Mathematical
Inequalities
& Applications
Volume 1, Number 3 (1998), 391–401

THE JENSEN–STEFFENSEN INEQUALITY

P. S. BULLEN

Abstract. In this paper a simple calculus proof is given of the Jensen-Steffensen inequality, and of an inverse inequality due to Pečarić.

1. Introduction and notations

In a previous note [3] the basic geometric properties of convex functions were given a simple calculus proof. Here we extend those arguments to give an equally simple proof of Jensen's inequality, a converse and the extension of Jensen's inequality due to Steffensen; see [1, and 4; pp. 23–30.]

If f is a real-valued function defined on $\mathbb{R}$ it is said to be *strictly convex*, or just *convex*, if

$$f''(x) \geqslant 0, \tag{1}$$

and every sub-interval contains a point x at which $f''(x) > 0$. It is easy to modify this definition to allow for functions that are defined on other kinds of intervals. If $-f$ is convex then we say that f is *concave*.

In most standard references the class of convex functions is defined more generally and then various deductions become more delicate; see [6]. However for applications to inequalities the above definition will suffice.

Only elementary calculus will be used, but it is worth noting a deduction that can be made from the mean value theorem.

The theorem itself is:

If f is continuous on $[a,b]$ and differentiable on $]a,b[$ then for some c, $a < c < b$,

$$f(b) - f(a) = (b-a)f'(c); \tag{2}$$

such a point c will be called a mean-value point of f on $[a,b]$.

Applying (2) to arbitrary subintervals of $[a,b]$ we get the usual consequences:

(i) if f takes the same value twice then at some point in between f' is zero;

(ii) if $f' \geqslant 0$ and every sub-interval contains a point at which $f' > 0$, in particular if $f' \geqslant 0$ with $f'(x) = 0$ at only a finite number of points, then f is strictly increasing.

Mathematics subject classification (1991): 26D15.

Key words and phrases: Inequalities, mean value theorem, convex, arithmetic mean, converse inequality, Jensen, Steffensen.

This last consequence implies that if f is convex then f' is strictly increasing. A simple deduction from this is that $\big(f(y)-f(x)\big)/(y-x)$ is strictly increasing as a function of either x or y.

If $\underline{a} = (a_1, a_2, \ldots)$ is a sequence of real numbers and $\underline{w} = (w_1, w_2, \ldots)$ is a sequence of positive numbers, with $W_n = \sum_{i=1}^{n} w_i$, $n = 1, 2, \ldots$ then

$$A_n(\underline{a};\underline{w}) = \frac{1}{W_n}\sum_{i=1}^{n} w_i a_i, \qquad n = 1, 2, \ldots \tag{3}$$

is the sequence of the *arithmetic means of* $\underline{a}$ *with weight* $\underline{w}$.

For a given $n \geqslant 1$ it is often convenient to put

$$\frac{w_1}{W_n} = 1 - t_2 - \cdots - t_n,\ \frac{w_2}{W_n} = t_2, \ldots, \frac{w_n}{W_n} = t_n.$$

Then $0 < 1 - t_2 - \cdots - t_n < 1$ and $0 < t_i < 1$, $2 \leqslant i \leqslant n$ and (3) will be written

$$A_n(\underline{a}; t_2, \ldots, t_n) = (1 - t_2 - \cdots - t_n)a_1 + t_2 a_2 + \cdots + t_n a_n, .$$

When $n = 2, 3, 4$ the notation is usually changed to avoid suffices.

For suitably defined functions we will write $f(\underline{a})$ for the sequence $(f(a_1), f(a_2), \ldots)$.

2. Jensen's inequality and a converse inequality

If f is convex then an important property is that it satisfies the following inequality due to Jensen: if $n \geqslant 1$ then

$$f\left(\frac{w_1 a_1 + \cdots + w_n a_n}{W_n}\right) \leqslant \frac{w_1 f(a_1) + \cdots + w_n f(a_n)}{W_n}; \tag{J}$$

or in a shorter notation,

$$f\,(A_n(\underline{a};\underline{w})) \leqslant A_n(f(\underline{a});\underline{w});$$

with equality only if $a_1 = a_2 = \cdots a_n$.

Since equality under the conditions stated is trivial, it suffices to prove (J) is strict under the assumption that not all such a_i are equal. Further since if any two a_i are equal this is equivalent to considering (J) for a smaller value of n we can without loss in generality assume when convenient that all the a_i are distinct. In addition we could allow the weights to be zero, provided only that they are not all zero. Again this is equivalent to considering (J) for smaller values of n so there is no loss in assuming if we want that all the weights are positive.

The usual proof of (J) is by an algebraic induction but here we give a simple geometric induction that is based on the method used for the inequality between the arithmetic and geometric means; see [2]. The case $n = 2$ was given in [3] and is repeated here as it is needed for the induction and also because the method is the basis of that induction.

2.1. The case $n = 2$. In this case (J) can be rewritten as

$$f\left(\overline{1-s}\,x+sy\right) \leqslant (1-s)f(x)+sf(y), \quad 0<s<1, \tag{J$_2$}$$

with equality only if $x = y$. To prove (J_2) consider the following function obtained from the difference between its two sides

$$D_2(x,y;s) = D_2(s) = (1-s)f(x)+sf(y)-f\left(\overline{1-s}\,x+sy\right)$$

where $0 \leqslant s \leqslant 1$ and, from the above remarks, $x<y$.

Then (J_2) is equivalent to $D_2(s) > 0$, $0 < s < 1$.

Since $D_2(0) = D_2(1) = 0$ we have, by (i) of section 1, that for some s_0, $0 < s_0 < 1$, $D_2'(s_0) = 0$; then $\overline{1-s_0}\,x + s_0y$ is a mean-value point for f on $[x,y]$. Further

$$\begin{aligned} D_2'(s) &= f(y)-f(x)-(y-x)f'\left(\overline{1-s}\,x+sy\right) \\ D_2''(s) &= -\,(y-x)^2 f''\left(\overline{1-s}\,x+sy\right) \end{aligned} \tag{4}$$

Since $D_2'' \leqslant 0$, and is negative in every subinterval it follows that D is concave and that D_2' is strictly decreasing. Hence s_0 is unique, with D_2' positive to the left of the mean-value point, and negative to the right. Since then D_2 is not constant we have that it is positive except at $s = 0,1$, which proves (J_2).

Of course $D_2(s)$ is defined for all $s \in \mathbb{R}$ and the argument given above shows that D_2' is strictly decreasing for all s. This implies that D_2 is negative outside $[0,1]$; that is

$$D_2(s)\begin{cases} =0 & \text{if } s=0,1; \\ >0 & \text{if } 0<s<1; \\ <0 & \text{if } s<0 \text{ or } 1<s. \end{cases} \tag{5}$$

In addition to proving (J_2) the above argument also proves a converse inequality, because $D_2(s) \leqslant D_2(s_0)$ with equality only if $s = s_0$. So if $0<s<1$,

$$\begin{aligned} 0 \leqslant (1-s)f(x)+sf(y)&-f\left(\overline{1-s}\,x+sy\right) \\ &\leqslant (1-s_0)f(x)+s_0f(y)-f\left(\overline{1-s_0}\,x+s_0y\right) \end{aligned}$$

with equality on the left, (J_2), only if $x = y$, or on the right, a converse of (J_2), if $x=y$ or if $s=s_0$.

Another point we need is to note that if we consider D_2 as a function of x with $0<s<1$ then

$$D_2'(x) = (1-s)\left(f'(x)-f'(\overline{1-s}\,x+sy)\right),$$

which is negative; while if we consider D_2 as a function of y with $0<s<1$ then

$$D_2'(y) = s\left(f'(y)-f'(\overline{1-s}\,x+sy)\right),$$

which is positive. Hence if $x' \leqslant x < y \leqslant y'$ with not both $x'=x$ and $y'=y$ then

$$D_2(x',y';s) > D_2(x,y;s), \quad 0<s<1; \tag{6}$$

in particular the the maximum value of $D_2(x',y';s)$ is larger than that of $D_2(x,y;s)$.

2.2. The case $n = 3$ and the general case. In a similar manner the case $n = 3$ of (J),

$$f\left(\overline{1-s-t}\,x + sy + tz\right) \leqslant (1-s-t)f(x) + sf(y) + tf(z), \tag{J_3}$$

when $0 < s+t < 1, 0 < s < 1, 0 < t < 1$ with equality only if $x = y = z$ can be decided by considering the function,

$$D_3(x,y,z;s,t) = D_3(s,t) = (1-s-t)f(x) + sf(y) + tf(z) - f\left(\overline{1-s-t}\,x + sy + tz\right)$$

where $x < y < z$ and

$$0 \leqslant s \leqslant 1,\ 0 \leqslant t \leqslant 1,\ 0 \leqslant s+t \leqslant 1. \tag{T}$$

Since D_3 is continuous it attains both its maximum and its minimum on (T) and if this occurs in the interior of (T) then it occurs at a turning point. Now

$$\begin{aligned} \frac{\partial}{\partial s} D_3(s,t) &= f(y) - f(x) - f'\left(\overline{1-s-t}\,x + sy + tz\right)(y-x), \\ \frac{\partial}{\partial t} D_3(s,t) &= f(z) - f(x) - f'\left(\overline{1-s-t}\,x + sy + tz\right)(z-x). \end{aligned} \tag{7}$$

So for a turning point at (s,t) we must have that

$$f'\left(\overline{1-s-t}\,x + sy + tz\right) = \frac{f(y) - f(x)}{y-x} = \frac{f(z) - f(x)}{z-x}.$$

By a remark following (ii) in section 1, $\big(f(y) - f(x)\big)/(y-x) < \big(f(z) - f(x)\big)/(z-x)$. So D_3 has no turning points in (T) and attains both its maximum and minimum values on the boundary of (T).

However on a side of (T) the problem reduces to the previous case; for instance when $t = 0$, $D_3(x,y,z;s,0) = D_2(x,y;s)$. Hence D_3 attains its minimum value of zero at the corners of (T), the points $(0,0),(0,1),(1,0)$. By (6) the maximum of D_3 occurs at $(0,t_0)$ where $\overline{1-t_0}\,x + t_0 z$ is the mean-value point for f on $[x,z]$.

Thus we have (J_3) as well as the following converse, if $0 < s < 1, 1 < t < 1, 0 < s+t < 1$ then

$$\begin{aligned} (1-s-t)f(x) + sf(y) + tf(z) - f\left(\overline{1-s-t}\,x + sy + tz\right) \\ \leqslant (1-t_0)f(x) + t_0 f(z) - f\left(\overline{1-t_0}\,x + t_0 z\right), \end{aligned}$$

with equality only if $x = y = z$.

Now it is clear that this argument easily extends to the general case to give

THEOREM 1. *If f is convex and if $a_1, a_2 \ldots, a_n$ are not all equal, $n \geqslant 1$ and if $w_i \geqslant 0, 1 \leqslant i \leqslant n$ with $W_n \neq 0$ then*

(a)

$$f\left(\frac{w_1 a_1 + \cdots + w_n a_n}{W_n}\right) \leqslant \frac{w_1 f(a_1) + \cdots + w_n f(a_n)}{W_n}$$

with equality only if all the a_i with associated non-zero w_i are equal; and

(b)

$$\frac{w_1 f(a_1) + \cdots + w_n f(a_n)}{W_n} - f\left(\frac{w_1 a_1 + \cdots + w_n a_n}{W_n}\right)$$
$$\leqslant (1-t_0) f(m) + t_0 f(M) - f\left(\overline{1-t_0}\, m + t_0 M\right),$$

where m, M are, respectively the smallest and the largest of the a_i with associated non-zero w_i, and where $\overline{1-t_0}\, m + t_0 M$ is the mean-value point for f on $[m, M]$. There is equality in (b) only if either all the a_i with associated non-zero w_i are equal or if all the w_i are zero except those associated with m and M, and then these have weights $1-t_0, t_0$ respectively.

3. The Steffensen inequality

We now turn to the interesting extension of Jensen's inequality due to Steffenesen that allows a certain range of negative weights.

3.1. The case $n = 3$.

While $D_2(s)$ is positive precisely on the interval $]0, s[$, see (5), the function $D_3(s,t)$ is positive on a region larger than the interior of the triangle (T). This is because D_3 is continuous and positive on the whole of (T) except for the corners. Precisely $D_3 \geqslant 0$ in the region bounded by the 0-level curve of D_3 that contains the triangle (T); this curve passes through the corners of (T).

This region depends, in general, on the values of x, y, z. Thus if $x = y$ it is the strip $0 \leqslant t \leqslant 1$, while if $y = z$ it is the strip $0 \leqslant s + t \leqslant 1$. The question to be taken up is to find, if possible, a region larger than (T) that does not depend on x, y, z, as (T) does not so depend, and on which $D \geqslant 0$. The region we are looking for is $S = \cap_{\{(x,y,z); x<y<z\}} \{(s,t); D_3(x,y,z;s,t) \geqslant 0\}$.

If T is a proper subset of S then Jensen's inequality will hold for certain negative values of the weights. This is the result of Steffensen.

It follows from the above that S is a subset of the parallelogram common to the two strips $0 \leqslant t \leqslant 1$ and $0 \leqslant s + t \leqslant 1$; that is the region:

$$0 \leqslant s + t \leqslant 1, \qquad 0 \leqslant t \leqslant 1. \tag{P}$$

Note that this parallelogram, unlike the triangle (T), is not symmetric with respect to the variables s, t and so the condition $x \leqslant y \leqslant z$ will be needed for Steffensen's extension to hold.

Since, as we have observed, D_3 reduces to a D_2 on each of the sides of (T) it follows from the observations made about D_2, (5), that $D_3 < 0$ on the extensions of these sides; that is $D_3(s,t) < 0$ if (i) $t = 0$ and $s < 0$ or $s > 1$; (ii) $s = 0$ and $t < 0$ or $t > 1$ (iii) $s + t = 1$ and $t < 0$ or $t > 1$.

In addition considerations of the partial derivatives of D_3 at the corners of (T) show that $D_3 < 0$ in the regions containing the external angles of (T); that is the regions bounded by two rays on which we have just seen that D_3 is negative. For instance consider the third quadrant which is one of these regions.

$$\begin{aligned}\frac{\partial}{\partial s}D_3(0,0) =& f(y)-f(x)-f'(x)(y-x)\\ =&(y-x)\left(\frac{f(y)-f(x)}{y-x}-f'(x)\right)=(y-x)\big(f'(c)-f'(x)\big)\\ \frac{\partial}{\partial t}D_3(0,0) =& f(z)-f(x)-f'(x)(z-x)\\ =&(z-x)\left(\frac{f(z)-f(x)}{z-x}-f'(x)\right)=(z-x)\big(f'(d)-f'(x)\big),\end{aligned}$$

for some c,d where $x<c,d<z$; in addition using properties of convexity mentioned in section 1, $c<d$. Hence both of these partial derivatives are positive which implies that D_3 is negative in the third quadrant.

The other corners can be handled in a similar manner.

The tangent to the 0-level curve at the origin makes an angle θ_1 with the positive s-axis where

$$\tan\theta_1=-\frac{\partial D_3/\partial s\,(0,0)}{\partial D_3/\partial t\,(0,0)}=-\frac{(y-x)\,(f'(c)-f'(x))}{(z-x)\,(f'(d)-f'(x))},$$

and so

$$-1<\tan\theta_1<0.$$

This implies that the line $s+t=0$ crosses the 0-level curve at the origin, being on the side of (T) when $s<0$.

Similarly the tangent to the 0-level curve at the $(0,1)$ makes an angle θ_2 with the positive s-axis where

$$\tan\theta_2=-\frac{(y-x)\,(f'(c)-f'(z))}{(z-x)\,(f'(d)-f'(z))},,$$

and so

$$-1<\tan\theta_2<0.$$

This implies that the line $s+t=0$ crosses the 0-level curve at this point, being above the curve when $s<0$.

A similar discussion at the point $(1,0)$ leads to an angle with a tangent that is sometimes positive and sometimes negative depending on the values of x,y,z,since there we have, with the obvious notation that

$$\tan\theta_3=-\frac{(y-x)\,(f'(c)-f'(y))}{(z-x)\,(f'(d)-f'(y))},$$

and so as is to be expected this corner is of no interest to us.

At the first two corners we have that at least locally D_3 is positive on the sides of (P). We now show that in fact D_3 is positive on these two sides of (P).

Let us put $\phi(s) = D_3(s,1)$ when

$$\phi'(s) = f(y) - f(x) - f'(-sx+sy+z)(y-x), \qquad \phi''(s) = -f''(-sx+sy+z)(y-x)^2.$$

So ϕ is concave, zero at $s = 0$, with a unique maximum at $s_0 < 0$, where $-s_0x+s_0y+z$ is a mean-value point of f on $[x,y]$.

Now consider $\gamma(s) = D_3(s,-s)$. A similar argument shows that γ is concave, zero at $s = 0$ with a unique maximum at $s_1 < 0$ where $x + s_1y - s_1z$ is a mean-value point of f on $[y,z]$

So finally consider

$$D_3(-1,1) = f(x) - f(y) + f(z) - f(x-y+z) = (f(x) - f(x+h)) - (f(y) - f(y+h)),$$

where $h = z - y$. Hence by the convexity of f, we get that $D_3(-1,1) > 0$.

So D_3 is positive on the sides of (P) except at the corners of (T) and so by the general properties of D_3 we have that on (P) $D_3 \geqslant 0$, being zero only at the corners of (T) ; in addition D_3 attains its maximum value on one of the sides of (P).

Thus in particular we have,on rewriting the inequalities that define (P), that if $x \leqslant y \leqslant z$ then (J_3) holds, if

$$0 < 1 - s - t < 1, 0 < 1 - t < 1,$$

with equality only if $x = y = z$. This is Steffensen's extension of Jensen's inequality in this case; see [1; p.25].

3.2. The case $n = 4$ and the general case. Since, as we have seen, there is no Steffensen extension in the case $n = 2$, the preceding result is the first step in the inductive proof of the Steffensen theorem. As a result, to see how the induction proceeds we will consider the case $n = 4$.

Here the function to look at is,

$$\begin{aligned} D_4(w,x,y,z;s,t,u) = D_4(s,t,u) = (1-s-t-u)f(w) + sf(x) + tf(y) + uf(z) \\ - f\left(\overline{1-s-t-u}\,w + sx + ty + uz\right), \end{aligned}$$

with $w < x < y < z$ and

$$0 \leqslant s+t+u \leqslant 1, \qquad 0 \leqslant t+u \leqslant 1, \qquad 0 \leqslant u \leqslant 1. \tag{S}$$

As before the function D_4 attains both its maximum and minimum values on (S) on the boundary of (S). Unlike the case of (T), and its higher dimensional analogues, the fundamental simplices, the restriction of D_4 to a face of (S) does not immediately reduce to a case of a lower value of n. However as we will see, it is possible on each face to make this reduction in at most two steps.

There are six cases to consider:

$$\text{(I)}\; u = 0; \quad \text{(II)}\; u = 1; \quad \text{(III)}\; t+u = 0;$$
$$\text{(IV)}\; t+u = 1; \quad \text{(V)}\; s+t+u = 0; \quad \text{(VI)}\; s+t+u = 1.$$

3.2.1. Case (I). Here a simple reduction occurs: $D_4(w,x,y,z;s,t,0) = D_3(w,x,y;s,t)$, and $0 \leqslant s+t \leqslant 1$, $0 \leqslant t \leqslant 1$. So on this face $D_4 \geqslant 0$ by the case $n=3$ discussed in section 3.1.

3.2.2. Case (II). In this case $0 \leqslant -s-t \leqslant 1$, $0 \leqslant -t \leqslant 1$ and we have to show that $D_4(w,x,y,z;s,t,1) \geqslant 0$. Now,

$$
\begin{aligned}
(-s-t)f(w)&+sf(x)+tf(y)+f(z)\\
&=(-s-t)f(w)+(s+t)f(x)-tf(x)+tf(y)+f(z),\\
&\geqslant(-s-t)f(w)+(s+t)f(x)+f(-tx+ty+z), \text{ by the case } n=3,\\
&\geqslant f\left(\overline{-s-t}\,w+\overline{s+t}\,x+[-tx+ty+z]\right), \text{ by the case }, n=3,\\
&= f\left(\overline{-s-t}\,w+sx+ty+z\right),
\end{aligned}
$$

which gives this case.

3.2.3. Case (III). Here $0 \leqslant s \leqslant 1$, $0 \leqslant -t \leqslant 1$ and we have to consider $D_4(w,x,y,z;s,t,-t)$. Now

$$
\begin{aligned}
(1-s)f(w)+sf(x)+tf(y)-tf(z) &\geqslant f(\overline{1-s}\,w+sx)+tf(y)-tf(z), \quad \text{by } (\mathrm{J}_2),\\
&\geqslant f(\overline{1-s}\,w+sx]+ty-tz), \text{ by the case } n{=}3,\\
&= f(\overline{1-s}\,w+sx+ty-tz).
\end{aligned}
$$

So $D_4(w,x,y,z;s,t,-t) \geqslant 0$.

3.2.4. Case (IV). Now $0 \leqslant -s \leqslant 1$, $0 \leqslant t \leqslant 1$ and we must show that $D_4(w,x,y,z;s,t,1-t) \geqslant 0$. Now,

$$
\begin{aligned}
-sf(w)+sf(x)+tf(y)&+(1-t)f(z)\\
&\geqslant -sf(w)+sf(x)+f(ty+\overline{1-t}\,z), \text{ by } (\mathrm{J}_2),\\
&\geqslant f(-sw+sx+[ty)+\overline{1-t}\,z]), \text{ by the case } n=3,
\end{aligned}
$$

which gives this case.

3.2.5. Case (V). Here $0 \leqslant -s \leqslant 1, 0 \leqslant -s-t \leqslant 1$ and

$$
\begin{aligned}
f(w)+sf(x)&+tf(y)+(-s-t)f(z)\\
&= f(w)+sf(x)-sf(y)+(s+t)f(y)+(-s-t)f(z),\\
&\geqslant f(w+sx-sy)+(s+t)f(y)+(-s-t)f(z), \text{ by the case } n=3,\\
&\geqslant f([w+sx-sy]+\overline{s+t}\,y+\overline{-s-t}z), \text{ by the case } n=3,\\
&= f(w+sx-sy+ty+\overline{-s-t}z),
\end{aligned}
$$

which show that $D_4(w,x,y,z;s,t,-s-t) \geqslant 0$.

3.2.6. Case (VI). Like Case (I) this case reduces directly to a case where $n=3$ for $D_4(w,x,y,z;1-t-u,t,u)$ is just $D_3(x,y,z;t,u)$ and $0 \leqslant t+u \leqslant 1, 0 \leqslant u \leqslant 1$.

These arguments can readily be extended to higher value of n to give

THEOREM 2. *If* $n > 2$ *inequality (J) holds for real weights provided* $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$ *and the weights satisfy*

$$W_n \neq 0, \qquad \text{and} \quad 0 \leqslant \frac{W_i}{W_n} \leqslant 1, \quad 1 \leqslant i \leqslant n. \tag{8}$$

There is equality only if the a_i *with non-zero weights are all equal.*

4. Some final remarks

4.1. A general induction argument. While the induction needed for Theorem 2 can be carried out from the above arguments the pattern used in the $n = 4$ case may not be immediately clear. So we give the general induction agument here; it can be found in [1] but that journal may not be readily available.

If (8) holds we can assume without loss in generality that $W_n > 0$, when, from (8) , $w_1 \geqslant 0, w_n \geqslant 0$. If every $w_i \geqslant 0, 1 \leqslant i \leqslant n$ then Theorem 2 follows from Theorem 1 so there is nothing to prove. Suppose then that for some $p, 1 < p < n$ we have $w_i \geqslant 0, 1 \leqslant i < p$ and $w_p < 0$. Then, with all the assumptions of Theorem 2:

$$\begin{aligned} A_n\left(f\left(\underline{a}\right);\underline{w}\right) &= \frac{W_{p-1}}{W_n} A_{p-1}\left(f\left(\underline{a}\right);\underline{w}\right) + \frac{1}{W_n}\sum_{i=p}^{n} w_i f\left(a_i\right), \\ &\geqslant \frac{W_{p-1}}{W_n} f\left(A_{p-1}(\underline{a};\underline{w})\right) + \frac{1}{W_n}\sum_{i=p}^{n} w_i f\left(a_i\right), \qquad \text{by (J)}, \\ &= \frac{W_{p-1}}{W_n} f\left(A_{p-1}(\underline{a};\underline{w})\right) - \frac{W_{p-1}}{W_n} f\left(a_p\right) + \frac{W_p}{W_n} f\left(a_p\right) + \frac{1}{W_n}\sum_{i=p+1}^{n} w_i f\left(a_i\right). \end{aligned} \tag{9}$$

Now the coefficients $W_p, w_{p+1}, \ldots, w_n$ satisfy condition (8) and so we can apply the induction hypothesis to last two terms in (9) to get

$$\begin{aligned} A_n\left(f\left(\underline{a}\right);\underline{w}\right) \geqslant{}& \frac{W_{p-1}}{W_n} f\left(A_{p-1}(\underline{a};\underline{w})\right) \\ &- \frac{W_{p-1}}{W_n} f\left(a_p\right) + f\left(\frac{W_p}{W_n} a_p + \frac{1}{W_n}\sum_{i=p+1}^{n} w_i a_i\right). \end{aligned}$$

Finally the coefficients $W_{p-1}/W_n,\ -W_{p-1}/W_n,\ 1$ allow us to apply the case $n = 3$ of Theorem 2 to get (J).

4.2. The case of a general interval. It was stated in section 1 that everything is easily extended to the case of a function that is convex on a general interval. However there is one point that needs to be made. If f is convex on the interval I then of course in Theorems 1,2 $a_i \in I, 1 \leqslant i \leqslant n$, but now we have to show that $A_n(\underline{a};\underline{w}) \in I$. This is obvious in the case of Theorem 1 since it is well known and trivial that in the case of non-negative weights

$$\min_{1\leqslant i\leqslant n} a_i \leqslant A_n(\underline{a};\underline{w}) \leqslant \max_{1\leqslant i\leqslant n} a_i,$$

see for instance [4; p.35]. However this is not so clear in the case of Theorem 2. The following lemma that appears in [1] covers this situation.

LEMMA 3. *If* $a_1 \leqslant \cdots \leqslant a_n$ *then*

$$a_1 \leqslant A_n(\underline{a};\underline{w}) \leqslant a_n,$$

iff the weights satisfy (8).

Proof. By the Abel summation formula

$$A_n(\underline{a};\underline{w}) = a_n - \frac{1}{W_n}\sum_{i=1}^{n-1} W_i(a_{i+1}-a_i),$$

which gives the sufficiency of (8).

Taking $a_1 = \cdots = a_k = -1, a_{k+1} = \cdots = a_n = 0$ for $1 \leqslant k \leqslant n$ gives the necessity of (8). □

4.3. A final remark. It should be remarked that the comments at the beginning of section 3.1 show that Steffensen's extension of (J) is best possible. There is no larger domain of weights than that given by (8) for which (J) can hold for all choices of $\underline{a}$.

5. An inverse inequality

Simple considerations of the case $n = 3$ show that a converse for (J) in the Steffensen extension is not obtainable. The maximum of D_3 in (P) occurs on the boundary, but which section of the boundary depends on the relation between x, y and z.

However we can ask when the inverse of (J) holds, that is (J) with the $\leqslant$ sign replaced by the $\geqslant$ sign; call this inequality $(\sim \mathrm{J})$.

If $n = 2$ then (5) shows that $(\sim \mathrm{J}_2)$ holds if either $s \leqslant 0$, or $s \geqslant 1$. If $n = 3$ the discussion in section 2.2 shows that $(\sim \mathrm{J}_3)$ will hold if either (i) $s+t \geqslant 1, t \geqslant 1$, (ii) $s+t \geqslant 1, t \leqslant 0$, (iii) $s+t \leqslant 0, t \leqslant 0$, or (iv) $s+t \leqslant 0, t \geqslant 1$.

In general Pečarić, [5], has proved the following theorem. First let us write

$$\overline{W}_k = w_n + \cdots + w_k = W_n - W_k,\ 2 \leqslant k \leqslant n; \quad \overline{W}_1 = W_n.$$

THEOREM 4. *If* f *is convex,* $n > 1$, $W_n > 0$ *and* $a_1 \leqslant \cdots \leqslant a_n$ *then* $(\sim J)$ *holds iff for some* m, $1 \leqslant m \leqslant n$,

$$W_k \leqslant 0,\ 1 \leqslant k < m \quad \text{and} \quad \overline{W}_k \leqslant 0,\ m < k \leqslant n. \tag{10}$$

[It is understood that if $m = 1$ then the first condition in (10) is empty and that second is empty if $m = n$.]

Proof.

If $n = 2, 3$ this theorem reduces to the two special cases discussed above. The necessity part of this theorem is given in [5]; here we will give an alternative proof of the sufficiency by modifying the argument of section 4.1.

As the case $n = 2$ is given above let us assume the result is known for all integers less than n and suppose that (10) holds for $m = p$, $1 < p < n$; (the cases of $p = 1$ or $p = n$ are handled in a similar manner but the details are simpler). Then as in section 4.1, but using the induction hypothesis for $(\sim \mathrm{J}_{p-1})$ with $m = p - 1$,

$$A_n\left(f\left(\underline{a}\right);\underline{w}\right) \leqslant \frac{W_{p-1}}{W_n} f\left(A_{p-1}(\underline{a};\underline{w})\right) - \frac{W_{p-1}}{W_n} f\left(a_p\right) + \frac{W_p}{W_n} f\left(a_p\right) + \frac{1}{W_n}\sum_{i=p+1}^{n} w_i f\left(a_i\right). \tag{11}$$

Now the coefficients $W_p, w_{p+1}, \ldots, w_n$ satisfy (10) with $m = 1$ and so again using the induction hypothesis we have from (11),

$$\begin{aligned} A_n\left(f\left(\underline{a}\right);\underline{w}\right) \leqslant{} & \frac{W_{p-1}}{W_n} f\left(A_{p-1}(\underline{a};\underline{w})\right) \\ & - \frac{W_{p-1}}{W_n} f\left(a_p\right) + f\left(\frac{W_p}{W_n} a_p + \frac{1}{W_n}\sum_{i=p+1}^{n} w_i a_i\right). \end{aligned}$$

Finally the coefficients $W_{p-1}/W_n, -W_{p-1}/W_n, 1$ satisfy (10) with $m = 3$ and so using the induction hypothesis again we get $(\sim \mathrm{J})$ holds for the integer n. □

As a final remark note that if f is convex on an interval I we need two extra conditions to ensure the validity of Theorem 4:

(i) $A_n(\underline{a};\underline{w}) \in I$; for obvious reasons;

(ii) $[a_1, a_n] \subset I$, for if $[a_1, a_n] = I$ then, by Lemma 3, (i) cannot hold.

REFERENCES

[1] P. S. BULLEN, *The Steffensen inequality*, Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz. No. **320**–No. **328** (1970), 59–63.

[2] ———, *Another look at some classical inequalities,*, Inequalities and Applications, World Sci. Ser. Appl. Anal. 3, World Sci. Publishing (1944), 127–137.

[3] ———, *Jensen's Inequality*, Menuii Matematika, 1994.

[4] P. S. BULLEN, D. S. MITRINOVIĆ & P. M. VASIĆ, *Means and Their Inequalities*, Reidel Publishing Co. Dordrecht–Boston, 1988.

[5] J. E. PEČARIĆ, *Inverse of Jensen–Steffensen's Inequality*, Glasnik Mat. **(36) 16** (1981), 221–229.

[6] A. W. ROBERTS & D. E. VARBERG, *Convex Functions*, Academic Press, New York–London, 1973.

P. S. Bullen
Department of Mathematics
University of British Columbia
Vancouver, Canada
e-mail: bullen@math.ubc.ca