22. Zariski, O., *On the linear connection index of the algebraic surfaces* $z^n = f(x, y)$, Proceedings of the National Academy of Sciences, vol. 15 (1929), pp. 494–501.

23. Zariski, O., *On the non-existence of curves of order* 8 *with* 16 *cusps*, American Journal of Mathematics, vol. 53 (1931), pp. 309–318.

24. Zariski, O., *On the irregularity of cyclic multiple planes*, Annals of Mathematics, (2), vol. 32 (1931), pp. 485–511.

25. Zariski, O., *Algebraic Surfaces*, Ergebnisse der Mathematik, vol. 3 (1935), pp. 160–181.

WELLS COLLEGE

JENSEN'S INEQUALITY*

BY E. J. MCSHANE

The simplest form of Jensen's inequality is that if $\phi(x)$ is a convex function, and m is the arithmetic mean of $x_1, \cdots, x_n$, then the mean of the numbers $\phi(x_n)$ is not less than $\phi(m)$. This inequality can be generalized in several different ways. The function $\phi(x)$ can be replaced by a convex function of several variables, and the arithmetic mean can be replaced by any one of several other means, as has been shown in various proofs. Since the inequality is of considerable utility, it seems worth while to have it established in a form which is general enough to cover a wide assortment of applications.

The proofs will rest on two well known properties of convex sets.† If K is closed and convex and a point p is not in K, then p can be separated from K by a hyperplane. If K is closed and convex and p is a boundary point of K, there is a hyperplane of support of K passing through p.

1. *The Inequality in Geometric and in Analytic Form.* It will be convenient in the following proofs to use these symbols and definitions:

R_n is n-dimensional Euclidean space. Its points will be denoted by $(z_1, \cdots, z_n)$ or by z. Linear functions $\sum a_i z_i$ or R_n will be symbolized by $l(z)$.

---

* Presented to the Society, December 31, 1936.

† A set is convex if for every pair P, Q of points of the set the line segment PQ is contained in the set.

L is a linear class of real valued functions $f(x)$ defined on a set E. It shall be supposed have the properties:

(1a) If f_1, f_2 are in L and k_1, k_2 are real numbers, $k_1f_1+k_2f_2$ is in L.

(1b) The function defined and constantly equal to 1 on E is in L.

Mf is a linear mean defined on L, with these properties:

(2a) $M1=1$.

(2b) If f_1, f_2 are in L and k_1, k_2 are real numbers, $M(k_1f_1+k_2f_2) = k_1Mf_1+k_2Mf_2$.

(2c) If $f(x)$ is in L and $f(x)\geqq 0$, then $Mf\geqq 0$.

If $\boldsymbol{f}(x)$ is an n-tuple of functions $(f_1(x), \cdots, f_n(x))$ of L, we denote by $\boldsymbol{Mf}$ the n-tuple $(Mf_1, \cdots, Mf_n)$. From (2) we obtain

(3) $Ml(\boldsymbol{f})=l(\boldsymbol{Mf})$ for every function $l(\boldsymbol{z})$ linear on R_n.

The geometric formulation of Jensen's inequality is as follows:

THEOREM 1. *Let* (1) *and* (2) *be satisfied. Let K be a closed convex point set in R_n. Let $f_1(x), \cdots, f_n(x)$ be functions of the class L such that $\boldsymbol{f}=(f_1, \cdots, f_n)$ is in K for all x in E. Then $\boldsymbol{Mf}$ is in K.*

Let $l(\boldsymbol{z})+c=0$ be a hyperplane in R_n such that K is entirely to one side, say $l(\boldsymbol{z})+c\geqq 0$ for $\boldsymbol{z}$ in K. Then $l(\boldsymbol{f})+c\geqq 0$ for all x, and $0\leqq M(l(\boldsymbol{f})+c)=Ml(\boldsymbol{f})+Mc=l(\boldsymbol{Mf})+c$, so that $\boldsymbol{Mf}$ lies on the same side of the hyperplane as K. That is, no hyperplane separates $\boldsymbol{Mf}$ from K. This is only possible if $\boldsymbol{Mf}$ is in K.

The more usual analytical formulation of the inequality is covered by the following theorem:

THEOREM 2. *Let* (1) *and* (2) *be satisfied. Let K be a closed convex point set in R_n, and let $\phi(\boldsymbol{z})$ be continuous and convex* on K. Let $f_1(x), \cdots, f_n(x)$ be functions of the class L such that $\boldsymbol{f}(x) = (f_1(x), \cdots, f_n(x))$ is in K for all x in E, and such, moreover, that $\phi(\boldsymbol{f}(x))$ is in the class L. Then $\phi(\boldsymbol{Mf})$ is defined and*

$$(J) \qquad \phi(\boldsymbol{Mf})\leqq M\phi(\boldsymbol{f}).$$

* We call ϕ convex on K if $\phi(\frac{1}{2}(\boldsymbol{z}_1+\boldsymbol{z}_2))\leqq\frac{1}{2}(\phi(\boldsymbol{z}_1)+\phi(\boldsymbol{z}_2))$ for all $\boldsymbol{z}_1, \boldsymbol{z}_2$ in K.

Denote the points $(z_1, \cdots, z_{n+1})$ of R_{n+1} by $\boldsymbol{z}^1$. These can also be denoted by $(\boldsymbol{z}, z_{n+1})$, where $\boldsymbol{z}$ is in R_n. Define K_1 to be the set in R_{n+1} of points $(\boldsymbol{z}, z_{n+1})$ such that $\boldsymbol{z}$ is in K and $z_{n+1} \geqq \phi(\boldsymbol{z})$. It is clear that K_1 is closed and convex, and for all x in E the point $(\boldsymbol{f}(x), \phi(\boldsymbol{f}(x))$ is in K_1. Hence by Theorem 1, the point $(\boldsymbol{Mf}, M\phi(\boldsymbol{f}))$ is in K_1; that is, $\boldsymbol{Mf}$ is in K and $M\phi(\boldsymbol{f}) \geqq \phi(\boldsymbol{Mf})$.

Obviously we could weaken our hypotheses somewhat by omitting the requirement that ϕ be continuous and assuming instead that the set K_1 is closed and convex. This would permit $\phi(\boldsymbol{z})$ to have infinite discontinuities on the boundary of K.

EXAMPLES: (1) Let $\alpha(x_1, \cdots, x_m)$ be a positively monotonic* function of the variables $(x_1, \cdots, x_m)$, and let E be a set measurable with respect to α and such that $0 < m_\alpha E < \infty$. Let L be the class of all functions $f(x)$ which are Lebesgue-Stieltjes integrable with respect to α over E. Let $Mf = \int_E f d\alpha / \int_E d\alpha$. Then conditions (1) and (2) are satisfied. The conclusion of Theorem 2 assumes the form

$$\phi\left(\int f_1 d\alpha \Big/ \int d\alpha, \cdots, \int f_n d\alpha \Big/ \int d\alpha\right)$$
$$\leqq \int \phi(f_1, \cdots, f_n) d\alpha \Big/ \int d\alpha.$$

In the next three examples requirements of convergence or integrability are too obvious to need statement:

(2) The range of x is $(1, \cdots, m)$ or $(1, 2, \cdots)$, so that $f(x)$ is a (finite or infinite) sequence $(a_1, a_2, \cdots)$, and $Mf = \sum c_i a_i / \sum c_1$, where $c_i \geqq 0$ and $0 < \sum c_i < \infty$.

(3) The functions f of L are continuous, and $Mf = \int_E dx / \int_E dx$.

(4) The functions f of L are Lebesgue measurable over the measurable set E, and $Mf = \int fp dx / \int p dx$, where $p(x) \geqq 0$ and $0 < \int p dx < \infty$.

(5) E is in the interval $(0,1)$, L is the class of *all* bounded functions on E, Mf is the Banach integral† of f over $(0, 1)$.

(6) E is the set of all real numbers, L the class of all uniformly almost periodic functions, Mf is the mean value of f.

* That is, a function whose mth difference is non-negative.

† Banach, *Théorie des Opérations Linéaires*, p. 31.

(7) More generally, E is any group, L is the class of all functions almost periodic on E, Mf is von Neumann's* mean value of f.

(8) In the space $(x_1, x_2, \cdots)$ of infinitely many dimensions, $(0 \leqq x_i \leqq 1)$, E is a set of finite positive measure,† L is the class of functions summable over E, $Mf = \int_E f dx / mE$.

2. *Conditions for Strict Inequality.* In §1 we have not mentioned conditions for strict inequality. To investigate this question it is convenient to define negligible sets. A set $S \subset E$ is negligible (with respect to L and M) if there exists a function f in the class L such that

$$\text{(a)}\quad f(x) \geqq 0 \text{ on } E, \qquad \text{(b)}\quad f(x) > 0 \text{ on } S, \qquad \text{(c)}\quad Mf = 0.$$

It follows readily that every subset of a negligible set is negligible, and so is every set which is the sum of a finite number of negligible sets. It is then easy to prove the following theorem:

THEOREM 3. *In Theorem* 1, $\boldsymbol{Mf}$ *is a boundary point of K only if all points $\boldsymbol{f}(x)$ except those corresponding to a negligible set of x belong to the intersection of K with one of its hyperplanes of support.*

If $\boldsymbol{Mf}$ is a boundary point of K, through it there passes a hyperplane of support π: $l(\boldsymbol{z})+c=0$ of K; say $l(\boldsymbol{z})+c \geqq 0$ for $\boldsymbol{z}$ in K. Let S be the set of x such that $\boldsymbol{f}(x)$ is not in the intersection πK; then $l(\boldsymbol{f}(x))+c>0$ on S. Since $l(\boldsymbol{f}(x))+c \geqq 0$ for all x and $M(l(\boldsymbol{f}(x))+c) = l(\boldsymbol{Mf})+c=0$, we see that S is negligible.

I omit the easy proof of the following theorem:

THEOREM 4. *If the set K is strictly convex, and ϕ is strictly convex on K, then in Theorem* 2 *equality holds only if $f_i(x) = Mf_i$* $=$const. $(i=1, \cdots, n)$ *except on a negligible set.*

I am unable to state whether the conditions $f_i = Mf_i$ except on a negligible set are sufficient as well as necessary for equality in Theorem 4. However, by adding a further postulate concern-

---

* J. von Neumann, *Almost periodic functions in a group*, Transactions of this Society, vol. 36 (1934); in particular, p. 452.

† P. J. Daniell, *Integrals in an infinite number of dimensions*, Annals of Mathematics, (2), vol. 20 (1919), p. 281.

B. Jessen, *The theory of integration in a space of an infinite number of dimensions*, Acta Mathematica, vol. 63 (1934), p. 249.

ing L and M we can establish this, even when K is not strictly convex. We henceforth restrict our attention to systems L, M such that the following condition holds:

(4) If S is any negligible subset of E, and $f(x)$ is any (real) function which vanishes on $E-S$, then $f(x)$ is in the class L and $Mf=0$.

In example (**1**) negligible sets are sets of measure $m_\alpha S=0$; hence condition (4) is satisfied. In example (**2**), a set S of integers is negligible if $\sum_{i\epsilon S} c_i=0$; in (**4**), S is negligible if $p(x)=0$ on almost all of S. In (**8**), S is negligible if $mS=0$. In (**3**), (**6**), and (**7**), only the empty set is negligible. For all of these (4) is valid. I do not know whether example (**5**) satisfies (4).

An immediate consequence of conditions (1), (2), and (4) is that if $f(x)$ is any function of class L, and $g(x)=f(x)$ except on a negligible set, then $g(x)$ is in L and $Mf=Mg$.

THEOREM 5. *If condition* (4) *is satisfied, then in inequality* (J) *equality holds if and only if the following condition holds: For all x except at most those belonging to a negligible set S, the point $(f_1(x), \cdots, f_n(x))$ belongs to a convex subset K' of K on which $\phi(\mathbf{z})$ is linear. In particular, if $\phi(\mathbf{z})$ is strictly convex* equality holds if and only if $f_i(x)=$*const. *except on a negligible set.*

The last statement is an immediate consequence of the first, for if ϕ is strictly convex the only subsets K' on which ϕ is linear consist of single points. Suppose then that the condition of Theorem 5 holds; by redefining $f_i(x)$ on S, it will be true that $\mathbf{f}$ is in K' for all x, without change in $M\mathbf{f}$ or $M(\phi(\mathbf{f}))$. On K' we have $\phi(\mathbf{z})=l(\mathbf{z})+c$; hence $M\phi(\mathbf{f})=M(l(\mathbf{f})+c)=l(M\mathbf{f})+c$. But since K' is convex, by Theorem 1 the point $M\mathbf{f}$ is in K', and so $\phi(M\mathbf{f})=l(M\mathbf{f})+c$. Hence equality holds in the inequality (J).

To prove the necessity of our condition we first observe that there may be linear relations $l(\mathbf{f}(x))+c=0$ holding for all x except those of a negligible set. We choose a maximal set of such linear relations; there is no loss of generality in assuming that these are of the form $f_{s+1}(x)=\cdots=f_n(x)=0$ except on S_1, where S_1 is negligible. Then $Mf_{s+1}=\cdots=Mf_n=0$.

We now change notation. Let R_s be the space of points $(z_1, \cdots, z_s)$; let H be the set of $(z_1, \cdots, z_s)$ such that $(z_1, \cdots, z_s, 0, \cdots, 0)$ is in K; let $\psi(z_1, \cdots, z_s)=\phi(z_1, \cdots, z_s,$

* We assume only that K is convex, not that it is strictly convex.

$0, \cdots, 0)$ on H; in the space R_{s+1} of points $(z_1, \cdots, z_s, z_{n+1})$ let H_1 be the set for which $(z_1, \cdots, z_s)$ is in H and $z_{n+1} \geqq \psi(z_1, \cdots, z_s)$. For x in $E - S_1$ we have $(f_1(x), \cdots, f_s(x))$ in H and $(f_1, \cdots, f_s, \psi(f_1, \cdots, f_s))$ in H_1. If equality holds in (J), then

$$\begin{aligned} M\psi(f_1, \cdots, f_s) &= M\phi(f_1, \cdots, f_s, 0, \cdots, 0) = M\phi(f_1, \cdots, f_n) \\ &= \phi(Mf_1, \cdots, Mf_s, 0, \cdots, 0) \\ &= \psi(Mf_1, \cdots, Mf_s), \end{aligned}$$

so the point $(Mf_1, \cdots, Mf_s, M\psi(f_1, \cdots, f_s))$ is a boundary point of H_1 in R_{s+1}.

By Theorem 3, for all points x of $E - S_1$ except a negligible set S_2 the point $(f_1(x), \cdots, f_s(x), \psi(f_1, \cdots, f_s))$ belongs to the intersection of H_1 with a hyperplane of support π: $a + b_1z_1 + \cdots + b_sz_s + cz_{n+1} = 0$. That is, except on $S_1 + S_2$ the equation $a + b_1f_1 + \cdots + b_sf_s + c\psi(f_1, \cdots, f_s) = 0$ holds. Here $c \neq 0$; otherwise we would have a new linear relation between the f_i independent of the maximal set $f_{s+1} = \cdots = f_n = 0$. We may therefore suppose that $c = 1$; hence π has the equation $a + \sum b_iz_i + z_{n+1} = 0$. The left member of this equation is positive for some points $(z_1, \cdots, z_s, z_{n+1})$ of H_1, since z_{n+1} is arbitrarily large. But π is a hyperplane of support, so $a + \sum b_iz_i + z_{n+1}$ does not change sign on H_1; therefore $a + \sum b_iz_i + z_{n+1} \geqq 0$ on H_1. For $(z_1, \cdots, z_s)$ in H the point $(z_1, \cdots, z_s, \psi(z_1, \cdots, z_s))$ is in H_1, so that $a + \sum b_iz_i + \psi(z_1, \cdots, z_s) \geqq 0$ on H. Moreover, if $z_{n+1} > \psi(z_1, \cdots, z_s)$, then $a + \sum b_iz_i + z_{n+1} > 0$. Hence the points of H_1 which lie in π are those of the form $(z_1, \cdots, z_s, \psi(z_1, \cdots, z_s))$ with $(z_1, \cdots, z_s)$ in H and $a + \sum b_iz_i + \psi(z_1, \cdots, z_s) = 0$. The points satisfying these conditions form the intersection πH_1, which, being the intersection of convex sets, must be convex, as is also its projection H' on the space R_s. For all x in $E - (S_1 + S_2)$ the point $(f_1(x), \cdots, f_s(x), \psi(f_1, \cdots, f_s))$ is in πH_1, so that $(f_1(x), \cdots, f_s(x))$ is in H'. If we define K' to be the set of all points $(z_1, \cdots, z_s, 0, \cdots, 0)$ with $(z_1, \cdots, z_s)$ in H', then K' is convex, and for all x in $E - (S_1 + S_2)$ the point $(f_1(x), \cdots, f_s(x), 0, \cdots, 0)$ is in K'. If $(z_1, \cdots, z_n)$ is in K' then $(z_1, \cdots, z_s)$ is in H', and

$$\begin{aligned} \phi(z_1, \cdots, z_n) &= \phi(z_1, \cdots, z_s, 0, \cdots, 0) \\ &= \psi(z_1, \cdots, z_s) = -a - \sum b_iz_i. \end{aligned}$$

This establishes the theorem.

3. *Extension to Banach Spaces.* It is possible to extend Theorems 1, 2, 3, and 4 to functions $\boldsymbol{f}(x)$ assuming values in a Banach space B. Suppose that (1) and (2) are satisfied, and that $\boldsymbol{L}$ is a class of functions $\boldsymbol{f}(x)$ defined on E and assuming values in B. We shall assume that for every linear function $l(\boldsymbol{z})$ on B and every $\boldsymbol{f}(x)$ in $\boldsymbol{L}$ the function $l(\boldsymbol{f}(x))$ is in L. Further, we shall assume that there is a linear mean $\boldsymbol{M}$ defined on $\boldsymbol{L}$ such that for every linear function $l(\boldsymbol{z})$ on B and every $\boldsymbol{f}$ in $\boldsymbol{L}$ we have

$$l(\boldsymbol{M}\boldsymbol{f}) = M(l(\boldsymbol{f})).$$

We then find that Theorems 1, 2, 3, and 4 extend with only one change. The only properties of convex sets which we used were these: through each boundary point of a convex set there passes a hyperplane of support, and each point which does not belong to a convex set can be separated from it by a hyperplane. These properties have been established for convex *bodies** (closed convex sets having interior points). Hence our theorems extend at once, provided that we replace the words "convex set" by "convex body."

UNIVERSITY OF VIRGINIA

* Cf. Banach, *Théorie des Opérations Linéaires*, p. 246.