PUBLIKACIJE ELEKTROTEHNIČKOG FAKULTETA UNIVERZITETA U BEOGRADU
PUBLICATIONS DE LA FACULTÉ D'ÉLECTROTECHNIQUE DE L'UNIVERSITÉ À BELGRADE

SERIJA: MATEMATIKA I FIZIKA — SÉRIE: MATHÉMATIQUES ET PHYSIQUE

№ 302 — № 319 (1970)

302. THE JENSEN-STEFFENSEN INEQUALITY*

Ralph P. Boas, Jr.

1. Let φ be a continuous convex function over the range of the continuous function f with bounded domain $[a, b]$; then JENSEN's inequality, in a form that includes both the inequality for sums and the inequality for integrals ([2], Theorems 86, 206), states that

$$\varphi\left\{\frac{\int f(x)\, d\lambda(x)}{\int d\lambda(x)}\right\} \leqq \frac{\int \varphi(f(x))\, d\lambda(x)}{\int d\lambda(x)}, \tag{1}$$

where all integrals are over $[a, b]$, provided that λ is increasing (i.e., nondecreasing — I use this convention throughout), and bounded. The requirement that $[a, b]$ is finite is not essential; the necessary modifications in the infinite case are easily supplied. The two familiar special cases come from taking λ to be a stepfunction with positive jumps at integers, or an absolutely continuous function with positive derivative. ("JENSEN's inequality" of [2] is Theorem 19, a quite different inequality.)

It would be reasonable to ask whether the requirement that λ is increasing can be relaxed at the expense of restricting f more severely. An answer was given by STEFFENSEN [4]; see also [3], where many interesting special cases are discussed. In a slightly generalized form, it is as follows.

Jensen-Steffensen inequality. *Inequality* (1) *holds if f is continuous and monotonic* (*in either sense*) *provided that λ is either continuous or of bounded variation, and satisfies*

$$\lambda(a) \leqq \lambda(x) \leqq \lambda(b), \quad \text{all} \quad x \in [a, b];\ \lambda(b) - \lambda(a) > 0. \tag{2}$$

We may regard (2) as a very weak version of monotonicity, namely that λ increases over every set of 3 points that contains both a and b; we might call this (3,2)-monotonicity (3 points, 2 prescribed). A natural generalization is $(2n-1, n)$-monotonicity; that is, λ increases over each set of $2n-1$ points including n prescribed points (two of which are a and b), with the other $n-1$ points lying one in each of the $n-1$ intervals between the prescribed points (and $\lambda(b) > \lambda(a)$). If we strengthen the hypothesis on λ in this way, we can correspondingly weaken the hypothesis on f. The result is as follows.

* Presented April 7, 1970 by D. S. MITRINOVIĆ.

Theorem 1. *Inequality* (1) *holds if* λ *is continuous or of bounded variation and satisfies*

$$\lambda(a) \leqq \lambda(x_1) \leqq \lambda(y_1) \leqq \lambda(x_2) \leqq \cdots \leqq \lambda(y_{n-1}) \leqq \lambda(x_n) \leqq \lambda(b), \qquad n \geqq 2,$$

for all x_k *in* (y_{k-1}, y_k) $(y_0 = a,\ y_n = b)$, *and* $\lambda(b) > \lambda(a)$, *provided that* f *is continuous and monotonic* (*in either sense*) *in each of the* $n-1$ *intervals* (y_{k-1}, y_k).

Thus for example in the (5,3) case

$$\lambda(a) \leqq \lambda(x) \leqq \lambda(h) \leqq \lambda(y) \leqq \lambda(b)$$

provided $a < x < h < y < b$, and f could be (for example) decreasing on (a, h) and increasing on (h, b). In the limit as $n \to \infty$, λ would be increasing and f would be required only to be continuous, so that JENSEN's inequality is a limiting case of the generalized JENSEN-STEFFENSEN inequality, Theorem 1.

The hypothesis on λ in the JENSEN-STEFFENSEN inequality is much weaker than that in JENSEN's inequality; if we try to weaken it still more, say to $\lambda(a) \leqq \lambda(x)$ for $a \leqq x \leqq b$ (and $\lambda(a) < \lambda(b)$), i.e. to (2,1) monotonicity, it can be shown by examples that (1) no longer holds when f is merely monotonic. However, there is a variant that does hold in this case; under additional assumptions on f and λ it provides a weaker conclusion than (1).

To see how this variant arises, let us consider (1) when φ is a power, $\varphi(u) = u^p$ with $p > 1$, and $f(x) \geqq 0$. Then (1) says

$$\left\{ \int_a^b f(x)\, d\lambda(x) \right\}^p \leqq \left\{ \int_a^b d\lambda(x) \right\}^{p-1} \int_a^b f(x)^p\, d\lambda(x).$$

Under either the JENSEN or JENSEN-STEFFENSEN hypotheses the last integral is positive (cf. § 3), so that replacing $\int_a^b d\lambda$ by a larger number preserves the inequality. This remains true whenever φ is continuous and convex and $\varphi(0) \leqq 0$. For, suppose

$$\varphi\left\{ y^{-1} \int_a^b f(x)\, d\lambda(x) \right\} \leqq y^{-1} \int_a^b \varphi(f(x))\, d\lambda(x)$$

for some number y, i.e.

$$\int_a^b \varphi(f(x))\, d\lambda(x) \geqq y\, \varphi\left\{ y^{-1} \int_a^b f(x)\, d\lambda(x) \right\}.$$

This inequality will be preserved for a larger y if the derivative of the right-hand side is negative, i.e.

$$\varphi\left\{ y^{-1} \int_a^b f(x)\, d\lambda(x) \right\} - y^{-1} \varphi'\left\{ y^{-1} \int_a^b f(x)\, d\lambda(x) \right\} \int_a^b f(x)\, d\lambda(x) \leqq 0,$$

i.e., with $\int f\, d\lambda = z$, and $w = z/y$,

$$\varphi(w) - w\varphi'(w) \leqq 0. \qquad (3)$$

(Here φ' denotes either the right-hand or left-hand derivative of φ.) But (3) is true whenever φ is continuous and convex and $\varphi(0)\leqq 0$ (cf. [2], Theorem 129). Indeed,

$$\frac{\varphi(w)}{w} \leqq \frac{\varphi(w)-\varphi(0)}{w} \leqq \sup_{0\leqq t\leqq w} \varphi'(t) \leqq \varphi'(w),$$

since φ' increases. Consequently we have the essentially trivial inequality

$$\varphi\left\{\frac{\int_a^b f(x)\,d\lambda(x)}{\int_a^b d\mu(x)}\right\} \leqq \frac{\int_a^b \varphi(f(x))\,d\lambda(x)}{\int_a^b d\mu(x)} \tag{4}$$

whenever λ and f satisfy the hypotheses of either the JENSEN or JENSEN--STEFFENSEN inequality, $f(x)\geqq 0$, φ is convex and continuous, $\varphi(0)\leqq 0$, $\mu(b)-\mu(a)\geqq \geqq\lambda(b)-\lambda(a)$, and $\mu(b)-\mu(a)>0$.

Although (4) is trivial, it has as a consequence an inequality that applies when we assume only that $\lambda(a)\leqq\lambda(x)$ for $a\leqq x\leqq b$.

Theorem 2. *If $\lambda(a)\leqq\lambda(x)$ for $a\leqq x\leqq b$, $\lambda(a)<\lambda(b)$, and $\lambda(x)\leqq\lambda(b)+\lambda^*$, where $\lambda^*>0$; f decreases and $f(x)\geqq 0$; φ is continuous and convex over $(0, f(a))$, and $\varphi(0)\leqq 0$; and $\int_a^b d\mu(x)\geqq\int_a^b d\lambda(x)+\lambda^*$, then (4) holds.*

If $\lambda^*\leqq 0$ we are back in the JENSEN-STEFFENSEN situation. If we choose $\lambda^*=\sup[\lambda(x)-\lambda(b)]$ we can take μ to be the total variation of λ. In this case (4) was proved by CIESIELSKI [1] under more restrictive hypotheses and by a different method.

It is interesting to note that the JENSEN-STEFFENSEN inequality and Theorem 2 provide verisions of HÖLDER's inequality when the measure $d\lambda$ is not necessarily positive. In particular,

$$\left\{\int_a^b f(x)\,d\lambda(x)\right\}^p \leqq \left\{\int_a^b d\lambda(x)\right\}^{p-1} \int_a^b f(x)^p\,d\lambda(x), \qquad p>1,$$

provided that f decreases, $f(x)\geqq 0$, and $\lambda(a)\leqq\lambda(x)\leqq\lambda(b)$ for $a\leqq x\leqq b$; and

$$\left\{\int_a^b f(x)\,d\lambda(x)\right\}^p \leqq \left\{\int_a^b |d\lambda(x)|\right\}^{p-1} \int_a^b f(x)^p\,d\lambda(x)$$

provided that f decreases, $f(x)\geqq 0$, and $\lambda(a)\leqq\lambda(x)$ for $a\leqq x\leqq b$.

STEFFENSEN's proof of the JENSEN-STEFFENSEN inequality depended on another inequality, now known as "STEFFENSEN's inequality"; for a detailed discussion of this inequality and its applications see [3]. Here I shall give a different proof which is easily adapted to prove Theorem 1. We have already seen that (4) is a corollary of the JENSEN-STEFFENSEN theorem, and Theorem 2 follows as a further corollary (see § 5).

2. We begin by reproducing ZYGMUND's proof of JENSEN's inequality ([**6**], vol. 1, p. 24). Put

$$M(g)=\frac{\int_a^b g(x)\,d\lambda(x)}{\lambda(b)-\lambda(a)},$$

the mean value of g, and note that $M(z)=z$ for a constant z. Since λ increases but is not constant, $M(f)$ is in the range of f, and so $\varphi(M(f))$ is defined. A convex φ is characterized by having a supporting line at each point, i.e.

$$\varphi(y)-\varphi(z)\geqq k(y-z) \tag{5}$$

for all y and z (where k depends on z; in fact, $k=\varphi'(z)$ when $\varphi'(z)$ exists, and k is any number between $\varphi'_+(z)$ and $\varphi'_-(z)$ at the countable set where these are different). Take $z=M(f)$ and $y=f(x)$; then

$$\varphi(f(x))-\varphi(M(f))-k\{f(x)-M(f)\}\geqq 0. \tag{6}$$

Since $d\lambda\geqq 0$, taking mean values preserves the inequality in (6). Hence

$$M[\varphi(f(x))]-\varphi(M(f))\geqq k\{M(f)-M(M(f))\}=0, \tag{7}$$

which is precisely what (1) says.

Note that this proof does not make full use of the convexity of φ: we need only that φ has a supporting line at $M(f)$, so for each particular f and λ we could use some nonconvex functions φ. On the other hand, the proof of the JENSEN-STEFFENSEN inequality will make more essential use of the convexity of φ.

We shall make repeated use of the "second mean-value theorem" for STIELTJES integrals. This states ([**5**], p. 18) that if $f(x)\geqq 0$ and f decreases and $\int_a^b f(x)\,d\lambda(x)$ exists then

$$f(a)\inf_{a\leqq c\leqq b}\int_a^c d\lambda(x)\leqq\int_a^b f(x)\,d\lambda(x)\leqq f(a)\sup_{a\leqq c\leqq b}\int_a^c d\lambda(x). \tag{8}$$

If f increases, the theorem reads

$$f(b)\inf_{a\leqq c\leqq b}\int_c^b d\lambda(x)\leqq\int_a^b f(x)\,d\lambda(x)\leqq f(b)\sup_{a\leqq c\leqq b}\int_c^b d\lambda(x);$$

thus in either case we "take out" f at its largest value.

Inequality (8) is not exactly what is stated in [5], but the proof given there establishes the more general result. For the convenience of the reader, we outline the proof of the right-hand side of (8):

$$\int_a^b f(x)\,d\lambda(x) = \int_a^b f(x)\,d[\lambda(x)-\lambda(a)]$$

$$= f(b)[\lambda(b)-\lambda(a)] - \int_a^b [\lambda(x)-\lambda(a)]\,df(x)$$

$$\leqq f(b) \sup_{a\leqq c\leqq b} \int_a^c d\lambda(x) + \Big\{ \sup_{a\leqq c\leqq b} \int_a^c d\lambda(x)\Big\} \int_a^b |df(x)|$$

$$= f(a) \sup_{a\leqq c\leqq b} \int_a^c d\lambda(x);$$

the lower bound is obtained similarly.

3. We now establish the JENSEN-STEFFENSEN inequality. We first have to show that $M(f)$ is in the range of f, i.e. that when f is (say) decreasing we have $f(b)\leqq M(f)\leqq f(a)$ under the hypothesis that $\lambda(a)\leqq\lambda(x)\leqq\lambda(b)$ if $a<x<b$.

We have $f(x)-f(b)$ decreasing and positive, so by the second mean-value theorem

$$\int_a^b f(x)\,d\lambda(x) - f(b)\int_a^b d\lambda(x) = \int_a^b [f(x)-f(b)]\,d\lambda(x)$$

$$\leqq [f(a)-f(b)] \sup_{a\leqq c\leqq b} [\lambda(c)-\lambda(a)] \leqq [f(a)-f(b)][\lambda(b)-\lambda(a)],$$

i.e. $M(f)\leqq f(a)$. Similarly

$$[\lambda(b)-\lambda(a)][M(f)-f(b)] = \int_a^b [f(x)-f(b)]\,d\lambda(x) \geqq 0,$$

and $M(f)\geqq f(b)$.

Now we can repeat the proof of JENSEN's inequality down to (6), and the proof of the JENSEN-STEFFENSEN inequality reduces to showing that inequality in (6) is preserved under taking mean values.

Although (5) says that the graph of φ is above its supporting line at each point, more than this is true. In fact $\varphi(y)-\varphi(z)-k(y-z)$ is the vertical distance between the graph of φ and the supporting line, and when z is fixed this distance is a decreasing function of y when $y<z$ and an increasing function of y when $y>z$. Let $\Delta(x)$ be the left-hand side of (6); then I claim that whether f decreases or increases, $\Delta(x)$ decreases when $a<x<c=f^{-1}(M(f))$ and $\Delta(x)$ increases when $b>x>c$.

In fact, suppose for example that f decreases and $a<x_1<x_2<c$; then $f(x_1)\geqq f(x_2)\geqq M(f)$, and consequently $\Delta(x_1)\geqq\Delta(x_2)$; i.e., Δ decreases on $a<x<c$. The proof is similar in the other cases.

We now have

$$\int_a^b \Delta(x)\,d\lambda(x) = \int_a^c \Delta(x)\,d\lambda(x) + \int_c^b \Delta(x)\,d\lambda(x).$$

Since $\Delta(x) \geqq 0$, the second mean-value theorem now yields

$$\int_a^b \Delta(x)\, d\lambda(x) \geqq \Delta(a) \inf_{a \leqq \xi \leqq c} \int_a^\xi d\lambda(x) + \Delta(b) \inf_{c \leqq \eta \leqq b} \int_\eta^b d\lambda(x) \geqq 0$$

because of our hypothesis on λ. That is, (7) is true and this is equivalent to (1).

4. The proof of Theorem 1 follows the same lines. For simplicity, we outline the proof for the (5,3) case, when $\lambda(a) \leqq \lambda(x) \leqq \lambda(h) \leqq \lambda(y) \leqq \lambda(b)$ whenever $a<x<h<y<b$; here h is a prescribed point. If f is monotonic on (a, b) we have nothing new, so we assume (for definiteness) that f decreases on (a, h) and increases on (h, b). We again must begin by verifying that $M(f)$ is in the range of f.

Let $\inf f(x) = m$; then $f(x) - m \geqq 0$ and again decreases on (a, h) and increases on (h, b). Then

$$[\lambda(b) - \lambda(a)][M(f) - m] = \int_a^b [f(x) - m]\, d\lambda(x) = \int_a^h + \int_h^b$$

$$\leqq [f(a) - m] \sup_{a \leqq \xi \leqq h} \int_a^\xi d\lambda(x) + [f(b) - m] \sup_{h \leqq \eta \leqq b} \int_\eta^b d\lambda(x).$$

Now for any particular ξ, η we have

$$M(f) - m = \frac{[f(a) - m] \int_a^\xi d\lambda(x) + [f(b) - m] \int_\eta^b d\lambda(x)}{\lambda(b) - \lambda(a)}$$

$$\leqq [f(a) - m] \frac{\lambda(h) - \lambda(a)}{\lambda(b) - \lambda(a)} + [f(b) - m] \frac{\lambda(b) - \lambda(h)}{\lambda(b) - \lambda(a)},$$

and the right-hand side is a convex linear combination of $f(a) - m$ and $f(b) - m$, hence less than the larger of these two numbers; and this remains true after taking least upper bounds on the left. Therefore $M(f)$ does not exceed the larger of $f(a)$, $f(b)$.

Similarly we can show that $M(f) \geqq m = \inf_{a \leqq x \leqq b} f(x)$.

Thus $M(f)$ is indeed in the range of f and we have (6) again. Our inequality will be established if we show that (6) still holds after we take mean values.

We are assuming that f decreases for $a<x<h$ and increases for $h<x<b$. Since $M(f)$ is in the range of f, there is at least one point c such that $f(c) = M(f)$, and there may be two, say c_1 in (a, h) and c_2 in (h, b). We consider the latter case first. The situation on (a, h) is the same as it was on (a, b) in the Jensen-Steffensen case, with f decreasing, so that $\Delta(x)$ decreases on (a, c_1) and increases on (c_1, h). The situation on (h, b) is the

same except that f increases; but this again leads to $\Delta(x)$ decreasing on (h, c_2) and increasing on (c_2, b). Consequently by the second mean-value theorem

$$\int_a^b \Delta(x)\,d\lambda(x) = \int_a^{c_1} + \int_{c_1}^{h} + \int_h^{c_2} + \int_{c_2}^{b}$$

$$\geqq \Delta(a) \inf \int_a^{\xi_1} d\lambda(x) + \Delta(h) \inf \int_{\xi_2}^{h} + \Delta(h) \inf \int_h^{\xi_3} + \Delta(b) \inf \int_{\xi_4}^{b} \geqq 0$$

by our hypotheses on λ.

Finally, if $f(c) = M(f)$ only at one point c, suppose for definiteness that $a < c < h$. Then $\Delta(x)$ decreases on (a, c), increases on (c, h), and decreases on (h, b), and $\int_a^b \Delta(x)\,d\lambda(x) \geqq 0$ as before.

5. We now deduce Theorem 2 from the fact, discussed at the end of § 1, that (4) holds trivially under the hypotheses of Theorem 2 strengthened so that λ satisfies the condition $\lambda(a) \leqq \lambda(x) \leqq \lambda(b)$.

Note first that we may suppose that $f(b) = 0$, instead of $f(b) \geqq 0$. For, if $f(b) > 0$ we replace $[a, b]$ by $[a, b+\varepsilon]$ with $\varepsilon > 0$, extend λ and μ to be constant on $[b, b+\varepsilon]$, and extend f so that f decreases from $f(b)$ to 0 on $[b, b+\varepsilon]$. If we have proved (4) for $[a, b+\varepsilon]$ with $f(b+\varepsilon) = 0$, we have clearly proved (4) for $[a, b]$ with $f(b) > 0$, since the integrals in (4) all reduce to integrals over $[a, b]$.

Now let u be a step function whose only jump is unity at b, let λ^* be a positive number (as specified in Theorem 2) such that $\lambda(x) \leqq \lambda(b) + \lambda^*$ for $a \leqq x \leqq b$, and apply (4) with the JENSEN-STEFFENSEN hypotheses to φ; f with f decreasing, $f(x) \geqq 0$ and $f(b) = 0$; and $\lambda(x) + \lambda^* u(x)$. We then have

$$\int_a^b d\mu(x) \geqq \int_a^b d[\lambda(x) + \lambda^* u(x)]$$

by hypothesis. We get

$$\varphi\left\{\frac{\int_a^b f(x)\,d\lambda(x)}{\int_a^b d\mu(x)}\right\} \leqq \frac{\int_a^b \varphi(f(x))\,d\lambda(x) + \varphi(0)\,\lambda^*}{\int_a^b d\mu(x)}.$$

Since $\varphi(0) \leqq 0$ and $\lambda^* \geqq 0$, we obtain the conclusion of Theorem 2.

REFERENCES

1. Z. Ciesielski: *A note on some inequalities of Jensen's type*. Ann. Polon. Math. **4** (1958), 269—274.

2. G. H. Hardy, J. E. Littlewood and G. Pólya: *Inequalities*. Cambridge 1934.

3. D. S. Mitrinović: *The Steffensen inequality*. Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz. № **247** — № **273** (1969), 1—14.

4. J. F. Steffensen: *On certain inequalities and methods of approximation*. J. Inst. Actuaries **51** (1919), 274—297.

5. D. V. Widder: *The Laplace transform*. Princeton 1941.

6. A. Zygmund: *Trigonometric series*, 2d ed. Cambridge 1959.

Northwestern University
Evanston, Ill. 60201, USA