Mathematical
Inequalities
& Applications
Volume 5, Number 4 (2002), 619–623

THE HERMITE–HADAMARD INEQUALITY FOR CONVEX FUNCTIONS OF A VECTOR VARIABLE

CONSTANTIN P. NICULESCU

(*communicated by A. M. Fink*)

Abstract. The Hermite-Hadamard inequality is discussed in the light of Choquet's theory.

It is well known that every convex function $f : [a,b] \to \mathbf{R}$ can be modified at the endpoints to become convex and continuous. An immediate consequence of this remark is the integrability of f. The *mean value* of f,

$$M(f) = \frac{1}{b-a} \int_a^b f(x)\,dx,$$

can then be estimated by the *Hermite-Hadamard Inequality*,

$$f\left(\frac{a+b}{2}\right) \leqslant M(f) \leqslant \frac{f(a)+f(b)}{2}\,, \tag{HH}$$

which follows easily from the midpoint and trapezoidal approximation to the middle term. Moreover, under the presence of continuity, equality occurs (in either side) only for linear functions. An updated account on (HH) are to be found in [2].

What about the case of functions of several variables? A recent paper by S. S. Dragomir [3] (see also [2]) describes the case of balls in $\mathbf{R}^3$, by proving that

$$f(a) \leqslant \frac{1}{\operatorname{Vol} \overline{B}_R(a)} \iiint_{\overline{B}_R(a)} f(x)\,dV \leqslant \frac{1}{\operatorname{Area} S_R(a)} \iint_{S_R(a)} f(x)\,dS$$

for every continuous convex function $f : \overline{B}_R(a) \to \mathbf{R}$. However, as we shall show in the sequel, more general results are already available in the existing literature. In fact, the right approach of the entire subject of Hermite-Hadamard type inequalities comes from Choquet's theory, a theory whose highlights were presented by R. R. Phelps in his booklet [5]. For a more advanced material, see the monograph of E. M. Alfsen [1].

The basic observation is that the middle point $(a+b)/2$ represents the barycenter of the given interval $[a,b]$ (with respect to a uniform distribution of mass), while the

Mathematics subject classification (2000): Primary 26D07, 26D15, 46A55. Secondary 52A07.
Key words and phrases: convex function, barycenter, extreme point.
Partially supported by CNCSIS Grant A143/2001.

right hand side of (HH) represents the mean value of f over the set of extreme points of the given interval.

Then the two sides of (HH) follow different routes, with different degrees of generality.

To enter the details, let K be a compact convex subset K of a locally convex Hausdorff space E and suppose there is given a Radon probability measure μ on K (which can be thought of as a mass distribution on K). The $\mu-$ *barycenter* of K is defined as the unique point x_μ of K such that

$$x'(x_\mu) = \int_K x'(x)\, d\mu(x) \tag{B}$$

for every continuous linear functional x' on E; see [5], Proposition 1.1. When E is the Euclidean $n-$ dimensional space, the normed and the weak convergence are the same, so that

$$x_\mu = \int_K x\, d\mu(x)$$

i.e., the barycenter coincides with the moment of first order of μ.

An immediate consequence of (B) is the validity of the inequality

$$f(x_\mu) \leqslant \int_K f(x)\, d\mu(x)$$

for every continuous convex function $f : K \to \mathbf{R}$, a fact which extends the left part of the classical Hermite-Hadamard inequality. For details, see the remark before Lemma 4.1 in [5]. Another remark is the following monotonicity property (noticed by S. S. Dragomir [3] in a particular case):

PROPOSITION 1. *Under the above hypothesis, the function*

$$M(t) = \int_K f(tx + (1-t)x_\mu)\ d\mu(x)$$

is convex and nondecreasing on $[0, 1]$.

When $E = \mathbf{R}^n$ and μ is the Lebesgue measure, the value of M at t equals the mean of $f|K_t$, where K_t denotes the image of K through the mapping $x \to tx + (1-t)x_\mu$, i.e.,

$$M(t) = \frac{1}{\mu(K_t)} \int_{K_t} f(x)\, d\mu(x).$$

Proposition 1 tells us that shrinking K to x_μ, via the sets K_t, the mean of $f|K_t$ decreases to $f(x_\mu)$. The proof will need the following approximation argument, which was shown to us by Prof. Gheorghe Bucur:

LEMMA 2. *Every Radon probability measure μ on K is the pointwise limit of a net of discrete Radon probability measures μ_α on K, which have the same barycenter as μ.*

Proof. We have to prove that for every $\varepsilon > 0$ and every finite family $f_1, \ldots, f_n$ of continuous real functions on K there exists a discrete Radon probability measure ν

such that

$$x_\nu = x_\mu \quad \text{and} \quad \sup_{1\leqslant k\leqslant n} |\nu(f_k) - \mu(f_k)| < \varepsilon.$$

As K is compact and convex and the f_k's are continuous, there exists a finite covering $(D_\alpha)_\alpha$ of K by open convex sets such that the oscillation of each of the functions f_k on each set D_α is $< \varepsilon$. Let $(\varphi_\alpha)_\alpha$ be a partition of the unity, subordinated to the covering $(D_\alpha)_\alpha$ and put

$$\nu = \sum\nolimits_\alpha \mu(\varphi_\alpha)\, \varepsilon_{x(\alpha)}$$

where $x(\alpha)$ is the barycenter of the measure $f \to \mu(\varphi_\alpha f)/\mu(\varphi_\alpha)$. As D_α is convex and the support of φ_α is included in D_α, we have $x(\alpha) \in \overline{D}_\alpha$. On the other hand,

$$\mu(h) = \sum\nolimits_\alpha \mu(h\varphi_\alpha) = \sum\nolimits_\alpha \frac{\mu(h\varphi_\alpha)}{\mu(\varphi_\alpha)} \cdot \mu(\varphi_\alpha) = \sum\nolimits_\alpha h(x(\alpha)) \cdot \mu(\varphi_\alpha) = \nu(h)$$

for every continuous affine function $h : K \to \mathbf{R}$. Consequenly, μ and ν have the same barycenter. Finally, for each k,

$$\begin{aligned} |\nu(f_k) - \mu(f_k)| &= \left|\sum\nolimits_\alpha \mu(\varphi_\alpha)\, f_k(x(\alpha)) - \sum\nolimits_\alpha \mu(\varphi_\alpha\, f_k)\right| \\ &= \left|\sum\nolimits_\alpha \mu(\varphi_\alpha) \left[f_k(x(\alpha)) - \frac{\mu(\varphi_\alpha\, f_k)}{\mu(\varphi_\alpha)}\right]\right| \\ &\leqslant \varepsilon \cdot \sum\nolimits_\alpha \mu(\varphi_\alpha) = \varepsilon. \quad \square \end{aligned}$$

Proof of Proposition 1. A straightforward computation shows that $M(t)$ is convex and $M(t) \leqslant M(1)$. Then, assuming the inequality $M(0) \leqslant M(t)$, from the convexity of $M(t)$ we infer

$$\frac{M(t) - M(s)}{t - s} \geqslant \frac{M(s) - M(0)}{s} \geqslant 0$$

for $0 \leqslant s < t \leqslant 1$ i.e., $M(t)$ is nondecreasing. To end the proof, it remains to show that $M(t) \geqslant M(0) = f(x_\mu)$. For, choose a net $(\mu_\alpha)_\alpha$ of discrete Radon probability measures on K, as in Lemma 2 above. Clearly,

$$f(x_\mu) \leqslant \int_K f(tx + (1-t)x_\mu)\, d\mu_\alpha(x) \quad \text{for all } \alpha$$

and thus the desired conclusion follows by passing to the limit over α. $\square$

The extension of the right hand inequality in (HH) is a bit more subtle and makes the object of Choquet's theory, briefly summarized in the sequel. Given two Radon probability measures μ and λ on K, we say that μ is *majorized* by λ (i.e., $\mu \prec \lambda$) if

$$\int_K f(x)\, d\mu(x) \leqslant \int_K f(x)\, d\lambda(x)$$

for every continuous convex function $f : K \to \mathbf{R}$. As noticed in [5], $\prec$ is a partial ordering on the set of all Radon probability measures on K.

THE CHOQUET THEOREM 3. ([5], ch. 3). *Let μ be a Radon probability measure on a metrizable compact convex subset K of a locally convex Hausdorff space E. Then there exists a maximal Radon probability measure $\lambda \succ \mu$ such that the following two conditions are verified:*

i) The barycenter of K with respect to λ and μ is the same;

ii) The set $\mathscr{E}xt\ K$ of all extremal points of K is a $G_\delta-$ subset of K and λ is concentrated on $\mathscr{E}xt\ K$ (i.e., $\lambda(K \setminus \mathscr{E}xt\, K) = 0$).

Under the hypotheses of the above result we get

$$f(x_\mu) \leqslant \int_K f(x)\, d\mu(x) \leqslant \int_{\mathscr{E}xt\, K} f(x)\, d\lambda(x) \tag{Ch}$$

for every continuous convex function $f : K \to \mathbf{R}$, a fact which represents a full extension of (HH) in the case of *metrizable* compact convex sets. Notice that the right part of (Ch) reflects the *maximum principle* for convex functions.

In general, λ is not unique, except for the case of simplices; see [5], ch. 9.

Another useful remark is that every Radon probability measure λ, concentrated on $\mathscr{E}xt\, K$, for which (Ch) holds, is maximal. Cf. [5], Corollary 9.8.

According to the above discussion, if $K = [a, b]$, then necessarily λ is a convex combination of the Dirac measures ε_a and ε_b, say $\lambda = (1-\alpha)\varepsilon_a + \alpha\varepsilon_b$. This remark yields Fink's Hermite-Hadamard type inequality [4] in the case of probability measures:

$$\int_a^b f(x)\, d\mu(x) \leqslant \frac{b - x_\mu}{b-a} \cdot f(a) + \frac{x_\mu - a}{b-a} \cdot f(b) \tag{F}$$

for every continuous convex functions $f : [a, b] \to \mathbf{R}$ and every Radon probability measure μ on $[a, b]$; as usually, x_μ denotes the barycenter of μ, i.e, $x_\mu = \int_a^b x\, d\mu(x)$. In fact, checking

$$\int_a^b f(x)\, d\mu(x) \leqslant (1-\alpha) \cdot f(a) + \alpha \cdot f(b)$$

for $f(x) = (x-a)/(b-a)$ and $f(x) = (b-x)/(b-a)$ we obtain

$$\alpha \geqslant \frac{x_\mu - a}{b-a} \quad \text{and respectively } 1-\alpha \geqslant \frac{b - x_\mu}{b-a}$$

i.e., $\alpha = (x_\mu - a)/(b-a)$.

The argument above can be extended easily for all continuous convex functions defined on $n-$dimensional simplices $K = [A_0, A_1, \ldots, A_n]$ in $\mathbf{R}^n$. Then the corresponding analogue of (F) for Radon probability measures μ *on* K will read as

$$f(X_\mu) \leqslant \int_K f(x)\, d\mu \leqslant \sum_{k=0}^n \mathrm{Vol}_n\left([A_0, A_1, \ldots, \widehat{A_k}, \ldots, A_n] \cdot f(A_k);\right.$$

here X_μ denotes the barycenter of μ, and $[A_0, A_1, \ldots, \widehat{A_k}, \ldots, A_n]$ denotes the sub-simplex obtained by replacing A_k by X_μ; this is the sub-simplex opposite to A_k, when adding X_μ as a new vertex. Vol_n represents the Lebesgue measure in $\mathbf{R}^n$.

In the case of closed balls $K = \overline{B}_R(a)$ in $\mathbf{R}^3$, $\mathscr{E}xt\,K$ coincides with the sphere $S_R(a)$; the paper by Dragomir [3] illustrates the aforementioned theorem of Choquet in the case where μ is the normalized Lebesgue measure on $\overline{B}_R(a)$. His argument, based on Calculus, avoids Choquet's theory, but it cannot be extended to arbitrary compact convex sets and arbitrary Radon probability measures on them.

The Choquet theory is today a well established subject in Mathematics, with many extensions and ramifications, and Theorem 3 above is just the beginning of the story. The reader will find much fun formulating many other results in the Choquet theory as Hermite-Hadamard type inequalities.

REFERENCES

[1] E. M. ALFSEN, *Compact convex sets and boundary integrals*, Springer–Verlag, Berlin, 1971.
[2] C. E. M. PEARCE AND S. S. DRAGOMIR, *Selected topics on Hermite-Hadamard Inequality and Applications*, Victoria University, Melbourne, 2000.
[3] S. S. DRAGOMIR, *On Hadamard's inequality for the convex mappings defined on a ball in the space and applications*, Math. Inequal. & Appl., **3** (2000), 177–187.
[4] A. M. FINK, *A best possible Hadamard inequality*, Math. Inequal. & Appl., **1** (1998), 223–230.
[5] R. R. PHELPS, *Lectures on Choquet's Theorem*, D. van Nostrand Company Inc., Princeton, 1966.

(Received July 21, 2001)

University of Craiova
Department of Mathematics
Craiova 1100, ROMANIA
e-mail: `tempus@oltenia.ro`