NON-PARAMETRIC SURFACES GIVEN BY LINEARLY CONTINUOUS FUNCTIONS

BY

CASPER GOFFMAN

Purdue University, Lafayette, Ind., U.S.A. (¹)

1. Area was defined, [2], in 1936 by L. Cesari for the "surface" given by any equivalence class of measurable functions. Subsequently, we gave a somewhat different, but equivalent, definition. We then showed, [5], that this area agrees with the Lebesgue area for surfaces given by continuous functions, and that the Tonelli theorems remain valid for this wider class of surfaces, provided only that the notions of absolute continuity and bounded variation are suitably modified.

On the other hand, for certain purposes, this class of non-parametric surfaces is too wide. For example, it has already been observed by us, [6], that a theory of similar scope is impossible in the parametric case, since the elementary area of quasi-linear mappings from the square into euclidean 3 space is not lower semi-continuous with respect to almost everywhere convergence. Moreover, for continuous non-parametric surfaces, it has been shown by Federer, [4], (also, see Mickle and Rado, [9]), using earlier work of Besicovitch, [1], for the case of a surface given by an ACT function, that the Lebesgue area is equal to the Hausdorff 2 dimensional measure of the graph of the function. Since, in our general theory, a surface is an equivalence class of measurable functions, and since changing the values of a function on a set of measure 0 may change the Hausdorff 2 dimensional measure of its graph, it is evident that Federer's theorem cannot remain fully valid. Moreover, we must contend with phenomena such as that exhibited by the function f defined on the closed unit square I given by $f(x, y) = 0$, $x \leqslant \frac{1}{2}$ and $f(x, y) = 1$, $x > \frac{1}{2}$. In our theory, the area here is 2, since the step is included, but the Hausdorff 2 dimensional measure of the graph is 1.

(¹) Research supported by National Science Foundation grant number NSF G-5867.

All this indicates that certain specialization is needed. In the first place, we have been able to successfully define a parametric area for surfaces given by functions which we call linearly continuous. This theory is being developed in another paper. In the second place, for non-parametric surfaces given by linearly continuous functions, we shall show that the area as defined by us and by Cesari is equal to the Hausdorff 2 dimensional measure of an "essential" part of the graph. For these reasons, and for others which arise as the text develops, it appears that the linearly continuous functions deserve special attention.

We wish to thank the referee for his careful study of the paper and for his suggestions leading to its improvement.

2. In 1942, I. Verchenko, [14], showed that to every non-parametric surface of finite area, given by a continuous function f on the unit square I, there corresponds a measure μ_f such that for every open rectangle $R \subset I$, $\mu_f(R)$ is the Lebesgue area of the surface given by f restricted to R. The main result of Verchenko was that for every Borel set $E \subset I$, if $A(f) < \infty$, $A(g) < \infty$, and $f = g$ on E, then $\mu_f(E) = \mu_g(E)$. This result also follows from Federer's theorem, alluded to above, that for every surface given by a continuous f the area of the surface is equal to the Hausdorff 2 dimensional measure of the graph.

For the linearly continuous case, we shall proceed as follows: We shall first extend Verchenko's work to surfaces given by linearly continuous functions. The resulting measure will then play the essential part in the extension of Federer's theorem to these surfaces.

Indeed, we first consider any equivalence class of measurable functions. (We use the function symbol f to designate such a class, whenever there is no danger of confusion). We recall that the area of the surface given by f is defined as follows: Let $E(p)$ be the elementary area of a quasi-linear function p; i.e., p is continuous and its graph consists of a finite set of triangles, and $E(p)$ is the sum of the areas of these triangles. Then

$$A(f) = \inf \left[\liminf_{n\to\infty} E(p_n)\right],$$

where the infimum is taken over the sequences $\{p_n\}$ of quasi-linear functions defined on the closed unit square I which converge almost everywhere to f.

3. We associate a measure with f in the following way. Observe that for every open rectangle $R \subset I$, the area $A(f|R)$, of f restricted to R, may be defined in the above way, and that if $R_1 \subset R_2$ then $A(f|R_1) \leqslant A(f|R_2)$. We use the notation S° for the interior of a set S. For I itself, it follows from the definition that $A(f|I) = A(f|I^\circ)$, so that the resulting measure μ_f will have the property that $\mu_f(I) = \mu_f(I^\circ)$.

Analogous definitions to those of $A(f|I)$ and $A(f|I^\circ)$ may be given for $A(f|R)$ and $A(f|R^\circ)$ for any closed rectangle $R \subset I$. As before, $A(f|R) = A(f|R^\circ)$. However, $\mu_f(R)$ may be different from $\mu_f(R^\circ)$ since $\mu_f(R) = \inf A(f|S)$, for open rectangles S containing the closed rectangle R, may be greater than $\mu_f(R^\circ)$.

A point $(x, y) \in I^\circ$ will be called *singular* if, for every open $R \subset I^\circ$, with $(x, y) \in R$, $A(f|R) = \infty$. The set S_f of singular points of f is evidently closed. Its complement $G_f = I^\circ \sim S_f$ will be called the set of *regular* points of f.

We shall show that, for every $\varepsilon > 0$ and regular point (x, y), there is an open R such that $(x, y) \in R$ and $A(f|R) < \varepsilon$. Moreover, if $\{R_n\}$ is a decreasing sequence of open rectangles, with $\bar{R}_1 \subset G_f$, where $\bar{R}_1$ is the closure of R_1, such that their intersection is empty, then $\lim_{n\to\infty} A(f|R_n) = 0$. These facts seem to be hard to prove by direct use of the definition. However, we have obtained elsewhere, [5], [7], expressions for $A(f|R)$, in terms of variation functions, which may be used for this purpose.

4. We now give a brief recapitulation of these results in a somewhat improved form. For a Lebesgue measurable function f, of a single real variable, defined on an interval (a, b), we may define the generalized variation as

$$\varphi(f;\ (a, b)] = \inf \{\liminf_{n\to\infty} V[g_n;\ (a, b)]\},$$

where the infimum is taken for all sequences $\{g_n\}$ of continuous functions converging almost everywhere to f, and $V[g;\ (a, b)]$ is the variation of g on (a, b). In [5], for summable f using L_1 convergence, the function φ was shown to be equal to the ordinary variation of f, restricted to its set of points of approximate continuity, or to the min $V[g;\ (a, b)]$, where g is any function equivalent to f. This holds as well in the present situation.

We use the term oriented rectangle for one whose sides are parallel to the coordinate axes. Thus oriented refers to direction; the term is often used for sense orientation in other works. For a function f, defined on an oriented rectangle $R = (a, b) \times (c, d)$, for every $x \in (a, b)$, we designate by f_x the function $f(x, y)$ of y defined for all $y \in (c, d)$ and, for every $y \in (c, d)$, we designate by f_y the function $f(x, y)$ of x defined for all $x \in (a, b)$. We then have the generalized variations.

$$\Phi_1(R) = \int_c^d \varphi[f_y;\ (a, b)]\, dy \quad \text{and} \quad \Phi_2(R) = \int_a^b \varphi[f_x;\ (c, d)]\, dx.$$

These variations were first defined by Cesari, [2], in a different way, and have been used by Krickeberg, [8], Pauc, [11], de Vito, [15], and others. We are interested in the variation $\Psi^{*}(f)$ of the rectangle function

$$\Phi(R) = [\Phi_1^2(R) + \Phi_2^2(R) + |R|^2]^{\frac{1}{2}}.$$

That is to say, for any finite set $R = \{R_1, \ldots, R_n\}$ of non-overlapping oriented rectangles consider

$$\Phi(R) = \sum_{i=1}^{n} \Phi(R_i),$$

and let $\Psi^{*}(f) = \sup \Phi(R)$ for all R. This is an example of a so-called lower, or Gëocze, area. In previous work, we had used the notion of admissible subdivision in defining a lower area. There we used the functions

$$\Theta_1(R) = \int_c^d |f(b, y) - f(a, y)|\, dy \quad \text{and} \quad \Theta_2(R) = \int_a^b |f(x, d) - f(x, c)|\, dx.$$

and defined
$$\Theta(R) = [\Theta_1^2(R) + \Theta_2^2(R) + |R|^2]^{\frac{1}{2}}.$$

We then called a rectangle admissible if f is approximately continuous, separately in each variable, almost everywhere on the boundary of R. We then defined a Gëocze area as the Burkill integral of $\Theta(R)$ restricted to admissible rectangles. We call this $X(f)$.

We point out that $\Psi^{*}(f) = X(f)$. Since, for every R, $\Phi_1(R) \geqslant \Theta_1(R)$ and $\Phi_2(R) \geqslant \Theta_2(R)$, it is obvious that $\Psi^{*}(f) \geqslant X(f)$. On the other hand, we showed in [5] that $\Phi_1(R)$ and $\Phi_2(R)$ are, respectively, the upper Burkill integrals of $\Theta_1(S)$ and $\Theta_2(S)$, where S is an arbitrary admissible oriented rectangle in R. Using the inequality

$$[(a_1 + \cdots + a_n)^2 + (b_1 + \cdots + b_n)^2 + (c_1 + \cdots + c_n)^2]^{\frac{1}{2}} \leqslant (a_1^2 + b_1^2 + c_1^2)^{\frac{1}{2}} + \cdots + (a_n^2 + b_n^2 + c_n^2)^{\frac{1}{2}}, \quad (*)$$

it follows by a straightforward calculation that $\Psi^{*}(f) \leqslant X(f)$.

Since we know from [5] that $X(f) = A(f)$, it follows that $\Psi^{*}(f) = A(f)$.

Moreover, it easily follows from the definition that the following inequality holds for every R:

$$\Phi_1(R) + \Phi_2(R) + |R| \geqslant \Psi^{*}(R) \geqslant \max [\Phi_1(R), \Phi_2(R), |R|].$$

5. Now, suppose (x, y) is a regular point. There is then an $R = (a, b) \times (c, d)$ with $(x, y) \in R$ and $\Psi^{*}(R) < \infty$. Since $\Phi_1(S) < \Psi^{*}(S)$, for every S, and

$$\Phi_1(S) = \int_{c'}^{d'} \varphi[f_y; (a', b')]\, dy,$$

for $S = (a', b') \times (c', d')$, there is a $\delta' > 0$ such that $\Phi_1((a, b) \times (y - \delta', y + \delta')) < \frac{1}{3}\varepsilon$. Similarly, there is a $\delta'' > 0$ such that $\Phi_2((x - \delta'', x + \delta'') \times (c, d)) < \frac{1}{3}\varepsilon$. Let

$$\delta = \min[\delta', \delta'', \sqrt{\varepsilon/12}].$$

Then, using (*), it follows that

$$\Psi((x - \delta, x + \delta) \times (y - \delta, y + \delta)) < \varepsilon.$$

Since $A(f|R) = \Psi(R)$, it follows that there is a decreasing sequence $\{R_n\}$ of open rectangles containing (x, y) such that $\lim_{n\to\infty} A(f|R_n) = 0$.

Suppose, next, that $\{R_n\}$ is a decreasing sequence of oriented rectangles, with $R_1 \subset G_f$, and $\bigcap_{n=1}^{\infty} R_n$ empty. Then either the lengths or the widths of the R_n converge to 0. Suppose the widths, $b_n - a_n$, converge to 0. Then it is evident that $|R_n|$ and $\Phi_2(R_n)$ converge to 0, provided $\Phi_2(R_n)$ is finite for some n. To show that $\Phi_1(R_n)$ converges to 0, we need a fact regarding the generalized variation of a function of a single variable. Let $\varphi[f; (a, b)] < \infty$, and let $I_n = (a_n, b_n)$ be a decreasing sequence of open intervals in (a, b) whose intersection is empty. Since there is a g epuivalent to f such that $V(g; (\alpha, \beta)) = \varphi[f; (\alpha, \beta)]$, for every $(\alpha, \beta) \subset (a, b)$, it follows that $\lim_{n\to\infty} \varphi[f; (a_n, b_n)] = 0$. Let $R_n = (a_n, b_n) \times (c_n, d_n)$. Then

$$\Phi_1(R_n) = \int_{c_n}^{d_n} \varphi[f_y; (a_n, b_n)]\, dy.$$

The sequence of functions $\{\varphi[f_y; (a_n, b_n)]\}$ are summable, decreasing, and converge to 0 almost everywhere. Hence

$$\lim_{n\to\infty} \Phi_1(R_n) = \lim_{n\to\infty} \int_{c_n}^{d_n} \varphi[f_y; (a_n, b_n)]\, dy = 0.$$

It now follows that $\lim_{n\to\infty} A(f|R_n) = 0$.

6. We now define the measure μ_f for all Borel sets in G_f. For every open oriented rectangle $R \subset G_f$, let $\mu_f(R) = A(f|R)$. For every open oriented line segment L we define $\mu_f(L) = \lim_{n\to\infty} \mu_f(R_n)$, where $\{R_n\}$ is a decreasing sequence of open rectangles whose intersection is L. Since, if $\{S_n\}$ were another such sequence of open rectangles, then for every n, there are m' and m'' such that $S_{m'} \subset R_n$ and $R_{m''} \subset S_n$, it is clear that $\mu_f(L)$ is uniquely defined. For every point $P \in G_f$, we define $\mu_f(P) = 0$. The definition is justified by the fact that if $P = \bigcap_{n=1}^{\infty} R_n$, where $\{R_n\}$ is a decreasing sequence of rectangles, then $\lim_{n\to\infty} A(f \mid R_n) = 0$.

Let $S \subset G_f$ be the union of a finite number of pair-wise disjoint sets $S_1, S_2, \ldots, S_n$ each of which is either an open oriented rectangle, an open oriented line segment, or a point. For every such $S = [S_1, S_2, \ldots, S_n]$, we let $\mu_f(S) = \sum_{i=1}^{n} \mu_f(S_i)$.

We have shown [5, § 6] that if R is a rectangle and $R = \bigcup_{i=1}^{n} S_i$, where the S_i are open rectangles, open lines, or points, and are pair-wise disjoint, then $\mu_f(R) = \sum_{i=1}^{n} \mu_f(S_i)$. It then follows that for if S has two representations $S = [S_1, S_2, \ldots, S_n]$ and $S = [S_1', S_2', \ldots, S_m']$ then $\sum_{i=1}^{n} \mu_f(S_i) = \sum_{i=1}^{m} \mu_f(S_i')$ so that $\mu(S)$ is well defined.

We note that the family $\mathfrak{S} = [S]$ is a ring of sets, for it is easily seen to be closed with respect to finite unions and differences. Furthermore, we show that if $\{S_n\}$ is a decreasing sequence of sets in $\mathfrak{S}$, with $\mu_f(S_1) < \infty$ and $\bigcap_{n=1}^{\infty} S_n$ empty, then $\lim_{n\to\infty} \mu_f(S_n) = 0$. For this proof, we use auxiliary measures μ_f^1 and μ_f^2, which are defined in exactly the same way as μ_f, except that the functions $\Phi_1(R)$ and $\Phi_2(R)$ are used instead of $A(f \mid R)$. The basic inequality

$$\mu_f^1(S) + \mu_f^2(S) + |S| \geqslant \mu_f(S) \geqslant \max\,[\mu_f^1(S),\ u_f^2(S),\ |S|]$$

then holds.

In order to show that $\lim_{n\to\infty} \mu_f(S_n) = 0$, we need only show that

$$\lim_{n\to\infty} \mu_f^1(S_n) = \lim_{n\to\infty} \mu_f^2(S_n) = \lim_{n\to\infty} |S_n| = 0.$$

It is only necessary to show that $\lim_{n\to\infty} \mu_f^1 = 0$. For this, for every $y_0 \in (0, 1)$, let S_{n, y_0} be the linear set of x for which $(x, y_0) \in S_n$. It is apparent that

$$\mu_f^1(S_n) = \int_0^1 \varphi\,[f_y\,;\,S_{n,y}]\,dy.$$

But, by the same argument as given before, the sequence of functions $\{\varphi\,[f_y\,;\,S_{n,y}]\}$ are summable, decreasing, and converge almost everywhere to 0. Hence,

$$\lim_{n\to\infty} \mu_f^1(S_n) = \lim_{n\to\infty} \int_0^1 \varphi[f_y; S_{n,y}]\,dy = 0.$$

It follows that μ_f is completely additive on S and may be extended to the Borel sets in G_f. Moreover, if $E \subset G_f$ is compact, then $\mu_f(E) < \infty$, since every point $P \in E$ is contained in an open rectangle R such that $\mu_f(R) < \infty$, and a finite number of these rectangles covers E. Since G_f is the union of countably many compact sets, the measure is totally σ finite for G_f. We summarize the above results in

THEOREM 1. *If f is an equivalence class of measurable functions defined on $I = (0, 1) \times (0, 1)$, there corresponds to f a closed set S_f and its complement $G_f = I \sim S_f$, and a measure μ_f such that $\mu_f(S) = \infty$ for every non-empty $S \subset S_f$, μ_f is totally σ finite on G_f, and $\mu_f(R) = A(f|R)$ for every open rectangle R.*

We remark that if $S \subset G_f$ and $S \subset Z_1 \times Z_2$, where $|Z_1| = |Z_2| = 0$, then $\mu_f(S) = 0$. For, evidently, $\mu_f^1(S) = \mu_f^2(S) = |S| = 0$, and $\mu_f(S) \leqslant \mu_f^1(S) + \mu_f^2(S) + |S|$.

7. We digress briefly to a discussion of the measure associated with a continuous, parametric surface. This theory has been developed by Cesari, [3], for surfaces of finite area. Consider the closed square $I = [0, 1] \times [0, 1]$. Let f be a continuous mapping from I into euclidean 3 space E_3. For every point $P \in E_3$ consider the set Γ_P of components of $f^{-1}(P)$. Let $\Gamma = \bigcup_{P \in E_3} \Gamma_P$. The sets $\gamma \varepsilon \Gamma$ are pair-wise disjoint and their union is I. We topologize Γ as follows: A subset $G \subset \Gamma$ will be open if $\bigcup_{\gamma \in G} \gamma$ is an open set in I. It is known, [3], [12], that Γ is then a Peano space. In particular, every open set is of type F_σ, and the space Γ is normal. It follows that, for every open $G \subset \Gamma$, there are

$$G_1 \subset \bar{G}_1 \subset \cdots \subset G_n \subset \bar{G}_n \subset \cdots$$

such that $G = \bigcup_{n=1}^{\infty} G_n$, and $\bar{G}_n$ is the closure of G_n, for every n. Evidently, $\bar{G}_n$ is compact.

A point $\gamma \in \Gamma$ will be called singular if, for every open $G \subset \Gamma$, $\gamma \in G$, we have $A(f|G) = \infty$. Otherwise, γ is called regular. The set G_f of regular points is open relative to Γ, and the set $\bigcup_{\gamma \in G_f} \gamma$ is open relative to I. Furthermore, $G_f = \bigcup_{n=1}^{\infty} G_n$, where $G_1 \subset \bar{G}_1 \subset \cdots \subset G_n \subset \bar{G}_n \subset \cdots$. For each G_n, we consider the Cesari measure for the Borel sets in G_n. This is defined since $A(f|G_n) < \infty$, the G_n being contained in compact sets composed of regular points. Moreover, every G_n is an admissible set à la Cesari,

[12]. It also follows that, for every n, $A(f|G_n)$ equals the Cesari measure of G_n as a subset of G_{n+1} obtained by considering the function f restricted to G_{n+1}.

Thus we obtain a uniquely defined measure, perhaps having infinite values for some sets, for the Borel subsets of G_f; recall that these are subsets of Γ not of I. It is a totally σ finite measure. It would seem that an improper surface integral may also be defined for such surfaces, and we intend to discuss these matters elsewhere.

8. In this section, we point out that an analog of Verchenko's theorem for non-parametric surfaces of finite area does not hold for continuous parametric surfaces of finite area, even for the case where the mappings involved are homeomorphisms.

For this purpose, we construct a closed, zero dimensional set in the unit square as follows: Consider a finite set of pair-wise disjoint, oriented, closed squares; call these squares of rank 1. In each square of rank $n \geqslant 1$, consider a finite set of pair-wise disjoint, oriented, closed squares; call these squares of rank $n+1$. For every $n > 1$, label the squares of rank n with n subscripts $R_{i_1 i_2 \ldots i_n}$ in such a way that $R_{i_1 i_2 \ldots i_{n-1} i_n} \subset R_{i_1 i_2 \ldots i_{n-1}}$. Let $S_n = \bigcup R_{i_1 i_2 \ldots i_n}$ be the union of all squares of rank n, and let $S = \bigcap_{n=1}^{\infty} S_n$. Moreover, choose the squares so that every square of rank n has diagonal of length less than $1/n$, and so that $|S| = \frac{1}{2}$. It is clear that S is zero dimensional.

Directly above the center of I, at height 2, place a square σ, and directly above the center of each $R_{i_1 i_2 \ldots i_k}$, at height $1/k$, place a square $\sigma_{i_1 i_2 \ldots i_k}$. From each of these squares punch as many holes as there are squares $R_{i_1 i_2 i \ldots kn}$ and join the boundaries of these holes by means of the boundaries of pipes to the boundaries of the squares $\sigma_{i_1 i_2 \ldots i_k n}$. For every n, the boundary of σ and its interior, minus the punched holes, together with the pipes leading to the boundaries of the σ_{i_1}, their interiors minus the punched holes, and so on, until finally the pipes leading to the boundaries of the $\sigma_{i_1 i_2 \ldots i_n}$ and the interiors of the $\sigma_{i_1 i_2 \ldots i_n}$, form a parametric continuous surface S_n, given by an f_n which maps each $R_{i_1 i_2 \ldots i_n}$ onto $\sigma_{i_1 i_2 \ldots i_n}$. Moreover, the pipes can be chosen so that there are no intersections and the mapping f_n is a homeomorphism. Furthermore, the squares and pipes can be chosen sufficiently small that $A(S_n) < \frac{1}{4}$, for every n.

The sequence $\{f_n\}$ converges uniformly to a continuous mapping f for which all points in S are fixed. Let S be the surface of this mapping. Then, since

$$A(S) \leqslant \liminf_{n \to \infty} A(S_n), \quad A(S) \leqslant \tfrac{1}{4}.$$

Accordingly, it necessarily follows that $\mu_f(S) \leqslant \frac{1}{4}$. Consider the identity mapping g of the unit square onto itself. Clearly $\mu_g(S) = \frac{1}{2}$.

9. A real function f, on the closed unit square I, will be called *essentially linearly continuous* if there are g and h, equivalent to f, such that g_x is continuous in y for almost all x, and h_y is continuous in x for almost all y. The function f will be called *linearly continuous* if f_x is continuous in y for almost all x and f_y is continuous in x for almost all y. A linearly continuous function can be infinite on a subset of a set of the form $Z_1 \times Z_2$ where Z_1, Z_2 are of linear Lebesgue measure zero.

A related property is one which we call *strong linear continuity.* f has this property if, for every line L, f is a continuous function of one variable on almost all lines parallel to L.

Certain special kinds of sets are of interest to us. Let $U = [0, 1]$, $V = [0, 1]$, and $I = U \times V$. A set $Z \subset I$ will be called *negligible* if there are $A \subset U$, $B \subset V$, with $Z = A \times B$, and $|A| = |B| = 0$. A set $\mathcal{E}$ will be called *elementary* if

$$\mathcal{E} = (E_1 \times V) \cup (U \times E_2),$$

where $E_1 \subset U$ and $E_2 \subset V$ are closed sets.

We first show that if f is linearly continuous, then for every $\varepsilon > 0$, there is an elementary set $\mathcal{E}$ such that $|\mathcal{E}| > 1 - \varepsilon$ and f is continuous on $\mathcal{E}$.

LEMMA 1. *If f is linearly continuous on I, then for every $\varepsilon > 0$, $\delta > 0$, the subset $E \subset U$, for which $x \in E$ and $|y - y'| < \delta$ implies $|f(x, y) - f(x, y')| < \varepsilon$, is measurable.*

Proof. Let $x \in E_n$ if and only if f_x is a continuous function of y, and there are y, y' with $|y - y'| > \delta$ and $|f(x, y) - f(x, y')| > \varepsilon - 1/n$. Let $x_0 \in E_n$, $|y - y'| < \delta$, and $|f(x_0, y) - f(x_0, y')| > \varepsilon - 1/n$. Then there are y_1, y_2 such that f_{y_1} and f_{y_2} are continuous functions of x, $|y_1 - y_2| < \delta$, and $|f(x_0, y_1) - f(x_0, y_2)| > \varepsilon - 1/n$. There is an open interval $G \subset U$, $x_0 \in G$, such that for every $x \in G$, $|f(x, y_1) - f(x, y_2)| > \varepsilon - 1/n$. Thus, E_n is measurable. But E differs from $C(\bigcup_{n=1}^{\infty} E_n)$ by a set of measure 0, so that E is measurable.

LEMMA 2. *If f is linearly continuous on I, then for every $\varepsilon > 0$ there is a closed set $E \subset U$, $|E| > 1 - \varepsilon$, such that the saltus of f, relative to $E \times V$, is less than ε at every point of $E \times V$.*

Proof. There is a $\delta > 0$, such that if E is the subset of U for which $x \in E$ and $|y_1 - y_2| < \delta$ implies $|f(x, y_1) - f(x, y_2)| < \varepsilon$, then the exterior measure of E exceeds $1 - \varepsilon$. But, by Lemma 1, E is measurable, so that we may assume it is closed.

Suppose $x_0 \in E$, $x_n \in E$, $n=1, 2, \ldots$, and that $\{(x_n, y_n)\}$ converges to (x_0, y_0). Consider the subsequence of $\{(x_n, y_n)\}$ for which $x_n \geqslant x_0$, $y_n \geqslant y_0$; call this subsequence also $\{(x_n, y_n)\}$. There is a y' such that $0 < y' - y_0 < \delta$ and $f_{y'}$ is a continuous function of x. There is then a $\delta' > 0$ such that $0 < x' - x_0 < \delta'$ implies $|f(x_0, y') - f(x', y')| < \varepsilon$. Now, if n is large enough, (x_n, y_n) is in the rectangle $[x_0, x_0 + \delta'] \times [y_0, y']$. Then

$$|f(x_n, y_n) - f(x_0, y_0)| \leqslant |f(x_n, y_n) - f(x_n, y')| + |f(x_n, y') - f(x_0, y')| + |f(x_0, y') - f(x_0, y_0)|.$$

But the first and third terms on the right hand side are each less than ε since $x_0, x_n \in E$ and $|y_n - y'| < \delta$, $|y_0 - y'| < \delta$, and the second term is less than ε since $|y_n - y_0| < \delta'$. The subsequences of $\{(x_n, y_n)\}$ for which $x_n \geqslant x_0$, $y_n \leqslant y_0$; $x_n \leqslant x_0$, $y_n \geqslant y_0$; and $x_n \leqslant x_0$, $y_n \leqslant y_0$ may be treated similarly. Putting all together, we obtain the lemma.

LEMMA 3. *If f is linearly continuous on I, then for every $\varepsilon > 0$ there are closed sets $E \subset U$, $F \subset V$, with $|E| > 1 - \varepsilon$, $|F| > 1 - \varepsilon$, such that if $\mathcal{E} = (E \times V) \cup (U \times F)$, then the saltus of f, relative to $\mathcal{E}$, is less than ε at every point of $\mathcal{E}$.*

Proof. By Lemma 2, there are closed sets $E \subset U$, $F \subset V$, such that $|E| > 1 - \varepsilon$, $|F| > 1 - \varepsilon$, and the saltus of f, relative to $E \times V$, is less than ε at every point of $E \times V$, and the saltus of f, relative to $U \times F$, is less than ε at every point of $U \times F$. Clearly, the saltus of f is less than ε, relative to $\mathcal{E}$, at every point of $\mathcal{E} \sim ((U \times F) \cap \cap (V \times E))$. Let $(x, y) \in (U \times F) \cap (E \times V)$. There is a circle K_1, center (x, y), such that the saltus of f is less than ε on the set $K_1 \cap (U \times F)$, and a circle K_2, center (x, y), such that the saltus of f is less than ε on the set $K_2 \cap (E \times V)$. Hence the saltus of f, relative to $\mathcal{E}$, is less than ε at (x, y).

LEMMA 4. *If f is linearly continuous on I, then for every $\varepsilon > 0$, there are closed sets $E \subset U$, $F \subset V$, with $|E| > 1 - \varepsilon$, $|F| > 1 - \varepsilon$, such that f is continuous on*

$$\mathcal{E} = (E \times V) \cup (U \times F),$$

relative to $\mathcal{E}$.

Proof. Let $\{\varepsilon_n\}$ be a sequence of positive numbers with $\sum_{n=1}^{\infty} \varepsilon_n = \varepsilon$. For every n, let E_n, F_n be such that $|E_n| > 1 - \varepsilon_n$, $|F_n| > 1 - \varepsilon_n$, and such that the saltus of f is less than ε_n at every point of $\mathcal{E}_n = (E_n \times V) \cup (U \times F_n)$, relative to $\mathcal{E}_n$. Let $E = \bigcap_{n=1}^{\infty} E_n$, $F = \bigcap_{n=1}^{\infty} F_n$. Then $|E| > 1 - \varepsilon$. $|F| > 1 - \varepsilon$, and f is continuous on $\mathcal{E} = (E \times V) \cup (U \times F)$ relative to $\mathcal{E}$.

In other words, for every $\varepsilon > 0$, there is an elementary set of "size" $1-\varepsilon$ on which f is uniformly continuous.

10. A measurable function f on I has finite area if and only if it is of bounded variation in the sense of Cesari (BVC). This means that

$$\int_0^1 \varphi(f_x)\,dx < \infty \quad \text{and} \quad \int_0^1 \varphi(f_y)\,dy < \infty;$$

or, that there are functions g and h, equivalent to f, such that

$$\int_0^1 V(g_x)\,dx < \infty \quad \text{and} \quad \int_0^1 V(h_y)\,dy < \infty.$$

We prove the existence of g with this property. We first note that if S is a planar measurable set then the linear density in the y direction of S exists and equals 1 almost everywhere on S. It then follows that f is approximately continuous in y almost everywhere. Thus if for every x, we let g_x be the upper measurable boundary of f_x then g_x is an m function. We then define $g(x, y) = g_x(y)$. The function g is measurable and such that $\int_0^1 V(g_x)\,dx = \int_0^1 \varphi(f_x)\,dx$.

Before going any further, we first observe that if f is BVC then it is summable. For this, we note that, for almost all x, the saltus, $\omega(g_x)$, of g_x is less than or equal to $V(g_x)$. Moreover, there is a y_0 (indeed, almost every y has this property) such that g_{y_0} has bounded variation relative to a subset E of $[0, 1]$ of measure 1. There is thus an M such that $|g(x, y_0)| < M$ for every $x \in E$. It follows that, for almost all x, $|f(x, y)| \leqslant \omega(g_x) + M$, for every y. Thus,

$$\int_0^1 \int_0^1 |f(x, y)|\,dy\,dx \leqslant \int_0^1 (\omega(g_x) + M)\,dx \leqslant \int_0^1 (V(g_x) + M)\,dx \leqslant \Phi_1(f) + M.$$

Hence, f is summable.

Since we shall first consider the finite area case, we shall, for the time being, deal only with summable functions.

Suppose f is essentially linearly continuous. We shall show that if f is also BVC, then it is equivalent to a linearly continuous g; indeed, to a strongly linearly continuous one. In the process of proving this, we shall also obtain the fact that for a linearly continuous f, of finite area, the integral means converge to f, in a quite strong sense, which we call linear convergence.

We need the known

LEMMA 5. *If f is a continuous, monotonically non-decreasing, real function on $[0, 1]$, if $\{f_n\}$ is a sequence of continuous, monotonically non-decreasing functions converging to f almost everywhere, then $\{f_n\}$ converges to f everywhere, and the convergence is uniform.*

11. By the integral means f^h, $h > 0$, of a summable f, we mean the functions

$$f^h(x, y) = h^{-2} \int_0^h \int_0^h f(x+u, y+v)\, du\, dv.$$

Observe that if f is defined on I, then f^h is defined on $[0, 1-h] \times [0, 1-h]$.

The first goal is to prove the summability of f.

Let f be essentially linearly continuous and BVC, and let g be equivalent to f and such that g_x is continuous for almost all x. Now, $V(g_x) < \infty$ almost everywhere, so that, for almost all x, g_x has a canonical representation

$$g_x = g_x^+ - g_x^-$$

as a difference of monotonically non-decreasing functions of y. For every $y \in [0, 1]$, and every real c, this representation may be defined so that $g_x^+(y) = c$. Now, choose y_0 so that there is an M such that $|g(x, y_0)| \leqslant M$ for almost all x, let E be the set of these x, and choose the representation functions g_x^+ and g_x^- so that $g_x^+(y_0) = g(x, y_0)$ and $g_x^-(y_0) = 0$, for all $x \in E$. Finally, define the functions g^+, g^- by $g^+(x, y) = g_x^+(y)$, $g^-(x, y) = g_x^-(y)$, for $x \in E$, and $g^+(x, y) = g^-(x, y) = 0$, everywhere else.

The functions g^+ and g^- are Lebesgue measurable functions with respect to 2 dimensional Lebesgue measure. We give the proof for g^+. g_{y_0} is measurable, as is g_y for almost all y. Fix n and consider $y_{-k} < \cdots < y_0 < y_1 < \cdots y_s$ where $y_{-k} < 1/n$, $y_s > 1 - 1/n$ and so that $y_{i+1} - y_i < 1/n$ and g_{y_i} is measurable, for every i. For every i, let E_i be the set of x for which $g(x, y_{i+1}) \geqslant g(x, y_i)$, and g_x is continuous. Then E_i is measurable.

Define the function k_n as follows: If g_x is not continuous, let $k_n(x, y) = 0$, for every y. If g_x is continuous, define $k_n(x, y)$ as the quasi-linear function of y which

a) is constant on $[y_i, y_{i+1}]$ if $x \notin E_i$,

b) is such that $k_n(x, y_{i+1}) - k_n(x, y_i) = g(x, y_{i+1}) - g(x, y_i)$, and linear on the interval $[y_i, y_{i+1}]$, if $x \in E_i$, and

c) $k_n(x, y_0) = g(x, y_0)$.

The functions k_n are measurable and converge almost everywhere to g^+. Hence g^+ is measurable.

Since $V(g_x) = V(g_x^+) + V(g_x^-)$, for almost all x, and $\int_0^1 V(g_x)\,dx < \infty$, it follows that

$$\int_0^1\int_0^1 |g^+(x, y)|\,dx\,dy \leqslant \int_0^1 (V(g_x^+) + M)\,dx \leqslant \int_0^1 V(g_x)\,dx + M < \infty,$$

and
$$\int_0^1\int_0^1 |g^-(x, y)|\,dx\,dy \leqslant \int_0^1 V(g_x^-)\,dx \leqslant \int_0^1 V(g_x)\,dx < \infty.$$

Hence the functions g^+ and g^- are summable.

It is necessary now to make some notational clarifications.

We consider an essentially linearly continuous f which is BVC and the corresponding g, defined above, such that $g = f$ almost everywhere and g_x is continuous for almost all x. Then g has corresponding g^+ and g^- defined in the above way. For every $k > 0$, we have the integral means of all three functions, which we designate by g^k, $(g^+)^k$, and $(g^-)^k$, respectively. We also have the associated functions of a single real variable. $(g^k)_x$, $(g^k)_y$, $((g^+)^k)_x$, $((g^+)^k)_y$, $((g^-)^k)_x$, and $((g^-)^k)_y$. For example, for every $x \in [0, 1]$, the function $((g^+)^k)_x$ is defined by

$$((g^+)^k)_x(y) = (g^+)^k(x, y).$$

Suppose, now, that x is fixed and $y_1 < y_2$. Then,

$$(g^+)^k(x, y_2) - (g^+)^k(x, y_1) = k^{-2}\int_0^k\int_0^k \{g^+(x+u, y_2+v) - g^+(x+u, y_1+v)\}\,du\,dv \geqslant 0,$$

since the integrand is positive almost everywhere. Thus, the functions $((g^+)^k)_x$ are monotonically non-decreasing functions of y, on $[0, 1-k]$, for every x, However, it is well known that $(g^+)^k$ converges almost everywhere to g^+. By Lemma 5, for almost all x, $((g^+)^k)_x$ converges everywhere to $(g^+)_x$, the convergence being uniform on $[0, 1-\varepsilon]$, for every $\varepsilon > 0$. Indeed, by extending the definition of $(g^+)^k$ to all of I, in the following way, we obtain that $((g^+)^k)_x$ converges uniformly to $(g^+)_x$ on $[0, 1]$ for almost all x:

For $x \in [0, 1-k]$, $y \in [1-k, 1]$, let $(g^+)^k(x, y) = (g^+)^k(x, 1-k)$.

For $x \in [1-k, 1]$, $y \in [0, 1-k]$, let $(g^+)^k(x, y) = (g^+)^k(1-k, y)$,

For $x \in [1-k, 1]$, $y \in [1-k, 1]$, let $(g^+)^k(x, y) = (g^+)^k(1-k, 1-k)$.

Similarly, the functions $((g^-)^k)_x$ converge uniformly to $(g^-)_x$ for almost all x. Since $(g^+ - g^-)^k = (g^+)^k - (g^-)^k$, it follows that $(g^k)_x$ converges uniformly to g_x for almost all x.

Since the integral means of g and h are the same, $(g^k)_y$ converges uniformly to h_y for almost all y. We thus see that $\lim_{k\to 0} g^k$ exists everywhere, except possibly at points belonging to a negligible set, and the limit is linearly continuous, and is equivalent to f. This proves

THEOREM 2. *If f, defined on I, is essentially linearly continuous and BVC, then f is equivalent to a linearly continuous g.*

One further remark may be made. If instead of averaging over squares, we had averaged over circles of radius k, with center at (x, y), in determining f^k, our proof shows that if f is essentially strongly linearly continuous then f is equivalent to a strongly linearly continuous g. We thus have

COROLLARY 1. *If f, defined on I, is essentially strongly linearly continuous and BVC, then f is equivalent to a strongly linearly continuous g.*

12. Just as continuous functions go with uniform convergence, summable functions with L_1 convergence, etc., linear continuity has its own particular kind of convergence. This is linear convergence, and the above proof implies that if f is linearly continuous and BVC, then f^k converges linearly to f.

A sequence $\{g_n\}$ of functions defined on I *converges linearly* to a function g, if $(g_n)_x$ converges uniformly to g_x, for almost all x, and $(g_n)_y$ converges uniformly to g_y, for almost all y. This type of convergence has an associated metric when restricted to the set of linearly continuous functions. Let f, g be linearly continuous, and let $d(f, g)$ be the infimum of the set of all k for which there are $E \subset U$, $F \subset V$, $|E| > 1-k$, $|F| > 1-k$ such that $|f(x, y) - g(x, y)| < k$ on $(E \times V) \cup (U \times F)$.

The space of linearly continuous functions, with this metric, is evidently a complete metric space $\mathfrak{C}$. Theorem 2 has the

COROLLARY 2. *If f is linearly continuous and BVC then the integral means f^k converge linearly to f.*

We also have

THEOREM 3. *The space $\mathfrak{C}$ is the metric space completion of the space $\mathfrak{D}$ of quasi-linear functions when given the metric of linear convergence.*

Proof. We need only show that $\mathfrak{D}$ is dense in $\mathfrak{C}$. Let $f \in \mathfrak{C}$. By Lemma 4, there is a set $\mathcal{E} = (E \times V) \cup (U \times F)$ such that E and F are closed, $|E| > 1 - \frac{1}{2}\varepsilon$, $|F| > 1 - \frac{1}{2}\varepsilon$, and f is continuous on $\mathcal{E}$ relative to $\mathcal{E}$. By the Tietze extension theorem, there is a continuous g on I such that $g = f$ on $\mathcal{E}$. There is a quasi-linear p such that $|p(x, y) - g(x, y)| < \varepsilon$ for every $(x, y) \in I$. But then $d(f, p) < \varepsilon$.

13. For non-parametric surfaces, the Lebesgue area is customarily defined for continuous surfaces as an extension from the quasi-linear ones using the topology of uniform convergence; and, the area we consider for measurable functions is obtained, in similar fashion, using the topology of convergence in measure (or convergence almost everywhere). We show that the area $\alpha(f)$, obtained, in this way, for the linearly continuous functions, using linear convergence, equals $A(f)$.

Let f be linearly continuous, and let

$$\alpha(f) = \inf\,[\liminf_{n\to\infty} E\,p_n)],$$

where the infimum is taken for all sequences $\{p_n\}$ of quasi-linear functions which converge linearly to f. We shall show that $\alpha(f) = A(f)$ by noting that $\alpha(f) = \Psi(f)$.

In the first place, α is lower semi-continuous on $\mathcal{D}$ with respect to linear convergence, since it is lower semi-continuous on $\mathcal{D}$ with respect to almost everywhere convergence. It then follows, from the definition, that α is lower semi-continuous on $\mathcal{C}$. Moreover, if Φ is a real valued functional on $\mathcal{C}$, which is lower semi-continuous and equals the elementary area on $\mathcal{D}$, then for every $f \in \mathcal{C}$, $\Phi(f) \leqslant \alpha(f)$. Since the functional Ψ is lower semi-continuous on $\mathcal{C}$, with respect to almost everywhere convergence, it is also lower semi-continuous with respect to linear convergence. Moreover, we know, [7], that $\lim_{k\to 0} \Psi(f^k) = \Psi(f)$, for every $f \in \mathcal{C}$; indeed, for every summable f. If f is BVC, f^k converges linearly to f. There then exists a sequence $\{p_n\}$ of quasi-linear functions, converging linearly to f, with $\lim_{n\to\infty} E(p_n) = \Psi(f)$. Hence $\alpha(f) \leqslant \Psi(f)$, so that $\alpha(f) = \Psi(f)$. We now know that if f is BVC then $\alpha(f) = A(f)$. It remains only to show that if f is not BVC then $\alpha(f) = \infty$. We need only observe that, for every linearly continuous f, $\alpha(f) \geqslant \Phi_1(f), \Phi_2(f)$. This clearly holds for quasi-linear p. Let $\{p_n\}$ converge linearly to f, and be such that $\alpha(f) = \lim_{n\to\infty} E(p_n)$, where if $\alpha(f) = \infty$, we mean that $E(p_n)$ approaches infinity. Since Φ_1 is lower semi-continuous with respect to linear convergence, we have

$$\Phi_1(f) \leqslant \liminf_{n\to\infty} \Phi_1(p_n) \leqslant \lim_{n\to\infty} E(p_n) = \alpha(f).$$

Also $\Phi_2(f) \leqslant \alpha(f)$, in similar fashion. If f is not BVC, either $\Phi_1(f) = \infty$ or $\Phi_2(f) = \infty$ so that $\alpha(f) = \infty$.

14. Let f be BVC and linearly continuous on I. Suppose $\mathcal{E} = (E \times V) \cup (U \times F)$ is an elementary set, and f is continuous on $\mathcal{E}$ relative to $\mathcal{E}$. We shall obtain an expression for $\mu_f(\mathcal{E})$ in terms of the measures μ_f^1 and μ_f^2.

Let $R=[a,b]\times[c,d]$ be an oriented rectangle. Since the functions f_x and f_y are continuous, for almost all x and almost all y, respectively, and of bounded variation, the measures corresponding to their variations are given by the expressions of Banach, [10], now to be described.

Let g be a continuous real function on $[0, 1]$, of bounded variation. The measure corresponding to the variation is then given, for a Borel set E, by

$$m_f(E)=\int_{-\infty}^{\infty} N(f, y, E)\,dy,$$

where $N(f, y, E)$ is the number of points, $x\in E$, for which $f(x)=y$. Now,

$$\mu_f^1(\mathcal{E}\cap R)=\int_c^d m_{f_y}((\mathcal{E}\cap R)_y)\,dy$$

and

$$\mu_f^2(\mathcal{E}\cap R)=\int_a^b m_{f_x}((\mathcal{E}\cap R)_x)\,dx.$$

Since, from the Banach expression, if f_1 and f_2 are continuous real functions, of bounded variation, on $[0, 1]$, and $f_1=f_2$ on a Borel set E, then $m_{f_1}(E)=m_{f_2}(E)$, it follows that if f and g are BVC and linearly continuous on I, then if $f=g$ on an elementary set $\mathcal{E}$, we have $\mu_f^1(\mathcal{E}\cap R)=\mu_g^1(\mathcal{E}\cap R)$ and $\mu_f^2(\mathcal{E}\cap R)=\mu_g^2(\mathcal{E}\cap R)$, for every oriented rectangle R. We consider the rectangle function

$$\Lambda_f^{\mathcal{E}}(R)=[\{\mu_f^1(\mathcal{E}\cap R)\}^2+\{\mu_f^2(\mathcal{E}\cap R)\}^2+|\mathcal{E}\cap R|^2]^{\frac{1}{2}}.$$

Let $\Psi^{\mathcal{E}}(f)$ be the variation of $\Lambda_f^{\mathcal{E}}(R)$ on I. Then, if $f=g$ on $\mathcal{E}$ it follows that $\Psi^{\mathcal{E}}(f)=\Psi^{\mathcal{E}}(g)$. We show that $\Psi^{\mathcal{E}}(f)=\mu_f(\mathcal{E})$. Let $\{R_n\}$ be the sequence of oriented, open rectangles, which are the components of the complement of $\mathcal{E}$. Then

$$\mu_f(\mathcal{E})=\mu_f(I)-\sum_{n=1}^{\infty}\mu_f(R_n).$$

For every n, let $Q_n=I\sim\bigcup_{k=1}^{n}R_k$. Then $\mu_f(Q_n)$ is the infimum of the variations of $\Psi(R)$ on figures containing the figure Q_n. Since $\mathcal{E}\subset Q_n$, it follows immediately from the definitions that $\mu_f(Q_n)\geqslant\Psi^{\mathcal{E}}(f)$, for every n. Let $\varepsilon>0$, and let Q_n be such that $|Q_n\sim\mathcal{E}|<\frac{1}{3}\varepsilon$, $\mu_f^1(Q_n\sim\mathcal{E})<\frac{1}{3}\varepsilon$ and $\mu_f^2(Q_n\sim\mathcal{E})<\frac{1}{3}\varepsilon$. Then, using the inequality (*), defined in § 5, for every finite set of pair-wise disjoint rectangles, $S_1, S_2, \ldots, S_k$, we have

$$\sum_{j=1}^{k}\Lambda_f^{\mathcal{E}}(S_j)>\sum_{j=1}^{k}\mu_f(S_j\cap Q_n)-\mathcal{E},$$

so that
$$\Psi^{\mathcal{E}}(f) \geq \mu_f(Q_n) - \mathcal{E}.$$

But $\mu_f(Q_n) \geq \mu_f(\mathcal{E})$ so that, for every $\varepsilon > 0$, $\Psi^{\mathcal{E}}(f) \geq \mu_f(\mathcal{E}) - \varepsilon$ and so $\Psi^{\mathcal{E}}(f) \geq \mu_f(\mathcal{E})$. This shows that $\Psi^{\mathcal{E}}(f) = \mu_f(\mathcal{E})$.

Since $\Psi^{\mathcal{E}}(f)$ depends only on the values of f on $\mathcal{E}$, we have the following extension of Verchenko's theorem.

THEOREM 4. *If f and g are BVC and linearly continuous, and if $\mathcal{E}$ is an elementary set such that f and g are continuous on $\mathcal{E}$, relative to $\mathcal{E}$, and if $f = g$ on $\mathcal{E}$, then $\mu_f(\mathcal{E}) = \mu_g(\mathcal{E})$.*

15. We are now ready to prove our main theorem. Let I be the open unit square. If f is a function, and $E \subset I$ is a Borel set, then $H(f, E)$ will designate the Hausdorff 2 dimensional measure of the set of $[(x, f(x)) : x \in E]$. Since this set is analytic it is measurable in the cases being considered.

For every Borel set $E \subset I$ and continuous f of finite area, $\mu_f(E) = H(f, E)$. This is true because $\mu_f(E)$ and $H(f, E)$ are both measures, which agree for open rectangles, and so for all Borel sets. We show that if f is linearly continuous and BVC, then for every $\varepsilon > 0$, there are closed sets $E \subset U$, $F \subset V$, with $|E| > 1 - \varepsilon$, $|F| > 1 - \varepsilon$, and a continuous g, of finite area, such that $f = g$ on $\mathcal{E} = (E \times V) \cup (U \times F)$. For this purpose, we choose E and F so that f is continuous on $\mathcal{E}$, for every $x \in E$, $(f_x)^h$ converges uniformly to f_x and $\varphi(f_x) < \infty$, and for every $y \in F$, $(f_y)^h$ converges uniformly to f_y and $\varphi(f_y) < \infty$.

Let S_n be a closed square in I whose boundary is in $\mathcal{E}$ and has distance less than $1/n$ from the boundary of S_n. Then $S_n - \mathcal{E}$ consists of a countable set of pairwise disjoint open rectangles $R_n = (a_n, b_n) \times (c_n, d_n)$, $n = 1, 2, \ldots$. Let $R = R_n$, for some fixed n. Then f is continuous, as a function of one variable, on the boundary of R, and is of bounded variation there. The functions f^h are continuous everywhere. Moreover, $\lim\limits_{n\to\infty} A(f^h | R) = A(f | R)$. This follows from the fact that $A(f^h | R) \leq A(f | R_h)$, where $R_h = (a_n, b_n + h) \times (c_n, d_n + h)$, that $\lim\limits_{n\to\infty} A(f | R_h) = A(f | R)$, and that $\lim\limits_{n\to\infty} A(f^h | R_{-h}) = A(f | R)$, where $R_{-h} = (a_n, b_n - h) \times (c_n, d_n - h)$. We have used here the additional fact, to be proved two sections hence, that if f is linearly continuous and BVC, then $\mu_f(L) = 0$ for every line L.

Let $\varepsilon > 0$. Choose h so that $A(f^h | R) < A(f | R) + \varepsilon$, and so that

$$\int_{a_n}^{b_n} |f^h(x, c_n) - f(x, c_n)| \, dx, \quad \int_{a_n}^{b_n} |f^h(x, d_n) - f(x, d_n)| \, dx,$$

$$\int_{c_n}^{d_n} |f^h(a_n, y) - f(a_n, y)|\, dy, \quad \text{and} \quad \int_{c_n}^{d_n} |f^h(b_n, y) - f(b_n, y)|\, dy$$

are all less than ε. Thus the boundary of R is spanned by a continuous parametric surfaces σ whose area satisfies $A(\sigma) < A(f \mid R) + 5\varepsilon$.

It is easy to replace σ by a non-parametric g such that $A(g \mid R) < A(f \mid R) + 5\varepsilon$ and $g = f$ on the boundary of R. Indeed, since the surface of minimal area, [15], spanning the curve given by f on the boundary of R, is a non-parametric surface, and since the above holds for every $\varepsilon > 0$, g may be chosen so that

$$\sup\,[|g(x, y)| : (x, y) \in R] = \sup\,[|f(x, y)| : (x, y) \in \text{bdry } R]$$

and so that $A(g \mid R) \leqslant A(f \mid R)$.

By extending f from $\mathcal{E}$ to S_n, by adding the above $g = g_n$ on each R_n, we obtain a function g, defined for all $(x, y) \in S_n$. It is evident that g is continuous and that $A(g) \leqslant A(f)$. Let $\mathcal{E}^n = \mathcal{E} \cup S_n$. Now $\mu_g(\mathcal{E}^n) = H(g, \mathcal{E}^n)$. But, $\mu_f(\mathcal{E}^n) = \mu_g(\mathcal{E}^n)$. It thus follows that $\mu_f(\mathcal{E}) = H(g, \mathcal{E})$, since $\mu_f(\mathcal{E}) = \lim_{n\to\infty} \mu_f(\mathcal{E}^n)$ and $H_g(\mathcal{E}) = \lim_{n\to\infty} H_g(\mathcal{E}^n)$.

Let $\mathcal{E}_n = (E_n \times V) \cup (U \times F_n)$, $n = 1, 2, \ldots$, where $|E_n| > 1 - 1/n$, $|F_n| > 1 - 1/n$, each have the above properties. Suppose, moreover, as we may, that $\mathcal{E}_n \subset \mathcal{E}_{n+1}$, $n = 1, 2, \ldots$. Then $\mu_f(\mathcal{E}_n) = H(f, \mathcal{E}_n)$. Let $\mathcal{E} = \bigcup_{n=1}^{\infty} \mathcal{E}_n$. Then $\mu_f(\mathcal{E}) = \lim_{n\to\infty} \mu_f(\mathcal{E}_n)$, and $H(f, \mathcal{E}) = \lim_{n\to\infty} H(f, \mathcal{E}_n)$. Hence, $\mu_f(\mathcal{E}) = H(f, \mathcal{E})$. Since $I \sim \mathcal{E}$ is negligible, $\mu_f(\mathcal{E}) = A(f)$, so that $A(f) = H(f, \mathcal{E})$. We have thus proved

THEOREM 5. *If f is linearly continuous on I, and BVC, there is a negligible set Z such that the Hausdorff 2 dimensional measure of the graph of f, restricted to $I \sim Z$, equals the area of the surface given by f. Moreover, the same holds for every negligible set containiny Z.*

In the process of proving Theorem 5, we have obtained a general version of the Plateau problem for the non-parametric case. This asserts that among the linearly continuous, non-parametric surfaces which span a continuous curve of finite length, defined over the boundary of a rectangle, there is one of smallest area, and it is given by a continuous function.

Also, from Theorem 5, there follows an extension of Verchenko's theorem. Let f, g be linearly continuous and BVC. Let Z_f and Z_g be the associated negligible sets of the theorem. They are Borel sets. Let E be a Borel set such that $f = g$ on E. Now, $\mu_f(E \sim Z_f \cup Z_g) = \mu_g(E \sim Z_f \cup Z_g)$ since $H(f, E \sim Z_f \cup Z_g) = H(g, E \sim Z_f \cup Z_g)$. Moreover, $\mu_f(Z_f)=\mu_g(Z_g)=0$. Finally, $\mu_f(Z_g \sim Z_f)=\mu_g(Z_f \sim Z_g)=0$. This last follows since, for example, $\mu_f^1(Z_g \sim Z_f)$, being majorized by the integral of a function defined on a set of measure 0, is 0. We thus have

THEOREM 6. *If f and g are linearly continuous and of finite area, and if $f=g$ on a Borel set E, then $\mu_f(E)=\mu_g(E)$.*

16. It remains for us to consider the case where f is linearly continuous and has infinite area. Then $A(f)=\infty$. We show that $H(f,\mathcal{E})=\infty$, whenever $I \sim \mathcal{E}$ is negligible. Suppose $\Phi_1(f)=\infty$. Then, for any given M, there is a finite set of pairwise disjoint rectangles $S_1, S_2, \ldots, S_n$ such that if $S_k=(a_k, b_k)\times(c_k, d_k)$, $k=1, 2, \ldots, n$,

then
$$\sum_{k=1}^{n} \int_{c_k}^{d_k} |f(b_k, y)-f(a_k, y)|\,dy > M.$$

Since
$$H(f, S_k) \geqslant \int_{c_k}^{d_k} |f(b_k, y)-f(a_k, y)|\,dy, \qquad k=1, \ldots, n,$$

it follows that
$$H(f, \mathcal{E}) \geqslant \sum_{k=1}^{n} H(f, S_k) \geqslant M.$$

Hence, $H(f,\mathcal{E})=\infty$. Theorem 5 thus has the

COROLLARY 3. *If f is linearly continuous, then the area of the surface given by f equals the Hausdorff 2 dimensional measure of the graph of f, provided that the values at a certain negligible set are deleted.*

We may also state

COROLLARY 4. *Let f and g be linearly continuous on I and let $G=G_f \cap G_g$. If $E \subset G$ is a Borel set, and $f=g$ on E, then $\mu_f(E)=\mu_g(E)$, finite or infinite.*

The proof, which is now easy to obtain, is omitted.

17. For any measurable f, μ_f is zero for the points in G_f. However, as we have noted, lines may have positive measure. We have also used the fact, now to be proved, that for linearly continuous f which are BVC, μ_f is zero for lines.

We call a set ridé if almost all lines, parallel to the coordinate axes, meet S in countable sets. We shall use the facts that, for every Borel set E,

$$\mu_f^1(E)+\mu_f^2(E)+|E| \geqslant \mu_f(E) \geqslant \max\,[\mu_f^1(E),\ \mu_f^2(E),\ |E|].$$

and that
$$\mu_f^1(E)=\int_0^1 m_{f_y}(E_y)\,dy,\quad \mu_f^2(E)=\int_0^1 m_{f_x}(E_x)\,dx.$$

Now, if S is a ridé, Borel set, and if f is linearly continuous and BVC, then $\mu_f^1(S)=\mu_f^2(S)=|S|=0$, so that $\mu_f(S)=0$. It follows that if f is linearly continuous, then $\mu_f(S)=0$ for every ridé Borel subset of G_f. This is not always true for functions which are not essentially linearly continuous. We conjecture that this property characterizes the essentially linearly continuous functions.

For linearly continuous f, there is a variant of Ψ which also gives the area.

Let f be linearly continuous and BVC on I. For every oriented $R\subset I$, let $\pi_1 R$ and $\pi_2 R$ be the projections of $[(x,y,f(x,y)):(x,y)\in R]$ into the xz and yz planes, respectively. Let

$$\xi_f(R)=[|\pi_1 R|^2+|\pi_2 R|^2+|R|^2]^{\frac{1}{2}}.$$

Then, let

$$\Omega(f)=\sup\sum_{k=1}^{n}\xi_f(R_k)$$

for all finite sets of pair-wise disjoint rectangles $R_1, R_2, \ldots, R_n$. Our assertion is that for linearly continuous f, which are BVC, $\Omega(f)=\Psi(f)=A(f)$. This is obviously false for BVC functions in general since, for the function $f(x,y)=0$, $x\leqslant\frac{1}{2}$, $f(x,y)=1$, $x>\frac{1}{2}$, $\Omega(f)=1$ but $\Psi(f)=2$.

For the linearly continuous case, the proof runs as follows: For every rectangle R, $\Phi_2(R)\geqslant|\pi_2(R)|$. On the other hand, for every $\varepsilon>0$, R contains pair-wise disjoint rectangles $R_1, R_2, \ldots, R_n$ such that $\sum_{k=1}^{n}|\pi_2(R_k)|>\Phi_2(R)-\varepsilon$. Then, for almost all x, f_x is continuous in y, so that there is an $n=n(x)$ for which

$$\sum_{i=1}^{N}|f(x,y_i)-f(x,y_{i-1})|>\varphi(f_x)-\tfrac{1}{2}\varepsilon,$$

where the y_i can be chosen arbitrarily, except that $|y_i-y_{i-1}|<1/n(x)$, $i=1,\ldots,N$. There is an n such that $n>n(x)$ for all x, except possibly for a set of measure less than δ, where δ is chosen so that for every set E, with $|E|>(b-a)-\delta$, it follows that $\int_E\varphi(f_x)\,dx>\Phi_2(R)-\frac{1}{2}\varepsilon$. Consider the rectangles

$$R_i=[a,b]\times[i/n,\ (i+1)/n],\ i=1,2,\ldots,n.$$

Then

$$\sum_{i=1}^{n}|\pi_2(R_i)|>\Phi_2(R)-\varepsilon.$$

Similarly, there is a set $S_1, S_2, \ldots, S_k$ of pair-wise disjoint rectangles in R, such that $\sum_{j=1}^{k}|\pi_1(S_j)|>\Phi_1(R)-\varepsilon$. Now, for a refinement of both the R_i and S_j, say Q_k, we have

$$\sum_{k=1}^{m} |\pi_1 Q_k)| > \Phi_1(R) - \varepsilon, \text{ and } \sum_{k=1}^{m} |\pi_2(Q_k)| > \Phi_2(R) - \varepsilon,$$

so that
$$\sum_{k=1}^{m} \xi_f(Q_k) > \Phi(R) - 2\varepsilon.$$

It follows from these inequalities that $\Omega(f) \geqslant \Psi(f)$. But since, for every R, $\Phi_i(R) \geqslant |\pi_i(R)|$, $i = 1, 2$, it is evident that $\Omega(f) \leqslant \Psi(f)$.

The converse is also true. If f is not essentially linearly continuous, and is BVC, then $\Omega(f) < \Psi(f)$. Let us call $\Omega_1(f)$ the upper Burkill integral of the rectangle function $|\pi_1(R)|$, with a similar meaning for $\Omega_2(f)$. Now,

$$\Psi(f) - \Omega(f) \geqslant \max [\Psi_1(f) - \Omega_1(f), \Psi_2(f) - \Omega_2(f)].$$

Suppose, now, that for a set E of values of x, of positive measure, there is a $y(x)$ such that $m_x(\{y(x)\}) > 0$. It then follows easily that for every subdivision $R_1, R_2, \ldots, R_n$

$$\sum_{k=1}^{n} |\pi_1(R_k)| \leqslant \Phi_2(f) - \int_E m_x(\{y(x)\})\, dx.$$

If the above were not true for a set of values of x of positive measure, then it would be true for a set of values of y of positive measure. We have thus proved

THEOREM 7. *If f is BVC, then $\Omega(f) = \Psi(f)$ if and only if f is equivalent to a linearly continuous g.*

We wish to remark that even if f is not BVC, then the linear continuity of f implies that $\Omega(f) = \Psi(f) = \infty$. However, Theorem 6 no longer holds, since $\Omega(f)$ may be infinite for an f which is not essentially linearly continuous.

18. Lest there be some misunderstanding regarding the scope of generality of the linearly continuous functions, we now discuss this matter. We first define a pole p of a function f as a point $p = (x_0, y_0)$ such that, for every M, there is a disk σ, center p, such that, for every $q \in \sigma$, $|f(q)| > M$.

Let R be an oriented square, with center (x_0, y_0), and S the concentric oriented square whose area is half that of R. We call the quasi-linear function which is 0 on $I \sim R$, k on S, and linear on lines joining bounding lines of R to the corresponding bounding lines of S, a *spine* function of center (x_0, y_0) and height k.

LEMMA 6. *If p is quasi-linear, ther for every (x_0, y_0), $k > 0$, and $\varepsilon > 0$, there is a spine function q, of center (x_0, y_0) and height k, such that $E(p+q) < E(p) + \varepsilon$.*

The proof is by elementary geometry, and is left to the reader.

Now, for every n, let (x_i^n, y_i^n), $i=1, 2, \ldots, i_n$, be a finite set of points in I such that every $(x, y) \in I$ has distance less than $1/n$ from at least one of them.

For every $i=1, 2, \ldots, i_1$, let p_i^1 be a spine function of center (x_i^1, y_i^1) and height 1, such that $E\left(\sum_{i=1}^{n_1} p_i^1\right) < \frac{3}{2}$. Let $p_1 = \sum_{i=1}^{n_1} p_i^1$. For every $i=i, 2, \ldots, i_2$, let p_i^2 be a spine function of center (x_i^2, y_i^2) and height 2, such that $E\left(p_1 + \sum_{i=1}^{n_2} p_i^2\right) < \frac{7}{4}$. Let $p_2 = \sum_{i=1}^{n_2} p_i^2$. Continue ad infinitum, so that, for every k, $E\,(p_1 + p_2 + \cdots + p_{k-1} + \sum_{i=1}^{n_k} p_i^k) < 2 - 2^{-k}$.

Moreover, the spine functions may be chosen so that, for every k, the sum of the perimeters of the R_i^k is less than 2^{-k}.

Let $q_n = p_1 + p_2 + \cdots + p_n$. for every n. Then, $E\,(q_n) < 2$, for every n. Since for every x, except for a set of measure less than 2^{-n}, $p_n = 0$, and similarly for y, it follows that the sequence $\{q_n\}$ converges linearly to a function f. f is linearly continuous. Moreover, $A\,(f) \leqslant \liminf_{n\to\infty} E\,(q_n)$, so that $A\,(f) \leqslant 2$.

Let T_i^k be the interior of S_i^k, and let $G_k = \bigcup_{i=1}^{n_k} T_i^k$. Now, let $H_k = \bigcup_{n=k}^{\infty} G_n$. For every k, H_k is an everywhere dense open set. Hence, the complement of $H = \bigcap_{k=1}^{\infty} H_k$ is of the first category. Let $(x_0, y_0) \in H$. Then, for every k, there is a circle σ, center (x_0, y_0), which is contained in H_k, so that $|f(x, y)| \geqslant k$ on σ. Hence H consists of poles of f. We have thus proved:

THEOREM 8. *If f is linearly continuous, and of finite area, then the complement of its set of poles can be of the first category.*

References

[1]. A. S. BESICOVITCH, On the definition and value of the area of a surface. *Quart. J. Math.*, 16 (1945), 86–102.

[2]. L. CESARI, Sulle funzioni a variazione limitata. *Ann. Scuola Norm. Sup. Pisa*, (2), 5 (1936), 299–313.

[3]. ——, *Surface Area*. Princeton, 1956.

[4]. H. FEDERER, The (φ, k) rectifiable subsets of n space. *Trans. Amer. Math. Soc.*, 62 (1947), 114–192.

[5]. C. GOFFMAN, Lower semi-continuity of area functionals, I. The non-parametric case. *Rend. Circ. Mat. Palermo*, (2), 2 (1954), 203–235.

[6]. ——, Lower semi-continuity of area functionals, II. The Banach area. *Amer. J. Math.*, 76 (1954), 679–688.

[7]. ——, Convergence in area of integral means. *Amer. J. Math.*, 77 (1955), 563–574.

[8]. K. KRICKEBERG, Distributionen, Funktionen beschränkter Variation und Lebesguescher Inhalt nichtparametrischer Flächen. *Ann. Mat. Pura Appl.*, IV 44 (1957), 105–133.

[9]. E. J. MICKLE & T. RADO, On a theorem of Besicovitch in surface area theory. *Revista Mat. Univ. Parma*, 2 (1951), 19–45.

10]. I. NATANSON, *Theorie der Funktionen einer reellen Veränderlichen.* Berlin, Akad. Verlag, 1954.

[11]. C. Y. PAUC, Considerations sur les gradients généralises de G. Fichera et E. de Giorgi. *Ann. Mat. Pura Appl.*, IV 44 (1957), 135–152.

[12]. T. RADO, Length and Area. *Amer. Math. Soc. Coll. Publ.*, 30, 1948.

[13]. ——, On the problem of Plateau. *Ergebn. der Math. und ihrer Grenzgeb.*, 2, 1933.

[14]. I. VERCHENKO, Über das Flächenmass von Mengen. *Mat. Sbornik* (N. S.), 10, 52 (1942), 11–32. (Russian, german summary).

[15]. L. DE VITO, Sulle funzioni di pui variabili a variazione limitata. *Arch. Rat. Mech. and Analysis*, 3, No. 1 (1959), 60–81.

Received July 30, 1959.