

THE ISOPERIMETRIC INEQUALITY AND THE LEBESGUE DEFINITION OF SURFACE AREA

BY
TIBOR RADÓ

1. **Introduction.** 1.1. If C is a closed curve in a Euclidean plane E_2 , and $l(C)$ and $a(C)$ denote the length of C and the area enclosed by C respectively, then we have *the plane isoperimetric inequality*

$$(1) \quad a(C) \leq l(C)^2/4\pi.$$

Similarly, if S is a closed surface in Euclidean three-space E_3 , and $A(S)$ and $V(S)$ denote the area of S and the volume enclosed by S respectively, then we have *the spatial isoperimetric inequality*

$$(2) \quad V(S)^2 \leq A(S)^3/36\pi.$$

The literature of these classical isoperimetric inequalities is very extensive (comprehensive presentations may be found in Blaschke [1] and Bonnesen [1]⁽¹⁾). In most instances, only convex curves and convex surfaces are considered, or else it is assumed that the curves and surfaces involved are sufficiently regular to permit the use of the classical formulas for the quantities $a(C)$, $l(C)$, $A(S)$, $V(S)$. Briefly, the greater part of the literature relates to what may be termed *the elementary range*. Within the elementary range, the concepts involved in the inequalities (1) and (2) have generally accepted meanings, and the validity of these inequalities is a foregone conclusion, even though the actual proofs are of great interest and of substantial difficulty. On the other hand, the situation is quite different beyond the elementary range, especially in the case of the spatial isoperimetric inequality (2). It is well known that the number of formal definitions that have been proposed for *surface area* is very large. It is perhaps less well known that most of the more relevant definitions of surface area were found to conflict with each other in relatively simple non-elementary cases (see, for example, Nöbeling [1]). Similarly the concept of *enclosed volume*, involved in the inequality (2), admits of several plausible formal definitions which are readily seen to conflict with each other beyond the elementary range (cf. 1.4, 5.6, 5.8). Finally, the concept of *closed surface* lends itself to several fundamentally different interpretations (cf. Youngs [1, 2]). Consequently, beyond the elementary range the spatial isoperimetric inequality (2) is by no means an a priori obvious geometrical fact. Rather, this inequality may be construed as a *test of adequate*

Presented to the Society, August 22 and 23, 1946; received by the editors August 13, 1946.

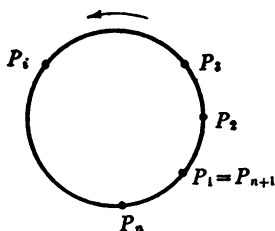
(¹) Numbers in brackets refer to the bibliography at the end of the paper.

adjustment between the concepts referred to by the symbols S , $V(S)$, $A(S)$ in the inequality (2). Examples show that entirely plausible formal definitions of S , $V(S)$, $A(S)$ may turn out to be incompatible with each other if tested by means of the spatial isoperimetric inequality (cf. 5.6).

The purpose of this paper is to study the Lebesgue area $A_L(S)$ (see, for instance, Youngs [2] and Huskey [1]) from the point of view of *fitness relative to the spatial isoperimetric inequality* (2). The Lebesgue area is of particular interest in this respect for several reasons. In the first place, examples due to Geöcze [1] and Besicovitch [1] reveal that the Lebesgue area has a tendency of yielding, beyond the elementary range, smaller values than one may expect on the basis of comparison with other area-definitions. Hence, other things being equal, the Lebesgue area may be expected to yield a sharper isoperimetric inequality. The results of this paper tend to confirm this expectation. A striking illustration is obtained by means of an example, due to Besicovitch, where $A_L(S) < \infty$, while any area $A_B(S)$ of a type preferred by Besicovitch himself is infinite (see 5.5). In the second place, the examples found by Geöcze and Besicovitch show that apparently quite natural definitions of the enclosed volume $V(S)$ may fail to pass the isoperimetric test in relation to the Lebesgue area $A_L(S)$ (see 5.4, 5.5). Thus the selection of a suitable definition of $V(S)$ is a relevant issue. The concept of enclosed volume $V(S)$ used in this paper is based on topological considerations (see 1.4), and seems to clear up certain paradoxical phenomena relative to "pathological" surfaces of unexpectedly small Lebesgue area. In the third place, in view of the examples of Geöcze and Besicovitch it would seem that previous discussions of the spatial isoperimetric inequality, in terms of the Lebesgue area $A_L(S)$, by Tonelli [1], Blaschke [1], and Bonnesen [1], do not take into account certain curious possibilities. Thus it seems worthwhile to re-examine the spatial isoperimetric inequality. The results obtained lead, by analogy, to simplifications and improvements concerning the more familiar plane isoperimetric inequality (1). For this reason, and also for purposes of comparison, we included a brief discussion of the inequality (1) also. We shall presently summarize the principal definitions and results of this paper.

1.2. As regards the concept of a *closed curve*, we shall use oriented Fréchet curves of the type of the 1-sphere in a given Euclidean xy -plane (see, for instance, Youngs [1, 2] for the formal definition of such curves). Since no other type of curve will be considered in this paper, we shall use the term *oriented closed curve* to refer to curves of this type. A plane oriented closed curve C is determined by a representation of the form $C: x = x(P)$, $y = y(P)$, $P \in \Gamma$, where Γ is the unit circle $u^2 + v^2 = 1$ in an auxiliary uv -plane, and $x(P)$, $y(P)$ are continuous functions on Γ . The orientation of C is determined by selecting one of the two possible orientation of Γ . For definiteness, we agree to choose the counterclockwise orientations on Γ . If we set $u = \cos \theta$, $v = \sin \theta$, then the representation of C appears in the form $C: x = x(\theta)$, $y = y(\theta)$, $0 \leq \theta \leq 2\pi$,

$x(0) = x(2\pi)$, $y(0) = y(2\pi)$, where $x(\theta)$, $y(\theta)$ are continuous functions of θ . The points (x, y) that correspond to the points P of Γ by means of the equations of C form a point-set that we shall denote by $[C]$. While the point P of Γ describes Γ in the counterclockwise sense, the corresponding point $[x(P), y(P)]$ travels over the set $[C]$, where certain points of $[C]$ may be crossed several times. Intuitively, C is thought of *not* as a point-set but rather as a *trip over a point-set*. Thus C and $[C]$ are distinct entities. The length $l(C)$ is defined, in



the usual sense, in terms of inscribed polygons, as follows. Let σ be a generic notation for a finite system of points $P_1, P_2, \dots, P_n, P_{n+1} = P_1$ of Γ that follow upon each other in the counterclockwise sense (see figure), and let (x_i, y_i) be the point that corresponds to the point P_i by means of the equations of C . Let us put

$$l_\sigma = \sum_{i=1}^n [(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]^{1/2}.$$

Then $l(C)$ is defined as the least upper bound of l_σ for all possible choices of σ . Thus $l(C)$ may be infinite. If $l(C) < \infty$, then C is termed rectifiable.

1.3. *Continuation.* The area enclosed by C has been defined in previous literature (see Blaschke [1]), in connection with the plane isoperimetric inequality, as a *signed* (or algebraic) area. Assuming that C is rectifiable, the signed enclosed area $a_s(C)$ is given by the formula

$$(3) \quad a_s(C) = 2^{-1} \int_{\Gamma} (x dy - y dx),$$

where the existence of the Stieltjes integrals involved follows from the assumption that C is rectifiable. The plane isoperimetric inequality (1) is then established in the form (see Blaschke [1])

$$(4) \quad |a_s(C)| \leq l(C)^2 / 4\pi.$$

Inspection of a lemniscate, for example, reveals that due to possible cancellations between the contributions of the "loops" of C the inequality (4) may yield an unduly weak statement. We shall establish a (generally) stronger inequality, using the topological index (see, for instance, Radó [3]). Let first (x, y) be a point that does not lie on the set $[C]$ (cf. 1.2). We define then the

index function $i(x, y)$ as the topological index of the point (x, y) relative to the oriented closed curve C . Intuitively, $i(x, y)$ indicates how many times C encircles the point (x, y) . If (x, y) lies on $[C]$, then we set $i(x, y) = 0$. The index-function $i(x, y)$ takes on only integral values. It is constant on each component of the complement of the set $[C]$, and in particular it vanishes on the unbounded component of the complement of $[C]$. Assume now that C is rectifiable. We have the formula (see Radó [3])

$$(5) \quad \iint i(x, y) dx dy = 2^{-1} \int_{\Gamma} (x dy - y dx),$$

where the double integral is extended over the whole xy -plane, and the integral is a Lebesgue integral. Since $i(x, y)$ vanishes outside of some sufficiently large rectangle, the range of integration may be taken as such a rectangle. By (3) and (5) we have for the signed area $a_s(C)$ the formula

$$(6) \quad a_s(C) = \iint i(x, y) dx dy \quad \text{if } l(C) < \infty.$$

We introduce now a quantity $a(C)$ defined as follows:

$$(7) \quad a(C) = \begin{cases} \iint |i(x, y)| dx dy & \text{if } i(x, y) \text{ summable,} \\ + \infty & \text{otherwise.} \end{cases}$$

The quantity $a(C)$ may be considered as *the absolute area* enclosed by C . We shall prove that the plane isoperimetric inequality (1) holds if $a(C)$ is defined by (7). Since generally $|a_s(C)| \leq a(C)$, this is an improvement over the inequality (4). The writer is not aware of any previous proof of this plausible result. As a matter of fact, our method of proof suggests the possibility of further improvements (cf. 5.8).

We noted above that C and $[C]$ are different entities. In fact, C , $l(C)$, $a(C)$ are not determined by the point-set $[C]$ alone. To illustrate this important point, let us consider the oriented closed curves C_1 , C_2 , C_3 given as follows.

$$\begin{array}{lll} C_1: & x = \cos \theta, & y = \sin \theta, & 0 \leq \theta \leq 2\pi, \\ C_2: & x = \cos 2\theta, & y = \sin 2\theta, & 0 \leq \theta \leq 2\pi, \\ C_3: & \begin{cases} x = \cos 2\theta, & y = \sin 2\theta, \\ x = \cos 2\theta, & y = -\sin 2\theta, \end{cases} & & \begin{array}{l} \text{for } 0 \leq \theta \leq \pi, \\ \text{for } \pi \leq \theta \leq 2\pi. \end{array} \end{array}$$

Then clearly the sets $[C_1]$, $[C_2]$, $[C_3]$ are identical, and coincide with the unit circle $x^2 + y^2 = 1$. On the other hand, the curves C_1 , C_2 , C_3 are entirely different "trips" over the unit circle. Clearly $l(C_1) = 2\pi$, $l(C_2) = 4\pi$, $l(C_3) = 4\pi$, $a(C_1) = \pi$, $a(C_2) = 2\pi$, $a(C_3) = 0$ (cf. (7)).

From the formal definition of the curves used here (cf. Youngs [2]) it follows readily that $[C]$, $l(C)$, and $a(C)$ are independent of the particular representation chosen for C .

1.4. Turning now to the spatial isoperimetric inequality (2), the main topic of this paper, we shall use the term *oriented closed surface* in the sense of oriented Fréchet surface of the type of the 2-sphere (see, for instance, Youngs [1, 2]). If U denotes the positively oriented unit sphere $u^2 + v^2 + w^2 = 1$ in an auxiliary Euclidean uvw -space, then an oriented closed surface S is determined by a representation

$$S: \quad x = x(p), \quad y = y(p), \quad z = z(p), \quad p \in U,$$

where $x(p)$, $y(p)$, $z(p)$ are continuous functions on U . The points (x, y, z) that are obtained by means of this representation form a point-set $[S]$ in Euclidean xyz -space. It is *not* assumed that the correspondence between the points p of U and the points (x, y, z) of $[S]$ is biunique. As regards *surface area*, we shall use the Lebesgue area $A_L(S)$ (see, for instance, Youngs [1, 2]). The *enclosed volume* $V(S)$ will now be defined in complete analogy with the formula (7), by the formula

$$(8) \quad V(S) = \begin{cases} \iiint |i(x, y, z)| \, dx dy dz & \text{if } i(x, y, z) \text{ summable,} \\ + \infty & \text{otherwise,} \end{cases}$$

where the triple integral is extended over the whole xyz -space, and the index-function $i(x, y, z)$ is defined as follows. Let (x_0, y_0, z_0) be any point. If (x_0, y_0, z_0) lies on the set $[S]$, then we put $i(x_0, y_0, z_0) = 0$. If $(x_0, y_0, z_0) \notin [S]$, then $i(x_0, y_0, z_0)$ is defined as the degree of a certain continuous mapping, obtained in the following manner. On introducing the vector-function $\mathfrak{x}(p)$ whose components are the functions $x(p)$, $y(p)$, $z(p)$ occurring in the given representation of S , this representation may be written in the more concise form $S: \mathfrak{x} = \mathfrak{x}(p)$, $p \in U$. Let us denote by \mathfrak{x}_0 the vector with components x_0, y_0, z_0 , and by $|\mathfrak{x}(p) - \mathfrak{x}_0|$ the length of the vector $\mathfrak{x}(p) - \mathfrak{x}_0$. Since $(x_0, y_0, z_0) \notin [S]$, we have $|\mathfrak{x}(p) - \mathfrak{x}_0| \neq 0$ for $p \in U$. The formula

$$\mathfrak{x} = \mathfrak{x}_0 + \frac{\mathfrak{x}(p) - \mathfrak{x}_0}{|\mathfrak{x}(p) - \mathfrak{x}_0|}, \quad p \in U,$$

defines then a continuous (generally not biunique) mapping from U onto a subset of the unit sphere with center at (x_0, y_0, z_0) , and the index $i(x_0, y_0, z_0)$ is defined as the degree of this mapping (see Alexandroff-Hopf [1]). From well known theorems on the degree of a mapping, it follows that $i(x, y, z)$ takes on only integral values (which may be positive, negative, or zero). On each component of the complement of the set $[S]$, $i(x, y, z)$ is constant. In particular, $i(x, y, z)$ vanishes on the unbounded component of the complement of

[S]. Using the definition (8) for the enclosed volume $V(S)$, we shall establish the isoperimetric inequality (2) in the form $V(S)^2 \leq A_L(S)^3/36\pi$, for every oriented closed surface S , the latter term being used in the sense explained at the beginning of the present §1.4.

In analogy with the formula (6), one may want to use the signed (or algebraic) enclosed volume $V_*(S)$ defined by the formula

$$V_*(S) = \iiint i(x, y, z) dx dy dz,$$

provided that the index-function $i(x, y, z)$ is summable. As a matter of fact, the brief remarks of Blaschke [1] seem to indicate that he had $V_*(S)$ in mind in connection with the spatial isoperimetric inequality (2). Since clearly $|V_*(S)| \leq V(S)$, our result is thus generally stronger than the result suggested by the comments of Blaschke. As regards the treatment of the spatial isoperimetric inequality (2) by Tonelli [1] and Bonnesen [1], comparisons are difficult due to the lack of an explicit definition of the concepts closed surface and enclosed volume in their work. However, it would seem that the method of proof used by these authors may fail to apply in certain cases (cf. 5.7).

Questions of considerable interest and difficulty arise if one attempts to compare our result, from the point of view of generality and strength, with results based on concepts of closed surface, surface area, and enclosed volume different from the concepts used in this paper (see, for instance, W. Gross [1]). A study of these questions is beyond the scope of this paper.

Remarks analogous to those in 1.3 show that the quantities $V(S)$ and $A_L(S)$, as well as the point-set $[S]$, are independent of the choice of a particular representation for S , and that the surface S and the point-set $[S]$ must be considered as entirely different geometrical objects. In particular, *the enclosed volume $V(S)$ and the surface area $A_L(S)$ are not determined by the point-set $[S]$ alone*. In fact, the point-set $[S]$, taken by itself, does not yield generally any information concerning $V(S)$ and $A_L(S)$ (cf. 5.4).

1.5. The general results stated in 1.3 and 1.4 will be derived by appropriate limit processes from elementary inequalities that we shall describe presently. In a Euclidean plane E_2 , let s_1, s_2, \dots, s_m be a finite system of (not necessarily distinct) straight segments. We put

$$(9) \quad l_* = l(s_1) + l(s_2) + \dots + l(s_m),$$

where $l(s_i)$ denotes the length of the segment s_i , $i = 1, 2, \dots, m$. Consider the set

$$(10) \quad F = s_1 + s_2 + \dots + s_m.$$

A straight segment s with end points p, q , located in the plane E_2 , will be termed admissible (relative to the system s_1, s_2, \dots, s_m) if (i) $p \notin F, q \notin F$ and (ii) s contains none of the end points of the segments s_1, s_2, \dots, s_m .

The set F is bounded and closed. The complement of F , relative to the plane E_2 , is an open set which has a finite number of components, one of which is unbounded. This unbounded component will be denoted by D_0 , while the bounded components, if present, will be denoted by D_1, \dots, D_n . If the bounded components are missing, then the inequality to be stated below is trivially true, and so we can assume that at least one bounded component is present.

Now let us take any function $\mu(p)$, defined for all points $p \in E_2$, subject to the following restrictions. (i) $\mu(p)$ is non-negative and takes on only integral values. (ii) $\mu(p) = 0$ for $p \in F$. (iii) $\mu(p)$ is constant on each one of the domains D_0, D_1, \dots, D_n . (iv) $\mu(p) = 0$ for $p \in D_0$. (v) If p, q are any two distinct points in E_2 , such that the straight segment with end points p, q is admissible with respect to the system s_1, s_2, \dots, s_m , then $|\mu(p) - \mu(q)|$ does not exceed the number of those of the segments s_1, s_2, \dots, s_m that are intersected by the segment pq .

Such a function $\mu(p)$ having been chosen, let us denote its constant value on the domain D_j by μ_j (note that $\mu_0 = 0$), and let us put

$$(11) \quad a_* = \sum_{j=1}^n \mu_j |D_j|,$$

where $|D_j|$ is the (two-dimensional) measure of the domain D_j , $j = 1, \dots, n$. If the bounded domains D_j are missing, we put $a_* = 0$. We shall establish in §2 the inequality

$$(12) \quad a_* \leq l_*^2/4\pi,$$

which will be shown, in §4, to yield the plane isoperimetric inequality (1) (in the strong form described in 1.3) by means of an immediate passage to the limit.

It should be noted that a_* depends not only upon the system s_1, s_2, \dots, s_m but also upon the choice of the function $\mu(p)$. For example, we may choose $\mu(p) \equiv 0$. Then $a_* = 0$, and the inequality (12) is trivial. However, we can choose $\mu(p) = 1$ for $p \in D_1 + \dots + D_n$, this choice being clearly compatible with the requirements (i)–(v) stated above. Then (12) yields the inequality $|D_1| + \dots + |D_n| \leq l_*^2/4\pi$. The inequality (12), in its general form, states that each one of the bounded domains D_j may be taken with a multiplicity μ_j , so long as the choice of the non-negative integers μ_j does not violate the requirement (v) stated above (cf. 5.8).

1.6. An entirely analogous elementary inequality will be used to derive, by means of a passage to the limit, the spatial isoperimetric inequality (2). In Euclidean three-space E_3 , let there be given a finite system of (not necessarily distinct) plane rectilinear triangles $\Delta_1, \Delta_2, \dots, \Delta_m$. We put

$$(13) \quad A_* = A(\Delta_1) + A(\Delta_2) + \dots + A(\Delta_m),$$

where $A(\Delta_i)$ denotes the area (in the elementary sense) of the triangle Δ_i , $i=1, 2, \dots, m$. Consider the set

$$(14) \quad F = \Delta_1 + \Delta_2 + \dots + \Delta_m.$$

A straight segment s , with end points p, q , will be termed admissible (relative to the system $\Delta_1, \Delta_2, \dots, \Delta_m$) if (i) $p \notin F, q \notin F$, and (ii) s does not intersect the perimeter of any one of the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$.

The set F , defined by (14), is bounded and closed. The complement of F is an open set which has a finite number of components, one of which is unbounded. This unbounded component will be denoted by D_0 , while the bounded components (if present) will be denoted by D_1, \dots, D_n . If the bounded components are missing, then the inequality to be stated below is trivially true, and so we can assume that at least one bounded component is present.

Now let us take any function $\mu(p)$, defined for all points $p \in E_3$, subject to the following restrictions. (i) $\mu(p)$ is non-negative and takes on only integral values. (ii) $\mu(p) = 0$ for $p \in F$ (see (14)). (iii) $\mu(p)$ is constant on each one of the domains D_0, D_1, \dots, D_n . (iv) $\mu(p) = 0$ for $p \in D_0$. (v) If p, q are any two points, such that the straight segment s with end points p, q is admissible with respect to the system $\Delta_1, \Delta_2, \dots, \Delta_m$, then $|\mu(p) - \mu(q)|$ does not exceed the number of those of the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$ that are intersected by the segment s .

Such a function $\mu(p)$ having been chosen, let us denote by μ_j its constant value on the domain D_j (note that $\mu_0 = 0$), and let us put

$$(15) \quad V_* = \sum_{j=1}^n \mu_j |D_j|,$$

where $|D_j|$ denotes the (three-dimensional) measure of the domain D_j , $j=1, \dots, n$. If the bounded domains D_j are missing, we put $V_* = 0$. We shall establish, in §3, the inequality

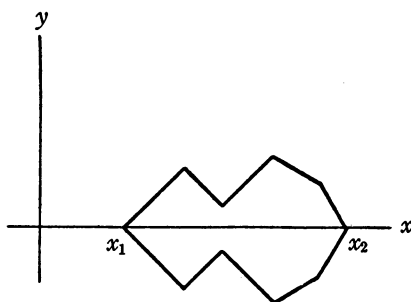
$$(16) \quad V_*^2 \leq A_*^3/36\pi,$$

which will be shown in §4 to yield, by an immediate passage to the limit, the spatial isoperimetric inequality (2) in the strong form described in 1.4.

The inequality (16) gives rise to remarks analogous to those made at the end of 1.5. If the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$ form a simple closed polyhedron, then we may choose $\mu(p) = 1$ on the interior of the polyhedron, and (16) yields the classical spatial isoperimetric inequality for this special case. The proofs given in the literature for this special case are all based in the last analysis upon processes of symmetrization first used by Steiner and H. A. Schwarz (see Bonnesen [1]), and our proof of the general inequality (16) reveals that these processes of symmetrization can be made to yield stronger conclusions than those found in previous literature. Similar remarks apply to (12).

We observed above that the spatial isoperimetric inequality may be construed as a test of fitness for the Lebesgue definition of surface area. In view of the fact that this definition of surface area has also been used in the solution of another classical variation problem, namely, the problem of Plateau (for literature, see Radó [2]), it appears that the Lebesgue area $A_L(S)$ proved a valuable tool in the study of two outstanding classical variation problems.

1.7. This paper is subdivided as follows. §§2 and 3 contain the proofs of the elementary inequalities (12) and (16) respectively. §4 is concerned with the passage to the limit that leads to the general isoperimetric inequalities. §5 describes various interesting special cases covered by our general results, and contains a brief discussion of certain gaps in previous literature.



2. Elementary inequalities in the plane. 2.1. In Euclidean xy -plane, let us consider a simply-connected polygonal region that is symmetric with respect to the x -axis (see figure). If l and a denote the length of the perimeter and the enclosed area, these terms being used in the elementary sense, then the isoperimetric inequality $a \leq l^2/4\pi$ can be established for this special case in a very simple and elementary manner (see Blaschke [1]). This special result being assumed, let $y = \eta(x)$, $x_1 \leq x \leq x_2$, be the equation of the upper half of the perimeter. Then the inequality $a \leq l^2/4\pi$ (assumed to hold in this elementary case) yields the inequality

$$(17) \quad \int_{x_1}^{x_2} \eta(x) dx \leq \frac{1}{2\pi} \left[\int_{x_1}^{x_2} (1 + \eta'(x)^2)^{1/2} dx \right]^2,$$

which is thus seen to hold for every continuous, piecewise linear function $\eta(x)$ such that $\eta(x_1) = \eta(x_2) = 0$ and $\eta(x) > 0$ for $x_1 < x < x_2$.

2.2. Now let $\psi(x)$ be a function that is continuous and piecewise linear in an interval $-K \leq x \leq K$ and satisfies the following conditions. (i) $\psi(-K) = \psi(K) = 0$. (ii) $\psi(x) \geq 0$ for $-K \leq x \leq K$. Let then E denote the set of those points in the interval $-K \leq x \leq K$ where $\psi(x) > 0$. We assert the inequality

$$(18) \quad \int_{-K}^K \psi(x) dx \leq \frac{1}{2\pi} \left[\int_E (1 + \psi'(x)^2)^{1/2} dx \right]^2.$$

Indeed, this inequality is obvious if $\psi(x) \equiv 0$. If $\psi(x) \not\equiv 0$, then the interval $-K \leq x \leq K$ can be subdivided into a finite number of intervals i_1, \dots, i_n , such that for each $k=1, \dots, n$ the function $\psi(x)$ is either identically zero in i_k , or else $\psi(x)$ is positive in the interior of i_k and vanishes at the end points of i_k . Let us define $\epsilon_k = 0$ if $\psi(x) \equiv 0$ in i_k and $\epsilon_k = 1$ if $\psi(x) > 0$ in the interior of i_k , and let us put $I_k = \int_{i_k} \psi(x) dx$, $J_k = \epsilon_k \int_{i_k} [1 + \psi'(x)^2]^{1/2} dx$. We have then, for $k=1, \dots, n$,

$$(19) \quad I_k \leq \frac{1}{2\pi} J_k^2.$$

Indeed, (19) is obvious if $I_k = 0$. If $I_k > 0$, then $\epsilon_k = 1$, and (19) follows from (17). From (19) we infer that

$$\int_{-K}^K \psi(x) dx = \sum_{k=1}^n I_k \leq \frac{1}{2\pi} \sum_{k=1}^n J_k^2 \leq \frac{1}{2\pi} \left[\sum_{k=1}^n J_k \right]^2.$$

Now clearly

$$\sum_{k=1}^n J_k = \int_E [1 + \psi'(x)^2]^{1/2} dx,$$

and the inequality (18) follows.

2.3. We proceed to prove the inequality stated in 1.5. Given a system s_1, s_2, \dots, s_m and a function $\mu(p)$ as described in 1.5, we choose a Cartesian coordinate system xy in the plane E_2 in such a manner that none of the segments s_1, s_2, \dots, s_m is parallel to the y -axis. The function $\mu(p)$ may then be denoted by $\mu(x, y)$. The quantity a_* occurring in (11) may be written now in the form

$$(20) \quad a_* = \int_{-K}^K \int_{-K}^K \mu(x, y) dx dy,$$

where $K > 0$ is so large that the set F (see (10)) is comprised in the interior of the square $-K \leq x \leq K$, $-K \leq y \leq K$. On introducing the function

$$(21) \quad q(x) = \int_{-K}^K \mu(x, y) dy, \quad -K \leq x \leq K,$$

we have the formula

$$(22) \quad a_* = \int_{-K}^K q(x) dx.$$

On setting

$$(23) \quad \psi(x) = q(x)/2,$$

the inequality (12) appears in the form

$$(24) \quad \int_{-K}^K \psi(x) dx \leq l_*^2/8\pi.$$

We shall verify presently this last inequality.

2.4. We assert that the function $q(x)$ (see (21)) is continuous in the interval $-K \leq x \leq K$. Indeed, let x_n be a sequence of points in this interval that converge to a point x_0 . The segment

$$(25) \quad \sigma_0: x = x_0, \quad -K \leq y \leq K,$$

intersects the set F (see (10)) at most in a finite number of points, since none of the segments s_1, s_2, \dots, s_m is parallel to the y -axis. If (x_0, y) is an interior point of the segment σ_0 that does not lie in F , then (x_0, y) lies in one of the domains D_0, D_1, \dots, D_n (see 1.5), and hence $\mu(x, y)$ is constant in a two-dimensional neighborhood of the point (x_0, y) . Hence clearly

$$(26) \quad \mu(x_n, y) \rightarrow \mu(x_0, y), \quad -K \leq y \leq K,$$

except perhaps for a finite set of y -values. Since the function $\mu(x, y)$ is bounded, termwise integration is permissible in (26), and we obtain the relation $q(x_n) \rightarrow q(x_0)$ which establishes the continuity of $q(x)$.

In view of the special character of the function $\mu(x, y)$ it is clear that $q(x)$ is piecewise linear in the interval $-K \leq x \leq K$. Since obviously $\mu(-K, y) \equiv 0$, $\mu(K, y) \equiv 0$, we have $q(-K) = q(K) = 0$. Since $\mu(x, y) \geq 0$, it is clear that $q(x)$ is non-negative.

2.5. *Continuation.* Clearly, we have a subdivision of the interval $-K \leq x \leq K$ into a finite number of intervals, such that the following holds for each interval I of the subdivision.

(i) No interior point of I is the projection of an end point of any of the segments s_1, s_2, \dots, s_m .

(ii) $q(x)$ is linear in I .

(iii) Either $q(x) \equiv 0$ or $q(x) > 0$ in the interior of I . In the first case we shall say that I is an *interval of the first kind*, while in the second case I will be termed an *interval of the second kind*.

Intervals of the second kind may be missing altogether. If this happens, then $q(x) \equiv 0$ and hence $a_* = 0$ (see (22)), and thus the inequality (12) is trivial in this case. So we can assume that intervals of the second kind are actually present.

2.6. *Continuation.* Let $I: x_1 \leq x \leq x_2$ be an interval of the second kind. For given k , the projection upon the x -axis of the segment s_k either has no interior point in common with I or else it contains the whole interval I (cf. (i) in 2.5). Let $\mathfrak{R}(I)$ be the set of those subscripts k for which the second alternative holds. For $k \in \mathfrak{R}(I)$, let $s_k(I)$ denote the portion of s_k that lies above I , and let $l[s_k(I)]$ be the length of the segment $s_k(I)$. If \sum^* denotes summation relative to all the intervals I of the second kind, then clearly (cf. (9))

$$(27) \quad \sum^* \sum_{k \in \mathfrak{R}(I)} l[s_k(I)] \leq l_*.$$

2.7. *Continuation.* Assuming that $I: x_1 \leq x \leq x_2$ is an interval of the second kind (see 2.5), we assert that the class $\mathfrak{R}(I)$ (see 2.6) contains at least two distinct subscripts. Indeed, let x_0 be any interior point of I . The points $(x_0, -K)$ and (x_0, K) are clearly comprised in the unbounded domain D_0 defined in 1.5. If the number of subscripts $k \in \mathfrak{R}(I)$ is less than two, then the segment $x = x_0$, $-K \leq y \leq K$ intersects the set F (see 1.5) in at most one point, and thus at most one point of this segment is not comprised in the unbounded domain D_0 . Hence (cf. 1.5) $\mu(x_0, y) = 0$ for $-K \leq y \leq K$, with the exception of at most one value of y . Hence (see (21)), $q(x_0) = 0$ for every interior point x_0 of I , in contradiction with the assumption that I is of the second kind.

2.8. *Continuation.* For $k \in \mathfrak{R}(I)$, let $y = \alpha_k x + \beta_k$ be the equation of the line that contains the segment $s_k(I)$ (see 2.6). Define, for $-K \leq y \leq K$, a function $n_k(y, I)$ as follows: $n_k(y, I) = 1$ if the point $(0, y)$ lies in the projection upon the y -axis of $s_k(I)$, and $n_k(y, I) = 0$ if this is not the case. Then the length of the projection of $s_k(I)$ upon the y -axis is equal to the integral of $n_k(y, I)$ from $-K$ to K , and it is also equal to $|\alpha_k| (x_2 - x_1)$. Hence

$$(28) \quad \int_{-K}^K n_k(y, I) dy = |\alpha_k| (x_2 - x_1) \quad \text{for } k \in \mathfrak{R}(I).$$

Now consider, for $-K \leq y \leq K$, the quantity $|\mu(x_2, y) - \mu(x_1, y)|$. Except for a finite number of values of y , this quantity can be estimated by means of the property (v) of μ stated in 1.5. Thus, except perhaps for a finite number of values of y , $|\mu(x_2, y) - \mu(x_1, y)|$ does not exceed the number of those segments $s_k(I)$ that are intersected by the segment with end points (x_1, y) and (x_2, y) . The number of these segments being clearly equal to $\sum n_k(y, I)$, $k \in \mathfrak{R}(I)$, we obtain the estimate

$$(29) \quad |\mu(x_2, y) - \mu(x_1, y)| \leq \sum n_k(y, I), \quad k \in \mathfrak{R}(I)$$

which is thus seen to hold with the possible exception of a finite number of values of y in the interval $-K \leq y \leq K$. Hence, in view of (28) and (21), integration of (29) yields the inequality

$$|q(x_2) - q(x_1)| \leq (x_2 - x_1) \sum |\alpha_k|, \quad k \in \mathfrak{R}(I).$$

Since $q(x)$ is linear in I , there follows the inequality

$$(30) \quad |q'(x)| \leq \sum |\alpha_k|, \quad k \in \mathfrak{R}(I), \quad x \in I^0,$$

where I^0 denotes the interior of I .

2.9. Let us now return to the function $\psi(x)$, defined by (23). From 2.4 it follows that $\psi(x)$ is continuous, non-negative, piecewise linear in the interval $-K \leq x \leq K$, and $\psi(-K) = \psi(K) = 0$. If I is an interval of the second kind (see 2.5), then we have by 2.8

$$\begin{aligned} \int_I [1 + \psi'(x)^2]^{1/2} dx &= \int_I [1 + 4^{-1} q'(x)^2]^{1/2} dx \\ &\leq 2^{-1} \left[4 + \left(\sum_{k \in \mathfrak{R}(I)} |\alpha_k| \right)^2 \right]^{1/2} (x_2 - x_1). \end{aligned}$$

Since the class $\mathfrak{R}(I)$ contains at least two distinct subscripts k (see 2.7), we obtain the further inequality

$$[4 + (\sum |\alpha_k|)^2]^{1/2} \leq [(\sum 1)^2 + (\sum |\alpha_k|)^2]^{1/2} \leq \sum (1 + \alpha_k^2)^{1/2},$$

where the summations are relative to the subscripts $k \in \mathfrak{R}(I)$. Now since (cf. 2.8)

$$l[s_k(I)] = (1 + \alpha_k^2)^{1/2} (x_2 - x_1),$$

there follows the inequality

$$\int_I [1 + \psi'(x)^2]^{1/2} dx \leq 2^{-1} \sum l[s_k(I)], \quad k \in \mathfrak{R}(I).$$

In view of (27), summation over all the intervals I of the second kind yields

$$\int_E [1 + \psi'(x)^2]^{1/2} dx \leq 2^{-1} l_*,$$

where E is the set of those points x in the interval $-K \leq x \leq K$ where $\psi(x) > 0$. In view of (18), the inequality (24) follows, and the proof of (12) is complete.

2.10. The idea of the preceding proof is to reduce the general inequality (12) to the ordinary isoperimetric inequality for a simply-connected, symmetric polygonal region, by means of the process of symmetrization due to Steiner (cf. Bonnesen [1, pp. 75-77]). In a certain sense, the inequality (12) represents the maximum amount of information that can be obtained by a careful analysis of a classical line of thought.

3. Elementary inequalities in three-space. 3.1. Let $\psi(x)$ be a function which is continuous and piecewise quadratic in an interval $-K \leq x \leq K$ (that is, this interval can be subdivided into a finite number of intervals in each of which $\psi(x)$ is a quadratic polynomial of x). Assume further that $\psi(x) \geq 0$ for $-K \leq x \leq K$, and $\psi(-K) = \psi(K) = 0$. We have then the inequality

$$(31) \quad \int_{-K}^K \psi(x) dx \leq \left\{ \int_{-K}^K [4\pi\psi(x) + \psi'(x)^2]^{1/2} dx \right\}^{3/2} / 6\pi^{1/2}.$$

Indeed, in view of the special character of the function $\psi(x)$, we can subdivide the interval $-K \leq x \leq K$ into a finite number of intervals $i_1, \dots, i_k, \dots, i_n$ in such a manner that the following holds. (i) $\psi(x)$ vanishes at the end points of i_k , $k=1, \dots, n$. (ii) For each k , either $\psi(x) \equiv 0$ in i_k or else $\psi(x) > 0$ in the

interior of i_k . Let us put

$$(32) \quad I_k = \int_{i_k} \psi(x) dx, \quad J_k = \int_{i_k} [4\pi\psi(x) + \psi'(x)^2]^{1/2} dx.$$

We assert that

$$(33) \quad I_k \leq J_k^{3/2} / 6\pi^{1/2}, \quad k = 1, \dots, n.$$

Indeed, if $\psi(x) \equiv 0$ in i_k , then (33) is obvious. If $\psi(x) > 0$ in the interior of i_k , then let us put $\eta_k(x) = [\psi(x)/\pi]^{1/2}$, $x \in i_k$. Let us introduce coordinate axes y, z that form a Cartesian system together with the x -axis, and let us consider the surface of revolution S_k obtained by rotating around the x -axis the curve $y = \eta_k(x)$, $x \in i_k$. Now in view of the elementary character of the function $\eta_k(x)$, the validity of the isoperimetric inequality $V^2 \leq A^3/36\pi$ for the surface of revolution S_k follows from a classical paper of H. A. Schwarz [1], where now

$$(34) \quad V = \pi \int_{i_k} \eta_k(x)^2 dx, \quad A = 2\pi \int_{i_k} \eta_k(x) [1 + \eta_k'(x)^2]^{1/2} dx.$$

As a matter of fact, as far as the argument used by H. A. Schwarz is concerned, the preceding familiar formulas may be taken as the definitions of V and A for the case of piecewise analytic surfaces of revolution. If we replace, in (34), $\eta_k(x)$ by $[\psi(x)/\pi]^{1/2}$, then the inequality $V^2 \leq A^3/36\pi$, established by H. A. Schwarz for the special case needed here, yields (33). From (33) we infer that

$$\begin{aligned} \int_{-K}^K \psi(x) dx &= \sum_{k=1}^n I_k \leq \frac{1}{6\pi^{1/2}} \sum_{k=1}^n J_k^{3/2} \leq \frac{1}{6\pi^{1/2}} \left(\sum_{k=1}^n J_k \right)^{3/2} \\ &= \frac{1}{6\pi^{1/2}} \left\{ \int_{-K}^K [4\pi\psi(x) + \psi'(x)^2]^{1/2} dx \right\}^{3/2}, \end{aligned}$$

and (31) is established.

3.2. We proceed to prove the inequality (16) in 1.6. Using the terminology and the notations introduced in 1.6, let us select a cartesian coordinate system xyz such that the z -axis is not parallel to any one of the planes containing the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$. For each $k=1, \dots, m$, let δ_k denote the projection of Δ_k upon the xy -plane. Then δ_k is a non-degenerate triangle. Let $K > 0$ be so large that the set $F = \Delta_1 + \Delta_2 + \dots + \Delta_m$ is interior to the cube $-K \leq x \leq K, -K \leq y \leq K, -K \leq z \leq K$. For $-K \leq \xi \leq K$, let $\Delta_k(\xi)$ denote the set of those points $(x, y, z) \in \Delta_k$ for which $x \leq \xi$, and let $\delta_k(\xi)$ have a similar meaning relative to the triangle δ_k . The sets $\Delta_k(\xi), \delta_k(\xi)$ are empty or non-empty for the same values of ξ . For example, clearly $\Delta_k(-K) = \delta_k(-K) = 0$. The (two-dimensional) measures of $\Delta_k(\xi), \delta_k(\xi)$ will be denoted by $A_k(\xi)$,

$a_k(\xi)$ respectively, while $b_k(\xi)$ will denote the (two-dimensional) measure of the projection of $\Delta_k(\xi)$ upon the yz -plane (it being understood that the measure of the empty set is equal to zero). The intersection of the plane $x = \xi$ with the triangle Δ_k is either empty or else it is a single point or a straight segment. In the third case, $l_k(\xi)$ will denote the length of this straight segment, while in the first two cases we set $l_k(\xi) = 0$. The symbol $\lambda_k(\xi)$ is defined in a similar manner relative to the triangle δ_k . We put $l(\xi) = l_1(\xi) + l_2(\xi) + \cdots + l_m(\xi)$. For $-K < x_1 < x_2 < K$, we define $g_k(y, z, x_1, x_2)$ as the characteristic function of the projection, upon the yz -plane, of the set of those points $(x, y, z) \in \Delta_k$ for which $x_1 \leq x \leq x_2$ (the characteristic function of a set E is equal to 1 for points of E and equal to zero for points not in E).

3.3. *Continuation.* For each $k = 1, \dots, m$, let $z = \alpha_k x + \beta_k y + \gamma_k$ be the equation of the plane that contains the triangle Δ_k . The following formulas are readily verified.

$$(35) \quad A_k(\xi) = (1 + \alpha_k^2 + \beta_k^2)^{1/2} a_k(\xi),$$

$$(36) \quad b_k(x_2) - b_k(x_1) = \int_{-K}^K \int_{-K}^K g_k(y, z, x_1, x_2) dy dz,$$

$$(37) \quad b_k(x_2) - b_k(x_1) = |\alpha_k| [a_k(x_2) - a_k(x_1)],$$

$$(38) \quad l_k(\xi) = (1 + \beta_k^2)^{1/2} \lambda_k(\xi).$$

Clearly, $a_k(\xi)$ is a non-negative, continuous, and piecewise quadratic function of ξ for $-K \leq \xi \leq K$, and hence the same holds for $A_k(\xi)$ in view of (35). Since clearly $a'_k(\xi) = \lambda_k(\xi)$, with the possible exception of a finite number of ξ -values, we have (with the same exceptions)

$$(39) \quad A'_k(\xi) = (1 + \alpha_k^2 + \beta_k^2)^{1/2} \lambda_k(\xi).$$

3.4. *Continuation.* The function μ of 1.6 becomes, on introducing the Cartesian coordinate system xyz , a function $\mu(x, y, z)$. Let us take any two numbers x_1, x_2 such that $-K < x_1 < x_2 < K$. Let $G(x_1), G(x_2)$ denote the sets in which the planes $x = x_1, x = x_2$ respectively intersect the set $F = \Delta_1 + \Delta_2 + \cdots + \Delta_m$. Furthermore, let G denote the set consisting of all the sides of all the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$. Finally, let $E(x_1, x_2)$ be the projection of the set $G(x_1) + G(x_2) + G$ upon the yz -plane. Then $E(x_1, x_2)$ is of (two-dimensional) measure zero, since none of the triangles $\Delta_1, \Delta_2, \dots, \Delta_m$ is parallel to the yz -plane. If the point $(0, y, z)$ does not lie in the set $E(x_1, x_2)$, then the segment σ with end points $(x_1, y, z), (x_2, y, z)$ is clearly admissible in the sense of 1.6. Hence, by property (v) of the function μ (see 1.6), the quantity $|\mu(x_2, y, z) - \mu(x_1, y, z)|$ does not exceed the number of triangles $\Delta_1, \dots, \Delta_m$ intersected by the segment σ . This remark yields the inequality (cf. 3.2)

$$|\mu(x_2, y, z) - \mu(x_1, y, z)| \leq \sum_{k=1}^m g_k(y, z, x_1, x_2),$$

for $(0, y, z) \notin E(x_1, x_2)$. Since $E(x_1, x_2)$ is of (two-dimensional) measure zero, integration yields (cf. (36), (37)) the inequality

$$(40) \quad \int_{-K}^K \int_{-K}^K |\mu(x_2, y, z) - \mu(x_1, y, z)| dy dz \leq \sum_{k=1}^m |\alpha_k| [a_k(x_2) - a_k(x_1)].$$

3.5. *Continuation.* Let us define

$$h(x, y) = \int_{-K}^K \mu(x, y, z) dz,$$

for $-K \leq x \leq K$, $-K \leq y \leq K$. Since none of the triangles $\Delta_1, \dots, \Delta_m$ is parallel to the z -axis, a reasoning entirely analogous to that used in 2.4 shows that $h(x, y)$ is continuous in the square $-K \leq x \leq K$, $-K \leq y \leq K$. In view of the special character of the function $\mu(x, y, z)$ it follows further readily that $h(x, y)$ is piecewise linear in this square (that is, the square can be subdivided into a finite number of polygonal regions in each of which $h(x, y)$ is a linear function of x and y). On setting

$$(41) \quad Q(x) = \int_{-K}^K \int_{-K}^K \mu(x, y, z) dy dz = \int_{-K}^K h(x, y) dy,$$

it follows that $Q(x)$ is non-negative, continuous, and piecewise quadratic in the interval $-K \leq x \leq K$. Since the planes $x = -K$ and $x = K$ are comprised in the unbounded domain D_0 (see 1.6), it is clear that $Q(-K) = Q(K) = 0$. Thus $Q(x)$ satisfies all the assumptions made in 3.1 concerning the function $\psi(x)$, and hence (see (31))

$$(42) \quad \int_{-K}^K Q(x) dx \leq \left\{ \int_{-K}^K [4\pi Q(x) + Q'(x)^2]^{1/2} dx \right\}^{3/2} / 6\pi^{1/2}.$$

From (40) and (41) we infer that

$$|Q(x_2) - Q(x_1)| \leq \sum_{k=1}^m |\alpha_k| [a_k(x_2) - a_k(x_1)].$$

Hence, with the possible exception of a finite number of x -values (cf. 3.3)

$$(43) \quad |Q'(x)| \leq \sum_{k=1}^m |\alpha_k| a'_k(x) = \sum_{k=1}^m |\alpha_k| \lambda_k(x), \quad -K \leq x \leq K.$$

In view of (35), (38) we have, with the possible exception of a finite number of x -values,

$$\begin{aligned} A'_k(x) &= (1 + \alpha_k^2 + \beta_k^2)^{1/2} a'_k(x) = (1 + \alpha_k^2 + \beta_k^2)^{1/2} \lambda_k(x) \\ &= [l_k(x)^2 + |\alpha_k \lambda_k(x)|^2]^{1/2}. \end{aligned}$$

On setting

$$(44) \quad A(x) = A_1(x) + \cdots + A_m(x),$$

there follows the relation (cf. 3.2 and (43))

$$(45) \quad \begin{aligned} A'(x) &= \sum_{k=1}^m [l_k(x)^2 + |\alpha_k \lambda_k(x)|^2]^{1/2} \\ &\geq \left\{ \left[\sum_{k=1}^m l_k(x) \right]^2 + \left[\sum_{k=1}^m |\alpha_k \lambda_k(x)|^2 \right] \right\}^{1/2} \geq [l(x)^2 + Q'(x)^2]^{1/2}. \end{aligned}$$

Now let ξ be any number between $-K$ and K , such that the plane $x = \xi$ does not contain any one of the vertices of the triangles $\Delta_1, \dots, \Delta_m$. We assert then the inequality

$$(46) \quad l(\xi)^2 \geq 4\pi Q(\xi).$$

Indeed, if the plane $x = \xi$ does not intersect any one of the triangles $\Delta_1, \dots, \Delta_m$, then this plane is comprised in the unbounded domain D_0 (see 1.6), and hence $Q(\xi) = 0$ in this case, and (46) is obvious. If the plane $x = \xi$ does intersect at least one of the triangles $\Delta_1, \dots, \Delta_m$, then let $\Delta_{k_1}, \dots, \Delta_{k_j}$ be the triangles intersected by this plane, and let s_{k_1}, \dots, s_{k_j} be the intersections. Then the segments s_{k_1}, \dots, s_{k_j} , jointly with the function $\mu(\xi, y, z)$, clearly satisfy the assumptions made in 1.5, and thus (46) is a direct consequence of the inequality (12). From (45), (46) we infer that

$$(47) \quad A'(x) \geq [4\pi Q(x) + Q'(x)^2]^{1/2}, \quad -K \leq x \leq K,$$

with the possible exception of a finite number of x -values. Now clearly (cf. (44), 3.3, 1.6) $A(K) = A_*$, $A(-K) = 0$, and hence, by (47),

$$(48) \quad A_* = A(K) - A(-K) = \int_{-K}^K A'(x) dx \geq \int_{-K}^K [4\pi Q(x) + Q'(x)^2]^{1/2} dx.$$

On the other hand (see 1.6, 3.5)

$$(49) \quad V_* = \int_{-K}^K \int_{-K}^K \int_{-K}^K \mu(x, y, z) dx dy dz = \int_{-K}^K Q(x) dx.$$

In view of (49), (42), (48), the inequality (16) follows.

4. The isoperimetric inequalities. 4.1. The limit processes, leading from the elementary inequalities (12) and (16) to the isoperimetric inequality for general curves and surfaces, are based on certain simple and general properties of the quantities $l(C)$, $a(C)$, $A_L(S)$, $V(S)$ that we shall first review.

Let C be an oriented closed curve (see 1.2), and let C_n be a sequence of such curves, all in a given Euclidean xy -plane. The sequence C_n is said to converge to C , in symbols $C_n \rightarrow C$, if there exist simultaneous representations (cf. 1.2)

$$\begin{aligned} C: \quad x &= x(p), & y &= y(p), & p &\in \Gamma, \\ C_n: \quad x &= x_n(p), & y &= y_n(p), & p &\in \Gamma, \end{aligned}$$

such that $x_n(p) \rightarrow x(p)$, $y_n(p) \rightarrow y(p)$ uniformly on Γ . Suppose that $C_n \rightarrow C$. We assert the inequalities

$$(50) \quad l(C) \leq \liminf l(C_n),$$

$$(51) \quad a(C) \leq \liminf a(C_n),$$

which express the fundamental fact that the length $l(C)$ and the enclosed area $a(C)$ (as defined by (7)) are *lower semicontinuous functionals*. This property of $l(C)$ is familiar. As regards $a(C)$, we may reason as follows. Let $i(x, y)$, $i_n(x, y)$ be the topological index-functions associated with C , C_n respectively (cf. 1.3). We assert that

$$(52) \quad |i(x, y)| \leq \liminf |i_n(x, y)|.$$

Indeed, if the point (x, y) lies on the set $[C]$ (see 1.2), then $i(x, y) = 0$ by definition, and (52) is trivial. If $(x, y) \notin [C]$, then $i_n(x, y) \rightarrow i(x, y)$ by well known properties of the topological index, and (52) is thus obvious in this case also. Now let us note that (51) is obvious if $\liminf a(C_n) = +\infty$. On the other hand, if $\liminf a(C_n) < +\infty$, then (51) follows by a well known theorem of Fatou from (52) (see, for instance, Saks [1]).

4.2. An oriented closed curve C will be termed an *oriented closed polygon* if it admits of a representation (cf. 1.2)

$$C: \quad x = x(p), \quad y = y(p), \quad p \in \Gamma,$$

with the following property: The unit circle Γ can be subdivided into a finite number of arcs $\gamma_1, \gamma_2, \dots, \gamma_m$, such that each one of the arcs γ_k is mapped, by the equations $x = x(p)$, $y = y(p)$, topologically onto a straight segment s_k in the xy -plane. It follows readily that in this case $l(C)$ is equal to the sum of the lengths of the segments s_1, \dots, s_m . We shall use the symbol P as a generic notation for an oriented closed polygon. From the definition of $l(C)$, it follows readily that for any given oriented closed curve C there exists a sequence of oriented closed polygons P_n such that $P_n \rightarrow C$ and $l(P_n) \rightarrow l(C)$.

4.3. Entirely similar considerations apply to surfaces. Let S be an oriented closed surface (see 1.4), and let S_n be a sequence of such surfaces. The sequence S_n is said to converge to S , in symbols $S_n \rightarrow S$, if there exist simultaneous representations (see 1.4)

$$\begin{aligned} S: \quad x &= x(p), & y &= y(p), & z &= z(p), & p &\in U, \\ S_n: \quad x &= x_n(p), & y &= y_n(p), & z &= z_n(p), & p &\in U, \end{aligned}$$

such that $x_n(p) \rightarrow x(p)$, $y_n(p) \rightarrow y(p)$, $z_n(p) \rightarrow z(p)$ uniformly on U . Suppose that $S_n \rightarrow S$. We assert the inequalities

$$(53) \quad A_L(S) \leq \liminf A_L(S_n),$$

$$(54) \quad V(S) \leq \liminf V(S_n),$$

which express the fundamental fact that the Lebesgue area $A_L(S)$ and the enclosed volume $V(S)$ (as defined by (8)) are *lower semi-continuous functionals*. This property of $A_L(S)$ is familiar. As regards $V(S)$, the proof is entirely similar to that given in 4.1 for the inequality (51).

4.4. An oriented closed surface S will be termed an *oriented closed polyhedron* if it admits of a representation (cf. 1.4)

$$S: \quad x = x(p), \quad y = y(p), \quad z = z(p), \quad p \in U,$$

with the following property: The sphere U can be subdivided into a finite number of curvilinear triangles t_1, \dots, t_m , such that each triangle t_k is mapped, by the equations $x=x(p)$, $y=y(p)$, $z=z(p)$, topologically onto a plane rectilinear triangle Δ_k in xyz -space. It can be shown (see, for instance, Youngs [2]) that the Lebesgue area of S is then equal to the sum of the areas of the triangles $\Delta_1, \dots, \Delta_m$. From the definition of the Lebesgue area it follows readily that for any given oriented closed surface S there exists a sequence of oriented closed polyhedra \mathfrak{P}_n such that $\mathfrak{P}_n \rightarrow S$ and $A_L(\mathfrak{P}_n) \rightarrow A_L(S)$.

4.5. Now let there be given an oriented closed polygon P in terms of a representation $P: x=x(p)$, $y=y(p)$, $p \in \Gamma$, with the property stated in 4.2, and let s_1, s_2, \dots, s_m be the straight segments described there. Let $i(x, y)$ be the corresponding index-function (see 1.3). Whenever the point (x, y) crosses a segment s_k , the index-function $i(x, y)$ changes by ± 1 . In view of this fact it follows that the segments s_1, s_2, \dots, s_m , jointly with $\mu(x, y) = |i(x, y)|$, satisfy the assumptions made in 1.5. Since, for this choice of $\mu(x, y)$, the a_* of formula (11) coincides with the enclosed area defined by (7), the elementary inequality (12) shows that we have the isoperimetric inequality $a(P) \leq l(P)^2/4\pi$ for every oriented closed polygon P .

4.6. Let there be given an oriented closed curve C in the xy -plane. By 4.2 we have then a sequence P_n of oriented closed polygons such that $P_n \rightarrow C$ and $l(P_n) \rightarrow l(C)$. By (51) we have then the relations

$$(55) \quad l(C) = \lim l(P_n), \quad a(C) \leq \liminf a(P_n)$$

at our disposal. By 4.5 we have the inequality

$$(56) \quad a(P_n) \leq l(P_n)^2/4\pi.$$

The relations (55), (56) yield the isoperimetric inequality $a(C) \leq l(C)^2/4\pi$ for every oriented closed curve C , where the enclosed area $a(C)$ is defined by formula (7) in 1.3.

4.7. The proof of the spatial isoperimetric inequality is entirely analogous. In the first place, a reasoning similar to that in 4.5 shows that for oriented closed polyhedra \mathfrak{P} the inequality $V(\mathfrak{P})^2 \leq A_L(\mathfrak{P})^3/36\pi$ is an immediate con-

sequence of the elementary inequality (16). Let then S be any oriented closed surface. By 4.4 we have a sequence of oriented closed polyhedra \mathfrak{P}_n such that $\mathfrak{P}_n \rightarrow S$, $A_L(\mathfrak{P}_n) \rightarrow A_L(S)$. By 4.3 we have then the relations

$$A_L(S) = \lim A_L(\mathfrak{P}_n), \quad V(S) \leq \liminf V(\mathfrak{P}_n).$$

Since each \mathfrak{P}_n is an oriented closed polyhedron, we already know that

$$V(\mathfrak{P}_n)^2 \leq A_L(\mathfrak{P}_n)^3/36\pi.$$

These relations yield the spatial isoperimetric inequality $V(S)^2 \leq A_L(S)^3/36\pi$, where the enclosed volume $V(S)$ is defined by the formula (8).

5. Miscellaneous comments. 5.1. The preceding discussion of the plane and spatial isoperimetric inequalities may be interpreted to show that the two cases are entirely analogous as regards concepts involved, methods used, and results achieved. And yet, the spatial isoperimetric inequality covers situations for which no analogues exist in the case of the plane isoperimetric inequality and which were, apparently, overlooked in previous treatments of the problem. The purpose of the following remarks is to amplify this general statement and to call attention to various further problems related to the spatial isoperimetric inequality.

5.2. While the theory of arc length and the theory of surface area show far-reaching analogies, there exist certain fundamental discrepancies that students of this field should keep in mind. One of the reasons for these discrepancies may be described as follows. Let γ be a simple polygonal line whose length $l(\gamma)$ is less than a given $\epsilon > 0$ and which passes through a given point p_0 . Then clearly γ is comprised in the sphere with center p_0 and radius ϵ . On the other hand, let Σ be a simple polyhedral surface whose area $A(\Sigma)$ (in the elementary sense) is less than a given $\epsilon > 0$ and which passes through a given point p_0 . Choosing Σ as a very long but very narrow rectangle, we see that a polyhedral surface may have very small area without being comprised in a small sphere. By properly folding the Σ of the preceding remark, one obtains an example of a simple closed polyhedral surface Σ^* with a surface area less than ϵ , such that Σ^* passes within ϵ of every point of an arbitrarily assigned bounded closed set F in xyz -space. Briefly, a polygonal line of small length is necessarily confined to a small sphere, while *a polyhedral surface of small area may be a surprisingly sprawling figure*. The Σ^* just mentioned also shows that a simple closed polyhedral surface may pass within ϵ of every point of a given solid cube and yet enclose a volume less than ϵ , where $\epsilon > 0$ is arbitrarily assigned, and all the terms involved are used in the most elementary sense.

The elementary phenomena just referred to reveal fundamental differences between arc length and surface area which account for many of the difficulties that arise in various fields, including the theory of double integral problems in Calculus of Variations. By means of appropriate passages to the limit, these elementary phenomena give rise to curious examples due to

Geöcze [1] and Besicovitch [1]. To simplify the presentation, let us introduce the following terminology. An oriented closed surface S , in the sense of 1.4, will be termed *simple* if it admits of a representation (see 1.4)

$$S: \quad x = x(p), \quad y = y(p), \quad z = z(p), \quad p \in U,$$

such that the correspondence between the points p of U and the image points (x, y, z) is biunique. The point-set $[S]$ (see 1.4) is then a simple closed surface in the sense of point-set-theoretical topology. By well known theorems, the point-set $[S]$ then divides the space xyz into two components, and the bounded component will be denoted by D_S and will be termed the interior of S . Furthermore, the index-function $i(x, y, z)$ is now equal to ± 1 on D_S and equal to zero on the unbounded component. Recalling that $i(x, y, z) = 0$ on $[S]$ by definition, we see that (cf. 1.4)

$$(57) \quad V(S) = |D_S|,$$

where $|D_S|$ denotes the three-dimensional measure of D_S .

5.3. Now let F be any bounded closed set in xyz -space, and let λ_n be any sequence of positive numbers. By the elementary remarks made above, we have then a sequence of oriented closed polyhedra

$$(58) \quad \mathfrak{P}_n: \quad x = x_n(p), \quad y = y_n(p), \quad z = z_n(p), \quad p \in U,$$

such that

$$(59) \quad A_L(\mathfrak{P}_n) < \lambda_n,$$

and such that the point-set $[\mathfrak{P}_n]$ passes within λ_n of every point of the set F . Suppose now that the sequences $x_n(p)$, $y_n(p)$, $z_n(p)$ converge uniformly on U , and let $x(p)$, $y(p)$, $z(p)$ denote the limit functions. For the oriented closed surface S defined by

$$S: \quad x = x(p), \quad y = y(p), \quad z = z(p), \quad p \in U,$$

we have then obviously the relation

$$(60) \quad F \subset [S].$$

Furthermore, since A_L is a lower semi-continuous functional (see 4.3), we infer from (59) the inequality

$$(61) \quad A_L(S) \leq \liminf \lambda_n.$$

As a matter of fact, we can impose further requirements upon the sequence \mathfrak{P}_n , and we obtain then interesting examples due to Geöcze and Besicovitch which we shall describe presently.

5.4. *Continuation.* As regards (60), we may ask whether we can achieve the relation $F = [S]$. Since $[S]$ is a continuous image of the unit sphere U , it is necessary for this purpose that F be a locally connected continuum. The

construction devised by Geöcze [1] (see also Radó [4]) yields the following result: by proper choice of the sequence \mathfrak{P}_n , the set $[S]$ can be made to coincide with any assigned locally connected continuum F , for arbitrary choice of the sequence λ_n . The following special cases are of interest.

(a) $\lambda_n \rightarrow 0$, while F is a solid cube. Then (61) yields $A_L(S) = 0$.

(b) $\lambda_n \rightarrow 0$, while F is the surface of a cube. Again $A_L(S) = 0$ by (61).

Now we noted in 1.4 that the surface S and the point-set $[S]$ are essentially different geometrical objects. If this important point is disregarded, the (a) and (b) may be described by the following picturesque statements.

(a*) A surface may fill a solid cube and yet have zero area.

(b*) A surface may cover the surface of a cube and have zero area.

The fundamental observations of Geöcze concerning surfaces of zero Lebesgue area lead to problems of great interest and difficulty (see Radó [4]). We restrict ourselves to remarks relevant for the spatial isoperimetric inequality. Since we have, by 4.7, $V(S)^2 \leq A_L(S)^3/36\pi$ for every oriented closed surface S , it follows that the relation $A_L(S) = 0$ implies the relation $V(S) = 0$. Hence we can amplify the above statement (b*), for example, as follows: there exists an oriented closed surface S that covers the whole surface of a cube, and yet has zero area and encloses zero volume. While these situations are immediate consequences of the elementary phenomena discussed in 5.2, they reveal in striking form the need for a careful distinction between the surface S and the point-set $[S]$. In the statement (b*), for instance, if the side length of the cube is equal to one, inspection of the point-set $[S]$ alone would suggest that the area of S should be equal at least to 6, and the enclosed volume should be equal to one, while actually $A_L(S) = 0$, $V(S) = 0$.

5.5. We may ask, in connection with the remarks in 5.3, whether the limit surface S may be required to be an oriented simple closed surface (cf. 5.2). Besicovitch [1] obtained the following example. On choosing F as a certain totally disconnected closed bounded set of three-dimensional measure $1/2$, he constructs a sequence of oriented simple closed polyhedra \mathfrak{P}_n with the following properties. (α) $A_L(\mathfrak{P}_n) < \epsilon$ for every n , where $\epsilon > 0$ is assigned arbitrarily. (β) The limit surface S is simple. (γ) The limit surface S satisfies the relation $F \subset [S]$. As noted in 5.3, the property (α) implies that $A_L(S) \leq \epsilon$. On the other hand, (γ) implies that $|[S]|$, the three-dimensional measure of the point-set $[S]$, is at least equal to $1/2$. Besicovitch expresses the view that this situation shows the inadequacy of the Lebesgue definition of surface area. If the writer is interpreting his remarks correctly, then he seems to feel that for a surface S of finite area the point-set $[S]$ should be of zero three-dimensional measure, at least if S is an oriented simple closed surface. Of course, the analogous requirement for arc length is satisfied: if C is a simple plane curve of finite length, then the point-set $[C]$ is of zero two-dimensional measure, and in fact this is true for general Fréchet curves also. Thus it seems entirely reasonable to require that a similar statement should hold for surface area.

But it seems equally reasonable to require that surface area should be a lower semi-continuous functional, since arc length has this property, and furthermore, the lower semicontinuity of surface area can be verified in all elementary cases. Thus the Besicovitch example merely shows that two equally reasonable postulates conflict with each other beyond the elementary range. In this connection, it is interesting to note that according to an investigation by Nöbeling [1], any two of the more relevant definitions of surface area can be shown to conflict with each other beyond the elementary range. Let us note also that the conflict, beyond the elementary range, between various equally plausible postulates concerning surface area has been fully understood in the literature on Lebesgue area (see, for example, Radó [1], where the conflict between the projection postulate and the lower semicontinuity postulate is discussed).

5.6. Let S denote the oriented simple closed surface constructed by Besicovitch. Since $|[S]| > 0$, we may want to distinguish between the *interior volume* $V_i = |D_S|$ (cf. 5.2) and the *exterior volume* $V_e = |D_S| + |[S]|$, and we may consider the isoperimetric inequality in either one of the forms

$$(62) \quad V_e^2 \leq A_L(S)^3 / 36\pi,$$

$$(63) \quad V_i^2 \leq A_L(S)^3 / 36\pi.$$

Besicovitch uses the term *volume of the enclosed solid*, meaning apparently V_e , and notes that (62) fails to hold. Indeed, in his example $V_e \geq |[S]| \geq 1/2$, while $A_L(S) \leq \epsilon$, where ϵ is arbitrarily assigned. In discussions that came to the attention of the writer, this situation was construed to show that the Lebesgue area is inadequate from the point of view of the isoperimetric inequality. It is therefore of interest to observe that the isoperimetric inequality holds for the interior volume V_i . Indeed, by (57) we have $V_i = V(S)$, and hence (63) holds by 4.7. In fact, by (57) and 4.7 we have the inequality

$$(64) \quad |D_S|^2 \leq A_L(S)^3 / 36\pi$$

for every oriented simple closed surface S . Thus the Besicovitch example merely shows that the concept of enclosed volume must be adjusted to the concept of surface area if the isoperimetric inequality is to hold, a fact which is strikingly illustrated by the classical example of the cube-filling surface of zero area constructed by Geöcze.

Since in the Besicovitch example $|[S]| > 0$, any surface area $A_B(S)$ conforming to the views of Besicovitch (see 5.5) should be infinite for his surface S . On the other hand, as (64) shows, the Lebesgue area $A_L(S)$ still yields a nontrivial inequality.

5.7. In previous treatments of the spatial isoperimetric inequality in terms of the Lebesgue area, no reference is made to the type of phenomena exhibited by the examples of Geöcze and Besicovitch (see Tonelli [1],

Blaschke [1], Bonnesen [1]). As a matter of fact, all these treatments seem to involve a relevant gap. After a discussion of the case of polyhedra, the case of a general surface is settled by these authors by the following remark. Let \mathfrak{P}_n be a sequence of polyhedra converging to the given surface S , such that the area of \mathfrak{P}_n converges to the area of S . Since the isoperimetric inequality is already known to hold for polyhedra, its validity follows for the general surface S , *since the volume enclosed by \mathfrak{P}_n converges to the volume enclosed by S* . The last statement is the one that we wish to investigate, restricting ourselves to the work of Tonelli which represents the most serious effort. In the first place, it should be noted that Tonelli does not seem to give a precise formal definition for enclosed volume, and so we shall use the symbol $V_T(S)$ for the volume he had in mind. Furthermore, we shall restrict ourselves to the case of oriented simple closed surfaces S , and we shall use the symbols D_S , V_i , V_e in the sense of 5.6. Let S be the surface constructed by Besicovitch, and let \mathfrak{P}_n be the sequence of polyhedra occurring in his construction (see 5.5). A somewhat more detailed study of the Besicovitch example reveals the following additional facts.

$$(65) \quad D_{\mathfrak{P}_1} \subset D_{\mathfrak{P}_2} \subset \cdots, \quad D_S = D_{\mathfrak{P}_1} + D_{\mathfrak{P}_2} + \cdots$$

$$(66) \quad A_L(\mathfrak{P}_n) \rightarrow A_L(S).$$

From (65) we infer that

$$(67) \quad |D_{\mathfrak{P}_n}| \rightarrow |D_S| = V_i < V_e.$$

Now the volume $V_T(S)$, used by Tonelli, may reasonably be assumed to coincide either with V_i or V_e , while for the polyhedron \mathfrak{P}_n we should have

$$(68) \quad V_T(\mathfrak{P}_n) = |D_{\mathfrak{P}_n}|.$$

According to Tonelli, (66) should imply that

$$V_T(\mathfrak{P}_n) \rightarrow V_T(S).$$

Hence, we should have one of the two relations

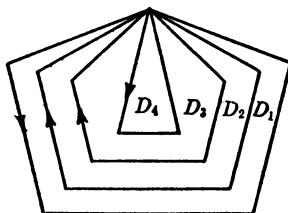
$$(69) \quad |D_{\mathfrak{P}_n}| \rightarrow V_e,$$

$$(70) \quad |D_{\mathfrak{P}_n}| \rightarrow V_i.$$

Now, in view of (67), it follows that (69) does not hold, and hence the interpretation $V_T(S) = V_e$ is *not* admissible. On the other hand, a slight modification of the Besicovitch example yields a surface S^* and a sequence of polyhedra \mathfrak{P}_n^* , such that (i) all these surfaces are simple, (ii) $\mathfrak{P}_n^* \rightarrow S^*$, (iii) $A_L(\mathfrak{P}_n^*) \rightarrow A_L(S^*)$, (iv) $|[S^*]| > 0$, (v) $D_{S^*} + [S^*] \subset D_{\mathfrak{P}_n^*} + [\mathfrak{P}_n^*]$ for every n . It follows that in this case the relation corresponding to (70) does not hold, and hence *the interpretation of V_T as the interior volume is also inadmissible*. Of course, the most plausible explanation is that Tonelli was unaware of the phe-

nomenon exhibited by the Bescovitch example. It should be noted that Bonnesen [1] assumes explicitly that the set $[S]$ should be of zero three-dimensional measure. Thus there arises the problem to decide whether the argument used by these authors is valid in this special case. The analogous problem for curves admits of an affirmative solution (see Radó [3]). The extension to surfaces seems to be an interesting problem.

5.8. In conclusion, let us call attention to further problems. Let P denote the oriented closed polygon suggested by the figure. Roughly, P consists of four simple closed polygons described in the sense indicated by the arrows. Let l be the length of P , and let $|D_i|$, $i = 1, 2, 3, 4$, denote the two-dimensional



measure of the domain D_i indicated by the figure. Then the isoperimetric inequality, as proved by Blaschke (see 1.3, formula (4)), yields

$$(71) \quad ||D_1| - |D_3|| \leq l^2/4\pi.$$

On the other hand, the isoperimetric inequality in the form proved in this paper yields (see 1.3)

$$(72) \quad |D_1| + |D_3| \leq l^2/4\pi.$$

Finally, direct application of the elementary inequality (12) yields

$$(73) \quad |D_1| + 2|D_2| + 3|D_3| + 4|D_4| \leq l^2/4\pi.$$

Thus (72) shows that we obtained an improvement over Blaschke's result. And yet, (73) reveals that for the case of polygons a generally sharper inequality may hold. It would be interesting to determine whether a similar improvement is possible for general oriented closed curves. Analogous remarks apply to the spatial isoperimetric inequality.

It should be observed that the Lebesgue area $A_L(S)$ is not the only lower semicontinuous area studied in the literature. For other lower semicontinuous areas (see Radó [5]) it is not known whether polyhedral approximations are possible. A study of such areas, from the point of view of the spatial isoperimetric problem, may lead to important issues.

The reader may have noted that we did not assert that the sign of equality holds in the spatial isoperimetric inequality *only* if the surface involved is a sphere. Indeed, Tonelli [1] already noted that this statement is generally false. This is due to the generality of the concept of surface used. For ex-

ample, it is easy to construct an oriented closed surface S such that the point-set $[S]$ consists of the surface U of the unit sphere, plus a straight segment attached to a point p_0 of U , where U is simply covered except for the point p_0 , and $V(S) = 4\pi/3$, $A_L(S) = 4\pi$. Thus $V(S)^2 = A_L(S)^3/36\pi$ in this case, while S is certainly not a sphere. It would be a matter of interest to study the structure of oriented closed surfaces for which the sign of equality holds in the spatial isoperimetric inequality.

BIBLIOGRAPHY

P. ALEXANDROFF and H. HOPF

1. *Topologie*, Berlin, 1935.

A. S. BESICOVITCH

1. *On the definition and the value of the area of a surface*, Quart. J. Math. Oxford Ser. vol. 16 (1945) pp. 86–102.

W. BLASCHKE

1. *Kreis und Kugel*, Leipzig, 1916.

T. BONNESEN

1. *Les problèmes des isopérimètres et des iséiphanes*, Paris, 1929.

Z. GEÖCZE

1. *Sur l'exemple d'une surface dont l'aire est égale à zéro et qui remplit un cube*, Bull. Soc. Math. France vol. 41 (1913) pp. 29–31.

W. GROSS

1. *Die Minimaleigenschaft der Kugel*, Monatshefte für Mathematik und Physik vol. 18 (1917) pp. 77–97.

H. D. HUSKEY

1. *Fréchet polyhedra*, Duke Math. J. vol. 11 (1944) pp. 417–425.

G. NÖBELING

1. *Über die Flächenmasse im Euklidischen Raum*, Math. Ann. vol. 118 (1943) pp. 687–701.

T. RADÓ

1. *Über das Flächenmass rektifizierbarer Flächen*, Math. Ann. vol. 100 (1928) pp. 445–479.
2. *On the problem of Plateau*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, 1933.
3. *A lemma on the topological index*, Fund. Math. vol. 27 (1936) pp. 212–225.
4. *On continuous path-surfaces of zero area*, Ann. of Math. vol. 44 (1943) pp. 173–191.
5. *On surface area*, Proc. Nat. Acad. Sci. U. S. A. vol. 31 (1945) pp. 102–106.

S. SAKS

1. *Theory of the integral*, Warsaw, 1937.

H. A. SCHWARZ

1. *Beweis des Satzes dass die Kugel kleinere Oberfläche besitzt als jeder andere Körper gleichen Volumens*, Gesammelte Mathematische Abhandlungen, vol. 2, pp. 327–340.

L. TONELLI

1. *Sulla proprietà minimo della sfera*, Rend. Circ. Mat. Palermo vol. 39 (1915) pp. 1–30.

J. W. T. YOUNGS

1. *Curves and surfaces*, Amer. Math. Monthly vol. 51 (1944) pp. 1–11.
2. *The topological theory of Fréchet surfaces*, Ann. of Math. vol. 45 (1944) pp. 753–785.

THE OHIO STATE UNIVERSITY,
COLUMBUS, OHIO.