only a finite number of (open) components $C_1, C_2, \cdots, C_m$, and $b_0(\operatorname{Fr}(A')) > m + n - 1$; thus from [5, Th. 4], we have $\sum b_0(\operatorname{Fr}(C_j)) > n$. Suppose now that $r(S)$ is finite, and write $A = \operatorname{Fr}(A')$; thus A is nowhere dense, and, from 2.3 (2), $b_0(A) < \infty$. Let $\{D_\lambda\}$ be the components of $\operatorname{Int}(A')$; then [5, Th. 4] we have $\sum b_0(\operatorname{Fr}(D_\lambda)] \geq b_0(A) = b_0[\operatorname{Fr}(A)]$. But the components of $\operatorname{Co}(A)$ are precisely the sets C_j, D_λ; hence $c(A) > b_0[\operatorname{Fr}(A)] + n$, contradicting (ii).

If $r(S) = \infty$, the above argument still applies provided that $b_0[\operatorname{Fr}(A')] < \infty$. Hence we may assume $b_0[\operatorname{Fr}(A')] = \infty$, so that there must exist some C_j, say C, for which $b_0[\operatorname{Fr}(C)] = \infty$. Now, the complement (say) F of C is closed and connected. If it is assumed that S is metrizable, then there exists a sequence of open sets G_n such that $G_n \supset \operatorname{Cl}(G_{n+1})$ $(n = 1, 2, \cdots)$, and $\bigcap G_n = F$. Let $X = C - \bigcup \operatorname{Fr}(G_n)$; from a theorem of Hewitt [3], there exist disjoint sets Y, Z such that $Y \cup Z = X$ and $\bar{Y} = \bar{Z} = \bar{X} = \bar{C}$. We take $A = C - Y$. Thus $\operatorname{Cl}(S - A) = S$; and $\operatorname{Fr}(A) = \bar{C}$, which is connected. But $\operatorname{Co}(A)$ can be separated, by one of the sets $\operatorname{Fr}(G_n)$, between F and any given point of Y; thus one of the components of $\operatorname{Co}(A)$ is F itself, and again (ii) is contradicted.

COROLLARY. *If S is unicoherent, and $\{C_\lambda\}$ are the components of an* *arbitrary set E, then*

$$b_0[\mathrm{Fr}(C_\lambda)] + b_0(\bar{E}) = b_0[\mathrm{Fr}(E)];$$

and this property characterizes unicoherence among metrizable (locally connected and connected) spaces.

It would be interesting to know whether the extra hypotheses on S imposed in (iii) are needed. It would be easy to replace them by others (for example, local compactness plus perfect normality).

5. Modified addition theorems.

5.1. A MODIFICATION. As in the case of two connected sets [5, Ths. 11 and 11a], special cases of Theorem 1 can be obtained under alternative hypotheses. As an example, we state:

THEOREM 7. *If A and B are any sets satisfying*

$$\mathrm{Fr}(A) \cap \mathrm{Fr}(B) \cap \mathrm{Fr}(A \cap B) = 0,$$

then

$$b_0(A \cup B) + b_0(A \cap B) \leq b_0(A) + b_0(B) + r(S);$$

and if there is finite equality here, then $A - B$ and $B - A$ are separated (so that Theorem 1 then in fact applies).[9]

The proof is a fairly straightforward generalization of that of [5, Th. 11], with 2.3(1) replacing [5, §7.4]. The extension of Theorem 7 to n sets, however, appears to present some difficulty.

5.2. ANOTHER MODIFICATION. A more interesting modification of Theorem 1 is the following, in which $r(S)$ does not enter explicitly; in some cases (in view of Theorem 6) it gives more information than does Theorem 1.

THEOREM 8. *If A and B are arbitrary sets such that*

$$\mathrm{Fr}(A) \cap \mathrm{Fr}(B) \cap \mathrm{Fr}(A \cup B) = 0,$$

then

$$h(\bar{A}, \bar{B}) + b_0(\bar{A}) \leq c(A).$$

---

[9] It follows (see 5.6 below) that, in the case of finite equality, we have for each component E of $A \cup B$ that $\mathrm{Fr}(A \cap E) \cap \mathrm{Fr}(B \cap E) = 0$. It is false, in general, that $\mathrm{Fr}(A) \cap \mathrm{Fr}(B) = 0$.

Proof. Write $C = \mathrm{Cl}[\mathrm{Co}(A)]$, and apply [4, Th. 6b] to the closed sets $\bar{A} \cup \bar{B}$, $\bar{A} \cap \bar{B}$, C. We obtain

(1) $b_0(\bar{A} \cup \bar{B}) + b_0(\bar{A} \cap \bar{B}) + b_0(C) \leq b_0[\bar{B} \cup \mathrm{Fr}(A)] + b_0[\bar{B} \cap \mathrm{Fr}(A)]$.

From the frontier relation satisfied by the sets A and B, it readily follows that $\mathrm{Fr}(A) \cap \mathrm{Co}(\bar{B})$ is closed, and thence that each component of $\mathrm{Fr}(A)$ which meets $\bar{B}$ is contained in $\bar{B}$. Hence we see that

$$b_0[\bar{B} \cup \mathrm{Fr}(A)] + b_0[\bar{B} \cap \mathrm{Fr}(A)] = b_0(\bar{B}) + b_0[\mathrm{Fr}(A)],$$

and consequently

(2) $b_0(\bar{A} \cup \bar{B}) + b_0(\bar{A} \cap \bar{B}) + b_0(C) \leq b_0(\bar{B}) + b_0[\mathrm{Fr}(A)]$.

But by 4.1(2), we have $b_0[\mathrm{Fr}(A)] \leq b_0(C) + c(A)$. Thus, *provided that $b_0(C)$ is finite,* we have proved

(3) $b_0(\bar{A} \cup \bar{B}) + b_0(\bar{A} \cap \bar{B}) \leq b_0(\bar{B}) + c(A)$,

from which the theorem follows immediately.

To complete the proof, we deduce that (3) continues to hold even when $b_0(C) = \infty$; and in doing so, we may assume that $b_0(\bar{B}) + c(A) < \infty$. Define B^* to be the union of those components of $\bar{B}$ which meet $\bar{A}$, and A^* to be the union of A with all components of $\mathrm{Co}(A)$ which have connected frontiers. It is easy to verify that

$$\mathrm{Fr}(A^*) \cap \mathrm{Fr}(B^*) \cap \mathrm{Fr}(A^* \cap B^*) = 0,$$

and that, since $c(A) < \infty$, $b_0[\mathrm{Co}(A^*)]$ is finite. Hence (3) holds for the sets A^*, B^*; and it is a routine matter to deduce that (3) also holds for A and B^*, and thence finally for A and B.

There is no difficulty in extending this theorem to any number of sets; for example, (2) can be extended to the following property, valid in an arbitrary topological space S (and generalizing [4, §7.4(1)]):

(4) If $A_1, \cdots, A_m, B_1, \cdots, B_n$ are arbitrary sets such that

$$\mathrm{Fr}(A_j) \cap \mathrm{Fr}(B_k) \cap \mathrm{Fr}(A_j \cup B_k) = 0 \qquad (1 \leq j \leq m,\ 1 \leq k \leq n),$$

and $C_j = \mathrm{Cl}[\mathrm{Co}(A_j)]$, then

$$\sum b_0[X_h(\overline{A}_1, \cdots, \overline{A}_m, \overline{B}_1, \cdots, \overline{B}_n)] + \sum b_0[X_j(C_1, \cdots, C_m)]$$
$$\leq \sum b_0[X_j[\mathrm{Fr}(A_1), \cdots, \mathrm{Fr}(A_m)]\}$$
$$+ \sum b_0[X_k(\overline{B}_1, \cdots, \overline{B}_n)] + mb_0(S),$$

the ranges of summation being $1 \leq h \leq m + n$, $1 \leq j \leq m$, $1 \leq k \leq n$; and (3) can be extended similarly.

5.3. AN INCLUSIVE RESULT. The next theorem includes both Theorem 1 and the extended Phragmén-Brouwer theorem (Theorem 6) as special cases. We shall need the following lemma.

LEMMA. *If G is a set with only finitely many components, then there exists a finite set of points $x_1, x_2, \cdots, x_q \in \mathrm{Fr}(G)$ such that*

$$b_0[G \cup (x_1) \cup \cdots \cup (x_q)] = b_0(\overline{G}).$$

For if G has components $G_1, G_2, \cdots, G_s$, we have only to take at least one point x_j in every nonempty set $\overline{G}_\lambda \cap \overline{G}_\mu$ $(\lambda \neq \mu)$.

5.4. THEOREM 9. *Let $A_1, A_2, \cdots, A_n$ be any subsets of S having separated differences and such that $A_j \cap A_k$ and $\mathrm{Co}(A_j \cup A_k)$ are separated whenever $j \neq k$; and suppose that no point belongs to A_j for more than m values of j, where $2 \leq m \leq n$. Then*

(1) $h(A_1, \cdots, A_n) + c(\overline{X}_1) + c(\overline{X}_2) + \cdots + c(\overline{X}_{m-1})$

$$\leq b_0(\overline{X}_1) + b_0(\overline{X}_2) + \cdots + b_0(\overline{X}_{m-1}) + (m-1)\,r(S),$$

where $X_j = X_j(A_1, \cdots, A_n)$. Further, if there is finite equality in (1), *then, for each $q \leq n - 1$, for each set J of $q + 1$ distinct suffixes $j_1, j_2, \cdots, j_{q+1}$ between 1 and n, and for each component E of X_q, we have*

(2) $\bigcap\{\mathrm{Fr}(A_j \cap E) \mid j \in J\} \subset E.$

To prove (1), we may assume throughout that $r(S)$ and $\sum b_0(A_j)$ are finite; it then follows from Theorems 1 and 6 that the numbers $b_0(X_j)$, $b_0(\overline{X}_j)$, and $c(X_j)$ are also finite. Further, we may obviously suppose that $X_{m-1} \neq 0$ (otherwise (1) would be derived with a smaller value of m). Again, by using the method of approximation, we may assume in addition that the sets A_j are all open and, by 1.3(2), satisfy

(3) $\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cap A_k) = 0 \qquad (j \neq k).$

For, in the general case, we apply 1.3 (3) to replace the sets A_j by slightly larger relatively connected sets A_j^* having the same incidences and satisfying (3); and, in view of Theorem 5, the truth of (1) for the sets A_j^* will imply (1) for the sets A_j.

From (3) and 1.3 (1), the open sets X_j satisfy

(4) $X_1 \supset \overline{X}_2 \supset X_2 \supset \overline{X}_3 \supset \cdots \supset X_{m+1} = 0.$

We shall define inductively, for $j = 1, 2, \cdots, m-1$, open sets G_j consisting of a finite number of components C_{jk} of $\mathrm{Co}(\overline{X}_j)$, and open sets $V_j \supset \mathrm{Fr}(G_j)$, such that[11]

(5) $G_j \cup V_j \subset G_k$ whenever $j < k$;

$\overline{V}_j \subset X_{j-1}$; $\overline{V}_j \cap \overline{X}_{j+1} = 0$; $\overline{V}_j \cap \overline{V}_k = 0$ if $j \neq k$;

$\mathrm{Fr}(V_j) \cap \mathrm{Fr}(A_k) = 0$ (for all j, k); and $\mathrm{Fr}(V_j) \cap \mathrm{Fr}(X_j) = 0.$

Further,

(6) $b_0(V_j) < \infty$, $b_0(X_j \cup V_j) \leq b_0(\overline{X}_j)$, and

$$b_0(V_j \cap G_j) \geq c(\overline{X}_j) + b_0(G_j).$$

For suppose this done for all $j < p$, where $1 < p < m$. Define G_p to be the union of all those components of $\mathrm{Co}(\overline{X}_p)$ which either (a) have disconnected frontiers, or (b) meet $G_{p-1} \cup V_{p-1}$. Since $G_{p-1} \cup V_{p-1} \subset \mathrm{Co}(\overline{X}_p)$, this gives $G_{p-1} \cup V_{p-1} \subset G_p$; and since further

$$b_0(G_{p-1} \cup V_{p-1}) < \infty,$$

Theorem 6, Corollary 1, shows that $b_0(G_p) < \infty$. Let G_p consist of the components C_{pk} of $\mathrm{Co}(\overline{X}_p)$ $(k = 1, 2, \cdots, n_p)$; thus

$$\textstyle\sum_k b_0[\mathrm{Fr}(C_{pk})] = c(\overline{X}_p).$$

Hence, if the components of $\mathrm{Fr}(C_{pk})$ are denoted by F_{pkl} $(l = 1, 2, \cdots, n_{pk})$, we have $\sum_k (n_{pk} - 1) = c(\overline{X}_p)$. For fixed p and k, there exist open sets $N_{pkl} \supset F_{pkl}$ with disjoint closures (for varying l); and it follows from the lemma in 4.2 that an open set U_p exists such that $U_p \supset \mathrm{Fr}(X_p)$ and $U_p \cap \mathrm{Cl}(C_{pk}) \subset \bigcup_l N_{pkl}$. We may further suppose, from (4), that $\overline{U}_p \subset X_{p-1}$, $\overline{U}_p \cap \overline{X}_{p+1} = 0$, and $\overline{U}_p \cap \overline{V}_{p-1} = 0$. It readily follows that $\mathrm{Fr}(U_p) \cap \mathrm{Fr}(A_k) = 0$ for each k $(1 \leq k \leq n)$. Again, from

[11] By convention, $X_0 = S$ and $X_{n+1} = 0$.

the lemma in 5.3, there exists a finite set $Q_p \subset \mathrm{Fr}(X_p)$ such that

$$b_0(X_p \cup Q_p) = b_0(\overline{X}_p).$$

Let V_p = union of those components of U_p which meet $Q_p \cup \bigcup F_{pkl}$; clearly $b_0(V_p)$ is finite; and, since $\mathrm{Fr}(V_p) \subset \mathrm{Fr}(U_p)$, it is easily seen that (5) continues to hold when $j = p$. Also the sets $V_p \cap C_{pk} \cap N_{pkl}$ are (for varying k and l) all pairwise separated and nonempty; hence the number of components of $V_p \cap G_p$ is at least $\sum_k n_{pk} = c(\overline{X}_p) + n_p$, so that (6) holds.

To start the induction, we take G_1 to consist of the components of $\mathrm{Co}(\overline{X}_1)$ with disconnected frontiers; the rest of the construction is exactly as in the general case. Thus (5) and (6) hold for $j = 1, 2, \cdots, m-1$. We remark that it follows trivially from (5) that

(7) $\mathrm{Fr}(V_j) \cap \mathrm{Fr}(G_k) = 0$ whenever $j \neq k$, and $V_j \cap G_k = 0$ if $j > k$.

Now consider the "elementary symmetric sets"

$$Y_j = X_j(G_1, G_2, \cdots, G_{m-1}, V_1, V_2, \cdots, V_{m-1}) \qquad (1 \leq j \leq 2m-2)$$

and

$$Z_k = X_k(A_1, A_2, \cdots, A_n, G_1, \cdots, G_{m-1}, V_1, \cdots, V_{m-1})$$

$$(1 \leq k \leq 2m+n-2).$$

Using (5) and (7), we obtain

(8) $Y_1 = G_{m-1} \cup V_{m-1}$, $Y_j = G_{m-j} \cup V_{m-j} \cup (G_{m-j+1} \cap V_{m-j+1})$ if

$2 \leq j \leq m-1$; $Y_m = G_1 \cap V_1$; and $Y_j = 0$ if $j > m$.

Thus, since $Z_k = X_k \cup \bigcup(X_{k-p} \cap Y_p) \cup Y_k$, we find:

(9) $Z_k \neq 0$ if $1 \leq k \leq m$; $Z_m = X_m \cup \bigcup(X_p \cap V_p) \cup \bigcup(G_q \cap V_q)$

$(p, q = 1, 2, \cdots, m-1)$; and $Z_k = 0$ if $k > m$.

Now the open sets $A_1, A_2, \cdots, A_n$; $G_1, \cdots, G_{m-1}$; $V_1, \cdots, V_{m-1}$ satisfy the hypotheses of Theorem 1, since this is true of $A_1, \cdots, A_n$ from (3), while we readily verify that

$$\mathrm{Fr}(A_j) \cap \mathrm{Fr}(G_k) \cap \mathrm{Fr}(A_j \cap G_k) = 0,$$

and that in all the remaining cases the sets have disjoint frontiers. Hence

$$(10)\quad \sum b_0(Z_p) \leq \sum b_0(A_j) + \sum b_0(G_k) + \sum b_0(V_k) + (m-1)\,r(S)$$

$$(1 \leq p \leq n + 2m - 2,\ \ 1 \leq j \leq n,\ 1 \leq k \leq m-1).$$

But (9) shows that

$$\sum b_0(Z_p) \geq b_0(X_m) + \sum b_0(X_k \cap V_k) + \sum b_0(V_k \cap G_k) + m - n.$$

Also

$$b_0(X_k \cap V_k) \geq b_0(X_k) + b_0(V_k) - b_0(X_k \cap V_k) \qquad \text{(cf. [4, §6.2])}$$

$$\geq b_0(X_k) + b_0(V_k) - b_0(\overline{X}_k),$$

and, from (6).

$$b_0(V_k \cap G_k) \geq c(\overline{X}_k) + b_0(G).$$

Thus finally, since all the numbers involved here are finite, (1) follows from (10).

5.5. THE CASE OF FINITE EQUALITY. Suppose now that there is finite equality in (1) above, and that a point y exists in (say)

$$\mathrm{Fr}(A_1 \cap E) \cap \mathrm{Fr}(A_2 \cap E) \cap \cdots \cap \mathrm{Fr}(A_{p+1} \cap E) - E,$$

where E is a component of X_p; thus $y \not\subset X_p$. It is easy to see that we may assume without loss of generality that $p \leq m-1$ and that the sets A_j are all open. Clearly $y \in \mathrm{Fr}(X_p)$; thus we may carry out the preceding construction in such a way that $y \in Q_p \subset V_p$. But, from the way in which (1) was derived from (10), we must now have $h(X_p, V_p) = 0$, so that the component W of V_p which contains y must meet E in a connected set; consequently, since $W \cap \mathrm{Cl}(X_{p+1}) = 0$, it follows that $W \cap E$ meets one and only one of the sets

$$A_J\ (= \cap \{A_j,\ j \in J\})$$

with $|J| = p$. Since W meets $A_1 \cap E$, we have $1 \in J$; similarly $2 \in J, \cdots$, and $(p+1) \in J$, giving a contradiction.

5.6. REMARKS. We observe that the preceding results contain those concerning modified addition theorems in [5, Ths. 11 and 11a]. For, in the first place, 1.3 (1) together with an "approximation" argument shows that the relation

(2) above is equivalent to the apparently stronger relation

(2a) $$\bigcap \{\mathrm{Fr}(A_j \cap E)\} \subseteq \mathrm{Cl}(X_{p+1}) \qquad (j \in J,\ |J| = p+1,\ p < n).$$

(In fact, the left side here is contained in

$$\mathrm{Fr}(X_{p+1}) \cup \mathrm{Fr}(X_{p+2}) \cup \cdots \cup \mathrm{Fr}(X_m).\,)$$

Hence if $A_1, \cdots, A_n$ also satisfy the condition (slightly stronger than in Theorem 9) that $\mathrm{Fr}(A_j \cap A_k) \cap \mathrm{Fr}(A_j \cup A_k) = 0$ $(j \neq k)$, finite equality in Theorem 9 will imply, again from 1.3(1), that

(2b) $$\bigcap\{\mathrm{Fr}(A_j \cap E)\} \subset \mathrm{Int}\,(E) \qquad (j \in J, |J| = p+1,\ p < n),$$

a relation which is slightly stronger than (2). And if the sets A_j satisfy the even stronger condition

$$\mathrm{Fr}(A_j) \cap \mathrm{Fr}(A_k) \cap \mathrm{Fr}(A_j \cap A_k) = 0 \qquad (j \neq k),$$

it can be deduced from (2b) that

(2c) $$\bigcap\{\mathrm{Fr}(A_j \cap E)\} = 0 \ \text{ if } \ 2p \geq n \qquad (j \in J,\ |J| = p+1,\ p < n).$$

Finally, if there is finite equality in Theorem 1, then there will be finite equality in Theorem 9, for $c(\overline{X}_j) \geq b_0(\overline{X}_j)$, by 4.1(3); and thus the above considerations will apply.

5.7. Other Inequalities. Many other inequalities can be derived from Theorem 9; for example:

Theorem 9a. *Under the hypotheses of Theorem 9, we have*

$$\text{(i)}\ \ c(\overline{X}_1) + c(\overline{X}_2) + \cdots + c(\overline{X}_{m-1}) + b_0(X_m) + b_0(X_{m+1}) + \cdots + b_0(X_n)$$
$$\leq \sum b_0(A_j) + (m-1)\,r(S).$$

Further, if there is finite equality in (i), *we have*

$$\text{(ii)}\ \ X_p(\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_n) = \overline{X}_p \qquad (1 \leq p \leq m).\text{[12]}$$

Proof. Relation (i) is a trivial consequence of Theorem 9, (1), since

$$b_0(\overline{X}_j) \leq b_0(X_j)\,.$$

[12] Condition (ii) need not hold for $p > m$.

Suppose there is finite equality in (i); as before it will suffice, by an approximation argument, to prove (ii) assuming that the sets A_j are open. Now, finite equality in (i) implies that Theorem 9, (2), holds, and also that $b_0(\overline{X}_j) = b_0(X_j)$ for all $j < m$. If J, K denote sets of p distinct suffixes between 1 and n, and $J \neq K$, we find (writing $A_J = \bigcap\{A_j \mid j \in J\}$) that

$$\overline{A}_J \cap \overline{A}_K \subset \overline{X}_{p+1} \qquad \text{if } p < m.$$

This includes (ii) when $p = 2$; and (ii) follows in general by an easy induction over p.

5.8. GEOMETRICAL CONSIDERATIONS. To illustrate the geometrical content of these theorems, we consider the case of two sets in more detail.

THEOREM 10. *Let A and B be sets, neither of which contains the other, having separated differences and connected complements, and suppose that $A \cap B$ and* $\mathrm{Co}(A \cup B)$ *are separated. Then*

$$b_0(A \cap B) + b_0(\mathrm{Co}(A \cup B)) \leq b_0(A) + b_0(B) + r(S) - 1.$$

If there is finite equality here, and further $\mathrm{Fr}(A) \cap \mathrm{Fr}(B) \cap \mathrm{Fr}(A \cup B) = 0$, *then each component of* $\mathrm{Co}(A \cup B)$ *has a frontier consisting of exactly two components.*

Proof. We can assume that $r(S)$ is finite. Write $P = \mathrm{Co}(A)$, $Q = \mathrm{Co}(B)$; then P and Q are connected, so that, from Theorem 1, $b_0(P \cap Q)$ is finite. Let $P \cap Q$ $(= \mathrm{Co}(A \cup B))$ have components $H_1, H_2, \cdots, H_n$. Then

(1) A and H_j are not separated,

since otherwise $Q = \{(Q \cap A) \cup (P \cap Q - H_j)\} \cup H_j$, a union of two nonempty separated sets. Similarly B and H_j are not separated. Hence

(2) $\mathrm{Fr}(H_j)$ meets both $\mathrm{Fr}(A)$ and $\mathrm{Fr}(B)$.

Let $A^* = A \cup (P \cap Q)$, $B^* = B \cup (P \cap Q)$; from (1), A^* is connected relative to A, so that $b_0(A^*) \leq b_0(A)$, and similarly $b_0(B^*) \leq b_0(B)$. It is easy to see that $A^* - B^*$ and $B^* - A^*$ are separated, and that $A^* \cap B^*$ and $\mathrm{Co}(A^* \cup B^*)$ are separated; hence 2.3 (1) gives

$$b_0(A^* \cap B^*) \leq b_0(A) + b_0(B) + r(S).$$

But $A^* \cap B^* = (A \cap B) \cup \mathrm{Co}(A \cup B)$, a union of two separated sets; thus the first part of the theorem follows.

From Theorem 9a and Theorem 5, Corollary, we also have

$$b_0(A \cap B) + \sum b_0(\mathrm{Fr}(H_j)) \leq b_0(A) + b_0(B) + r(S).$$

If further $\mathrm{Fr}(A) \cap \mathrm{Fr}(B) \cap \mathrm{Fr}(A \cup B) = 0$, (2) shows that $b_0[\mathrm{Fr}(H_j)] \geq 1$ for each j. Hence finite equality in Theorem 10 requires $b_0[\mathrm{Fr}(H_j)] = 1$ for each j, and the proof is complete.

COROLLARY. *If A and B are simple sets*[13] *with disjoint frontiers, and neither A nor B contains the other, then $b_0(A \cap B) + b_0[\mathrm{Co}(A \cup B)] \leq r(S) - 1$; and if there is finite equality here, then each component of $A \cap B$ or of $\mathrm{Co}(A \cup B)$ has a frontier with exactly two components.*

This follows on applying Theorem 10 first to A, B and then to $\mathrm{Co}(A)$, $\mathrm{Co}(B)$. If S is unicoherent, the first part of this corollary reduces to [4, §4.5].

6. Simple sets with disjoint frontiers.

6.1. FINITELY MULTICOHERENT SPACES. Throughout this section, we shall assume that *$r(S)$ is finite.*

THEOREM 11. *Let $A_1, A_2, \cdots, A_n$ be simple*[14] *subsets of S, every two of which meet, and which have disjoint frontiers. Then there exist N or fewer of the sets A_j whose union is $\bigcup_1^n A_j$, where*

$N = 2r(S)$ *if* $r(S) > 1$, *or if* $r(S) = 1$ *and* $\bigcap A_j \neq 0$.

$N = 3$ *if* $r(S) = 1$ *and* $\bigcap A_j = 0$, *and*

$N = 2$ *if* $r(S) = 0$.

These values of N are the smallest possible.

It is easy to see by examples (it suffices to take S to be a linear graph) that no smaller values of N are possible in general. To prove the rest of the theorem, we need two graph-theoretic lemmas.

6.2. LEMMA 1. *Let G be a connected linear graph having no end-points, and let $E_1, E_2, \cdots, E_n$ be closed connected subsets of G, every two of which meet. If $r(G) > 1$, or if $r(G) = 1$ and $\bigcap E_j \neq 0$, then $\bigcup E_j$ is the union of $2r(G)$ or fewer of the sets E_j; if $r(G) = 1$ and $\bigcap E_j = 0$, then $\bigcup E_j$ is the union of at most 3 sets E_j.*

---

[13] A set E is "simple" if E and $S - E$ are both connected.

[14] It would suffice to require only that $\mathrm{Cl}(A_j)$ and $\mathrm{Cl}[\mathrm{Co}(A_j)]$ be connected ($1 \leq j \leq n$).

The proof is by induction over $r(G)$. If $r(G) \leq 2$, the lemma can be verified by inspection of the possible graphs G. Suppose, then, that $r(G) \geq 3$, and that the lemma is true for all graphs of smaller degree of multicoherence but not for G, and that n is the smallest number of sets for which the lemma fails for G. Thus no E_j is contained in the union of the others.

From G we derive a homeomorphic graph G^* by suppressing all vertices of order 2; the (open) 1-cells of G^* will thus be the components of G minus its vertices of orders other than 2; we call them the "maximal 1-cells" of G. (A maximal 1-cell may have coincident end-points.) We consider three cases:

(1) If G has a cut-point R which is not a vertex of G, let PQ be the maximal 1-cell of G which contains R; thus here $P \neq Q$, and $G - PQ$ is a union of two disjoint, closed connected nonempty subgraphs H, K, neither of which has an end-point. From [6, §3.2(1)], we see that $r(H) + r(K) = r(G)$, while, since G has no end-points, $r(H) \geq 1$ and $r(K) \geq 1$. For the moment we assume that neither $r(H)$ nor $r(K)$ is 1. Write $E_j{}' = E_j \cap H$, $E_j{}'' = E_j \cap K$; it is easy to see that these sets are closed and connected, though possibly empty. Further, every two nonempty sets $E_j{}'$ must meet, since both must contain P unless one of the corresponding sets E_j is contained in H. Hence the hypothesis of induction applies to H and the nonempty sets $E_j{}'$, and $\bigcup E_j{}'$ must be contained in the union of at most $2r(H)$ sets E_j. Similarly $\bigcup E_j{}''$ is contained in the union of at most $2r(K)$ sets E_j. Thus we obtain at most $2r(G)$ sets E_j in all, which together contain $\bigcup E_j{}' \cup \bigcup E_j{}''$; further, their union is connected and so contains PQ and thus $\bigcup E_j$, unless $\bigcup E_j{}'$ or $\bigcup E_j{}''$ is empty.

If $\bigcup E_j{}''$, say, is empty but $\bigcup E_j{}' \neq 0$, it is easy to see that at most $2r(H) + 1 < 2r(G)$ sets E_j will suffice, namely those selected to contain $\bigcup E_j{}'$, together with the set E_j which contains the largest subarc of PQ. If $\bigcup E_j{}' = \bigcup E_j{}'' = 0$, all the sets E_j are contained in PQ, and two of them will suffice.

If $r(H)$, say, is 1 (so that H is a circle), the above argument needs modification only if one of the given sets is contained in $H - (P)$; we leave the details to the reader.

(2) If G has a cut-vertex R, but no cut-point other than a vertex, the argument is essentially the same as before, with PQ degenerating to R.

(3) Finally, if G has no cut-points, pick $x \in E_1 - (E_2 \cup \cdots \cup E_n)$; replacing x by a sufficiently nearby point if necessary, we can suppose that x is not a vertex and so belongs to a unique maximal 1-cell PQ of G. Here $P \neq Q$, since G has no cut-points, and the subgraph $H = G - PQ$ is connected and has no end-lines. We easily find $r(H) = r(G) - 1$. Write $E_j{}' = E_j \cap H$; as before, at most $2r(H)$ of the sets $E_2, \cdots, E_n$, say $E_2, \cdots, E_m$ $(m \leq 2r(H) + 1)$, must contain

$\bigcup E_j{}'$ $(j \geq 2)$. The connected set $E_1 \cup E_2$ joins x to H (for we may clearly assume $\bigcup E_j{}' \neq 0$), and so contains one of the arcs Px, Qx, say Px. If none of $E_{m+1}, \cdots, E_n$ meets Qx, the m sets $E_1, E_2, \cdots, E_m$ contain $\bigcup E_j$. If $Qx \cap (E_{m+1} \cup \cdots \cup E_n) \neq 0$, let y be its point on Qx closest to x, and let $y \in E_k$; then the connected set $E_2 \cup E_k$ joins y to H without containing x, and so contains QX; thus the $m+1$ sets $E_1, E_2, \cdots, E_m, E_k$ contain $\bigcup E_j$. Since $m + 1 \leq 2r(G)$, the proof is complete.

6.3. LEMMA 2. *Let $B_1, B_2, \cdots, B_n$ be n simple closed subsets of a connected linear graph G, every two of which meet. If $r(G) > 1$, or if $r(G) = 1$ and $\bigcap B_j \neq 0$, then $\bigcup B_j$ is the union of $2r(G)$ or fewer of the sets B_j; if $r(G) = 1$ and $\bigcap B_j = 0$, then $\bigcup B_j$ is the union of at most 3 sets B_j.*

As before, we may assume that the lemma is false, and that n is the smallest number of sets for which it fails; thus no B_j is contained in the union of the others. Define a "maximal end-line" PQ of G to be a maximal 1-cell PQ of G in which Q is an end-point of G; thus $P \neq Q$. If B_1, say, meets a maximal end-line PQ which it does not contain, then (being closed and simple) B_1 must be either a closed arc xQ, where $x \in PQ$, or the closure of the complement in G of such an arc. In the latter case, it is clear that B_1 together with one other set B_j will contain the rest; in the former case, we see similarly that either $B_1 \cup B_2 = G$, or $B_2 \supset B_1$, or $B_1 \supset B_2$—all of which are excluded. This proves, then, that each B_j contains all maximal end-lines of G which it meets. Let H be the graph obtained from G by removing all end-points and maximal end-lines, and write $E_i = B_i \cap H$. On applying Lemma 1 to the sets $E_1, \cdots, E_n$ in the graph H, we see that $\bigcup E_j$ is the union of the desired number of sets E_j; the analogous conclusion for the sets B_j follows.

6.4. *Proof of Theorem* 11. We shall consider only the case $r(S) > 1$ explicitly; the modifications needed when $r(S) = 1$ will be obvious, and the case $r(S) = 0$ is covered by [4, §4.5]. It will thus suffice to prove that, if $n > 2r(S) \geq 4$, one of the sets A_j is contained in the union of the others. Consider the 2^n intersections $Y_k = \overline{D}_1 \cap \overline{D}_2 \cap \cdots \cap \overline{D}_n$ $(1 \leq k \leq 2^n)$, where each D_j takes the two values A_j, $\mathrm{Co}(A_j)$, in all possible combinations. The sets Y_k are closed and cover S; and, since the sets $\mathrm{Fr}(A_j)$ are disjoint, no three of them have a common point. Further, from Theorem 1, $b_0(Y_k)$ is finite, and so the sets Y_k are of finite incidence. Let G denote the modified nerve (cf. [6]) of the sets Y_k; as in [6, §6.4], G is connected and $r(G) \leq r(S)$. Let B_p denote the subgraph of G consisting of (i) all vertices which correspond to intersections Y_k in which the pth "factor" D_p is A_p, and (ii) all edges of G both of whose end-points have been assigned to B_p. Let C_p be defined similarly, but with $\mathrm{Co}(A_p)$ replacing A_p. Thus, for each p $(1 \leq p \leq n)$, B_p and C_p are disjoint subgraphs which together contain all the vertices of G; and it is easy to see that B_p and C_p are connected, since A_p and $\mathrm{Cl}[\mathrm{Co}(A_p)]$ are. Hence $B_1, B_2, \cdots, B_n$ are simple closed subsets of G. Further, if $p \neq q$, B_p and B_q have at least a common vertex. Thus, by Lemma 2, one of the sets B_p is contained in the union of the others; say $B_1 \subset B_2 \cup \cdots \cup B_n$. It readily follows that $A_1 \subset A_2 \cup \cdots \cup A_n$, whence the proof is completed.

6.5. COROLLARY. *For any collection of more than* $2r(S)$ *simple subsets of* S *with disjoint frontiers, the union of some two of the sets contains the intersection of the rest.*

6.6. FURTHER RESULTS. Evidently the method which was employed to prove Theorem 11 is of more general applicability; it shows, roughly speaking, that the incidences of a system of sets with disjoint frontiers are no worse than if S were a linear graph of the same degree of multicoherence. In the same way we may prove:

THEOREM 11a. *Let* $A_1, A_2, \cdots, A_n$ *be* n *simple subsets of* S, *every two of which meet, and which have disjoint frontiers. If* n *is large enough compared with* $r(S)$ *(assumed finite), then some* A_j *is contained in the union of two others.*

(Note that no A_j need be contained in *one* other, irrespective of how large n is.) Here the determination of the "best" bound for n seems to be difficult: it can be shown, however, that, disregarding the trivial case $r(S) = 0$, it lies between $\exp\{\exp[c_1 r(S)]\}$ and $\exp\{\exp[c_2 r(S)]\}$, where c_1, c_2 are positive constants.

Another related theorem, proved in a similar way, is:

THEOREM 11b. *Let* $A_1, A_2, \cdots, A_n$ *be connected subsets of* S *such that* $b_0[\mathrm{Co}(A_j)] \leq q$ $(j = 1, 2, \cdots, n)$. *Suppose that every two of the sets* A_j *meet, and that they have disjoint frontiers. Then there exists a function* N *of* q *and* $r(S)$ *(independent of* n*) such that* $\bigcup A_j$ *is contained in the union of* N *or fewer of the sets* A_j.

It is easy to show by examples that, with $q \geq 1$, we have

$$N \geq (q+1)(q+2)\, r(S) \quad \text{if} \quad r(S) \geq 1,$$

and

$$N \geq q^2 + q + 2 \quad \text{if} \quad r(S) = 0;$$

but the author does not know if these values are in fact the best.

For theorems of this type, the conditions that the sets A_j (or, more generally, their closures) be connected, and that the numbers $b_0\{\mathrm{Cl}[\mathrm{Co}(A_j)]\}$ be bounded, cannot be omitted. In [4, §8] a theorem in a similar order of ideas was obtained for *arbitrary* connected sets in a unicoherent space; it can indeed be extended to the multicoherent case, but at the cost of requiring not only that certain intersections of the sets A_j be nonempty, but that they have sufficiently many components. For example, the theorem for three sets becomes:

(1) *If A_1, A_2, A_3, are connected subsets of S such that $A_1 \cap A_2 \cap A_3 = 0$, and $b_0(A_j \cap A_k) \geq r(S)$ whenever $j \neq k$, then every two of the sets* $\mathrm{Fr}(A_j)$ *meet.*

The proof of (1) is an easy consequence of [5, §7.2].

We finally remark that the present technique can be used to give a direct "elementary" proof of Theorem 1, without using mappings in S^1. However, though the basic idea (showing that the sets have the same incidences as if S were a linear graph) is simple, a quite lengthy and tedious argument is needed to reduce the general theorem to the case in which the complements of the sets are of finite incidence; and the proof given in 2.2 above is considerably shorter.

References

1. S. Eilenberg, *Sur les espaces multicohérents* I, Fund. Math. 27 (1936), 153-190.

2. ———, *Sur les espaces multicohérents* II, Fund. Math. 29 (1937), 101-122.

3. E. Hewitt, *A problem of set-theoretic topology*, Duke Math. J. 10 (1943), 309-333.

4. A. H. Stone, *Incidence relations in unicoherent spaces*, Trans. Amer. Math. Soc. 65 (1949), 427-447.

5. ———, *Incidence relations in multicoherent spaces* I, Trans. Amer. Math. Soc. 66 (1949), 389-406.

6. ———, *Incidence relations in multicoherent spaces* II, Canadian J. Math. 2 (1950), 461-480.

MANCHESTER UNIVERSITY

Pacific Journal of Mathematics

Vol. 2, No. 1 January, 1952