School of Mathematical Sciences, Lahore

PROBLEM SOLVING SEMINAR

February 12, 2007

A Journey into Convex Functions Theory

through the Hermite-Hadamard Inequality

Constantin P. Niculescu

---

[0] University of Craiova, Department of Mathematics, Craiova 200585, ROMANIA
E-mail address: cniculescu@central.ucv.ro
URL address: http://www.inf.ucv.ro/~niculescu

A great discovery solves a great problem, but there is a grain of discovery in the solution of any problem. Your problem may be modest, but if it challenges your curiosity and brings into play your inventive faculties, and if you solve it by your own means, you may experience the tension and enjoy the triumph of discovery.

George Polya

Topics

1. From elementary to higher mathematics:

The Hermite-Hadamard inequality at first glance

2. The highlights of Choquet's theory

The Hermite-Hadamard inequality from the point of view of higher mathematics.
The role played by the law of the lever

3. Choquet's theory for signed measures

A bunch of old and new inequalities in the 1-dimensional case. What about the higher dimensional case?

4. Looking for alternative theories

1 The Hermite-Hadamard inequality[1]

The *Hermite-Hadamard inequality*: If $f : [a,b] \to \mathbb{R}$ is a continuous convex function, then

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(x)\,dx \le \frac{f(a)+f(b)}{2}\,. \qquad \text{(HH)}$$

Equality holds in either side only for affine functions (i.e., for functions of the form $mx+n$).

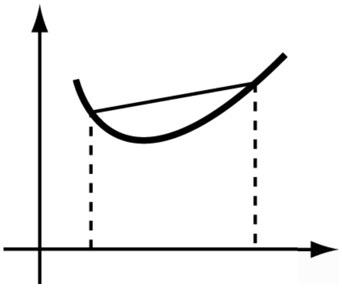

Figure 1: The geometric meaning of convexity.

(RHH): can be discarded from the picture.

For (LHH):

$$\begin{aligned}\frac{1}{b-a}\int_a^b f\,dx &= \frac{1}{b-a}\left(\int_a^{(a+b)/2} f\,dx + \int_{(a+b)/2}^b f\,dx\right)\\ &= \frac{1}{2}\int_0^1 [f\left(\frac{a+b-t(b-a)}{2}\right) + f\left(\frac{a+b+t(b-a)}{2}\right)]dt\\ &\ge f\left(\frac{a+b}{2}\right).\end{aligned}$$

---

[1] This inequality was discovered by Ch. Hermite [11] in 1883 (ten years before Hadamard [8]) but his priority remained unnoticed until 1985. See [13] for a full account of this story.

Applications

For $f = \exp$, (HH) yields

$$e^{(a+b)/2} < \frac{e^b - e^a}{b-a} < \frac{e^a + e^b}{2} \quad \text{for } a \neq b$$

that is,

$$\sqrt{xy} < \frac{x-y}{\log x - \log y} < \frac{x+y}{2} \quad \text{for } 0 < x < y, \tag{GLA}$$

which represents the Geometric, Logarithmic and Arithmetic Mean Inequality.

(Ch. Hermite). For $f(x) = 1/x$, $x \in [n, n+1]$, the inequality (HH) reads as

$$\frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right) > \log(n+1) - \log n > \frac{1}{n+1/2},$$

a fact which is instrumental in deriving the Stirling's formula,

$$n! \sim \sqrt{2\pi} \cdot n^{n+1/2} e^{-n}.$$

For $f(x) = \sin x$, $x \in [0, \pi]$ we obtain

$$\frac{\sin a + \sin b}{2} < \frac{\cos a - \cos b}{b-a} < \sin\left(\frac{a+b}{2}\right),$$

which is equivalent to

$$\tan x > x > \sin x,$$

for $x \in (0, \pi/2)$.

2 An alternative approach to (LHH)

It is possible to prove (LHH) by a dual geometric argument, the existence of support lines at each interior point.

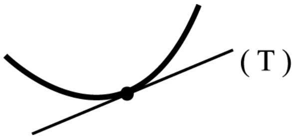

Figure 2: The existence of support lines at interior points

L. Galvani [7]: a real-valued function f is convex on an interval I if and only if for each $c \in I$, the function

$$s_c : x \to \quad \frac{f(x) - f(c)}{x - c} \tag{G}$$

is increasing on $I \setminus \{c\}$. This yields the existence of the lateral derivatives at all interior points of I. Moreover, for x < y in the interior of I we have

$$f'(x-0) \leq f'(x+0) \leq f'(y-0) \leq f'(y+0).$$

Lemma 1 *A function $f : (a, b) \to \mathbb{R}$ is convex if and only if for each point $c \in (a, b)$ there exists a number λ such that*

$$f(x) \geq f(c) + \lambda(x - c) \quad \textit{for every } x \in [a, b]. \tag{SL}$$

Moreover, λ can be chosen arbitrarily in the interval $[f'(c-0), f'(c+0)]$.

Thus (LHH) extends as

$$\frac{1}{b-a} \int_a^b f(x)\, dx \geq f(c) + \lambda \cdot \frac{b+a-2c}{2},$$

for any $c \in (a, b)$. Equality occurs for linear functions (i.e., for $f = f(c) + \lambda(x - c)$ on (a, b)).

3 The precision in the Hermite-Hadamard inequality

For functions of class C^2:

Theorem 2 *Let $f : [a,b] \to \mathbb{R}$ be a twice differentiable function such that there exist real constants m and M so that*

$$m \leq f'' \leq M.$$

Then

$$m \cdot \frac{(b-a)^2}{24} \leq \frac{1}{b-a} \int_a^b f(x)\,dx - f\left(\frac{a+b}{2}\right) \leq M \cdot \frac{(b-a)^2}{24},$$

and

$$m \cdot \frac{(b-a)^2}{12} \leq \frac{f(a)+f(b)}{2} - \frac{1}{b-a} \int_a^b f(x)\,dx \leq M \cdot \frac{(b-a)^2}{12}.$$

Estimates in the framework of Lipschitz functions $f : [a,b] \to \mathbb{R}$,

$$||f||_{Lip} := \sup\left\{\left|\frac{f(x)-f(y)}{x-y}\right|;\ x \neq y\right\} = M < \infty.$$

- *Inequality of Ostrowski,*

$$\left|f(x) - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \left[\frac{1}{4} + \left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right] M(b-a); \tag{O}$$

- *Inequality of Iyengar,*

$$\left|\frac{f(a)+f(b)}{2} - \frac{1}{b-a}\int_a^b f(t)\,dt\right| \leq \frac{M(b-a)}{4} - \frac{1}{4M(b-a)}\left(f(b)-f(a)\right)^2. \tag{I}$$

The proof of (O):

$$\begin{aligned}\left|f(x)-\frac{1}{b-a}\int_a^b f(t)\,dt\right| &= \left|\frac{1}{b-a}\int_a^b (\,f(x)-f(t))\;dt\right| \\ &\leq \frac{1}{b-a}\int_a^b |f(x)-f(t)|\;dt \\ &\leq \frac{M}{b-a}\int_a^b |x-t|\;dt \\ &= \left[\frac{1}{4}+\left(\frac{x-\frac{a+b}{2}}{b-a}\right)^2\right] M(b-a)\end{aligned}$$

for every $x \in [a,b]$.

4 Some straightforward improvements on the Hermite-Hadamard inequality

Suppose that $f : [a, b] \to \mathbb{R}$ is a convex function. By applying the Hermite-Hadamard inequality on each of the intervals $[a, (a+b)/2]$ and $[(a+b)/2, b]$ we get

$$f\left(\frac{3a+b}{4}\right) \leq \frac{2}{b-a}\int_a^{(a+b)/2} f(x)\,dx \leq \frac{1}{2}\left(f(a) + f(\frac{a+b}{2})\right)$$

and

$$f\left(\frac{a+3b}{4}\right) \leq \frac{2}{b-a}\int_{(a+b)/2}^{b} f(x)\,dx \leq \frac{1}{2}\left(f(\frac{a+b}{2}) + f(b)\right),$$

so summing up (side by side) we obtain the following refinement of (HH):

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &\leq \frac{1}{2}\left(f\left(\frac{3a+b}{4}\right) + f\left(\frac{a+3b}{4}\right)\right) \\ &\leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq \frac{1}{2}\left[f(\frac{a+b}{2}) + \frac{f(a)+f(b)}{2}\right] \\ &\leq \frac{1}{2}\left(f(a) + f(b)\right). \qquad \text{(IHH)} \end{aligned}$$

By continuing the division process, the integral mean of f can be approximated as well as we want by the mean values of f at the dyadic points of $[a, b]$.

Remark. *If $f : [a, b] \to \mathbb{R}$ is a convex function, then*

$$\frac{1}{2}\left(f\left(\frac{a+b}{2} - c\right) + f\left(\frac{a+b}{2} + c\right)\right) \leq \frac{1}{b-a}\int_a^b f(x)\,dx$$

for every $c \in [0, (b-a)/4]$, and $c = (b-a)/4$ is maximal with this property.

5 Various quadrature formulae as sources of Hermite-Hadamard type inequalities

By applying twice the integration by parts $f \in C^2([a,b],\mathbb{R})$ we get the following quadrature formula,

$$\frac{1}{b-a}\int_a^b f(x)\,dx = \frac{1}{2}\left[f(a)+f(b)\right] - \frac{1}{b-a}\int_a^b f''(x)\frac{(b-x)(x-a)}{2}\,dx$$

which implies (RHH).

In the case of C^4-functions,

$$\begin{aligned}\frac{1}{b-a}\int_a^b f(x)\,dx &= \frac{1}{8}\left[f(a)+3f\left(\frac{3a+b}{4}\right)+3f\left(\frac{a+3b}{4}\right)+f(b)\right]\\ &\quad -\frac{1}{b-a}\int_a^b f^{(iv)}(x)\varphi(x)\,dx,\end{aligned}$$

where φ is a piecewise polynomial nonnegative function such that

$$\frac{1}{b-a}\int_a^b \varphi(x)\,dx = \frac{(b-a)^4}{6480}.$$

This yields the following improvement on (RHH) for functions $f \in C^4([a,b])$ with $f^{(iv)} \geq 0$:

$$\begin{aligned}\frac{1}{b-a}\int_a^b f(x)\,dx &\leq \frac{1}{8}\left[f(a)+3f\left(\frac{3a+b}{4}\right)+3f\left(\frac{a+3b}{4}\right)+f(b)\right]\\ &\leq \frac{f(a)+f(b)}{2}. \qquad \text{(RHH}^*\text{)}\end{aligned}$$

Examples. *i*) (F. Burk [3]) For $f(x) = e^x|[\ln a, \ln b])$ the inequality (RHH*) yields the following inequality of Tung Po-Lin [27]:

$$\frac{b-a}{\ln b - \ln a} < \left(\frac{a^{1/3}+b^{1/3}}{2}\right)^3. \qquad \text{(LP}_3\text{)}$$

ii) In the case of $f(x) = \ln x$, $x \in [a, b]$, we are led to the following upper estimate for the *identric mean*:

$$\frac{1}{e}\left(\frac{b^b}{a^a}\right)^{1/(b-a)} \leq a^{1/8}\left(\frac{3a+b}{4}\right)^{3/8}\left(\frac{a+3b}{4}\right)^{3/8} b^{1/8},$$

and, according to the AG-inequality, this is a refinement of the last inequality in the following chain of inequalities,

$$\sqrt{ab} < \frac{b-a}{\ln b - \ln a} < \frac{1}{e}\left(\frac{b^b}{a^a}\right)^{1/(b-a)} < \frac{a+b}{2} \tag{GLIA}$$

valid for $a, b > 0$, $a \neq b$. See [1] and compare also the second inequality with (LP_3).

The theory of higher-order convexity, initiated by T. Popoviciu [22], [24], in 1934. See also [20]. A function $f : [a, b] \to \mathbb{R}$ is said to be *n-convex* ($n \in \mathbf{N}$) if for all choices of $n + 1$ distinct points $x_0 < ... < x_n$ in $[a, b]$, the nth order divided difference of f verifies

$$f[x_0, ..., x_n] \geq 0.$$

The divided differences are given inductively by

$$\begin{aligned} f\left[x_0, x_1\right] &= \frac{f(x_0) - f(x_1)}{x_0 - x_1} \\ f\left[x_0, x_1, x_2\right] &= \frac{f\left[x_0, x_1\right] - f\left[x_1, x_2\right]}{x_0 - x_2} \\ &\dots \\ f[x_0, ..., x_n] &= \frac{f\left[x_0, ..., x_{n-1}\right] - f\left[x_1, ..., x_n\right]}{x_0 - x_n}. \end{aligned}$$

As Popoviciu noticed, if f is n-times differentiable with $f^{(n)} \geq 0$ then f is n-convex.

6 The case of subharmonic functions

The idea of integrating by parts (used at the beginning of this section) can be adapted to the several variables setting, via the Green formula. This leads us to Hermite-Hadamard type formulae for subharmonic functions.

Let Ω be a bounded open subset of $\mathbb{R}^n$ with smooth boundary. Then the Dirichlet problem

$$\begin{cases} \Delta\varphi = 1 \text{ on } \Omega \\ \varphi = 0 \text{ on } \partial\Omega \end{cases} \tag{DP}$$

has a unique solution, which is < 0 on Ω (according to the maximum principle).

By Green's formula, for every $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ we have

$$\int_\Omega \begin{vmatrix} u & \varphi \\ \Delta u & \Delta\varphi \end{vmatrix} dV = \int_{\partial\Omega} \begin{vmatrix} u & \varphi \\ \nabla u & \nabla\varphi \end{vmatrix} \cdot n \, dS$$

i.e., in view of (DP),

$$\begin{aligned} \int_\Omega u \, dV &= \int_\Omega u\Delta\varphi \, dV \\ &= \int_\Omega \varphi\Delta u \, dV + \int_{\partial\Omega} u \left(\nabla\varphi \cdot n\right) \, dS - \int_{\partial\Omega} \varphi \left(\nabla u \cdot n\right) \, dS \\ &= \int_\Omega \varphi\Delta u \, dV + \int_{\partial\Omega} u \left(\nabla\varphi \cdot n\right) \, dS \end{aligned}$$

for every $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$. We are then led to the following Hermite-Hadamard type inequality:

Theorem 3 (*See* [17])*If $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ is subharmonic (i.e., $\Delta u \geq 0$ on Ω) and φ satisfies (DP), then*

$$\int_\Omega u \, dV < \int_{\partial\Omega} u \, \left(\nabla\varphi \cdot n\right) dS$$

except for harmonic functions (when equality occurs) .

The equality case needs the remark that $\int_\Omega \varphi \Delta u \, dV = 0$ yields $\varphi \Delta u = 0$ on Ω, and thus $\Delta u = 0$ on Ω; notice that $\varphi \Delta u$ is continuous and nonpositive since $\varphi < 0$ on Ω.

By using the same technique we also see that the following multidimensional version of Theorem 2 holds:

Theorem 4 (*See* [17]) *Let $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ be such that $m \leq \Delta u \leq M$ on Ω and let φ satisfy (DP). Then*

$$m \int_\Omega \varphi \, dV \leq \int_\Omega u \, dV - \int_{\partial\Omega} u \left(\nabla\varphi \cdot n\right) \, dS \leq M \int_\Omega \varphi \, dV.$$

In the case of balls in $\mathbb{R}^3$, the conclusion of Theorem 3 reads as follows:

$$u(a) \leq \frac{1}{Vol\, \overline{B}_R(a)} \iiint_{\overline{B}_R(a)} u(x) \, dV < \frac{1}{Area\, S_R(a)} \iint_{S_R(a)} u(x) \, dS \tag{VS}$$

for every $u \in C^2(B_R(a)) \cap C^1(\overline{B}_R(a))$ with $\Delta u \geq 0$, which is not harmonic. In fact, in this case

$$\varphi(x) = \frac{1}{6} \left(|x|^2 - R^2\right)$$

satisfies (DP) and

$$\nabla\varphi \cdot n = x/3 \cdot x/|x| = R/3.$$

The inequality (VS) works also for all convex functions on balls $\overline{B}_R(a)$.

7 Choquet's Theory

Let K be a compact convex subset K of a locally convex Hausdorff space E and suppose there is given a Borel probability measure μ on K (which can be thought of as a mass distribution on K). The *μ-barycenter* of K : 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. When E is the Euclidean n-dimensional space,

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

The general (LHH) inequality:

Lemma 5 *Let K and μ be defined as above. Then, for every continuous convex function $f : K \to \mathbb{R}$,*

$$f(x_\mu) \leq \int_K f(x)\,d\mu(x).$$

Details in [14] and [21].

The extension of the right-hand inequality in (HH).

The notion of majorization. G. H. Hardy, J. E. Littlewood and G. Pólya [9], [10]. Given two Borel probability measures μ and λ on K, we say that μ is *majorized* by λ (denoted $\mu \prec \lambda$) if

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

for every continuous convex function $f : K \to \mathbb{R}$. The relation $\prec$ is a partial ordering on the set of all Borel probability measures on K.

Theorem 6 (*G. Choquet*; *see* [21], [18]). *Suppose that K is a metrizable compact convex set (in a locally convex Hausdorff space). Then the set Ext K of all extreme points of K is a $G_\delta-$subset of K and for every Borel probability measure μ on K there exists a Borel probability measure λ on K supported by* Ext K (*i.e.*, $\lambda(K \setminus \text{Ext } K) = 0$) *such that*

$$f(x_\mu) \leq \int_K f(x)\,d\mu(x) \leq \int_{Ext\ K} f(x)\,d\lambda(x) \tag{Ch}$$

for every continuous convex function $f : K \to \mathbb{R}$.

When K is the interval $[a, b]$ endowed with the normalized Lebesgue measure $dx/(b-a)$, then x_μ is exactly the midpoint $(a+b)/2$ and Ext $K = \{a, b\}$. Any probability measure λ concentrated by Ext K is necessarily a convex combination of Dirac measures i.e.,

$$\lambda = \alpha\delta_a + (1-\alpha)\delta_b$$

for some $\alpha \in [0, 1]$. Checking the right side inequality in (Ch) for $f = x - a$ and $f = b - x$ we get

$$1 - \alpha \geq 1/2 \text{ and } \alpha \geq 1/2$$

i.e., $\alpha = 1/2$. Consequently, in this case (Ch) coincides with (HH).

The connection between the Choquet theory and the Hermite-Hadamard inequality was first noticed by Niculescu in 2001. See [14].

8 Proof of Choquet's Theorem

The *lower envelope of* a function f in $C(K, \mathbb{R})$ is given by

$$\underline{f}(x) = \sup\{h(x) : h \in A(K, \mathbb{R}) \text{ and } f \geq h\}$$

while its *upper envelope* is

$$\overline{f}(x) = \inf\{h(x) : h \in A(K, \mathbb{R}) \text{ and } h \geq f\}.$$

Clearly

$$\underline{f} = -\overline{(-f)},$$

so it suffices to investigate the properties of one type of envelopes, say the upper ones.

Lemma 7 (See [18], [21]) The upper envelope $\overline{f}$ is concave, bounded and upper semicontinuous (that is, for every real number α, *the set* $\{x : \overline{f}(x) < \alpha\}$ is open*).* Moreover*:*
i) $f \leq \overline{f}$ and $f = \overline{f}$ if f is concave*.*
ii) $\overline{f+h} = \overline{f} + h$ if $h \in A(K, \mathbb{R})$;
iii) The map $f \to \overline{f}$ is sublinear.

The Proof of Choquet's Theorem will be done in four steps.

Step 1. We start by proving that $\operatorname{Ext} K$ is a countable intersection of open sets (and thus it is a Borel set). Here the assumption on metrizability is essential.

Suppose that the topology of K is given by the metric d and for each integer $n \geq 1$ consider the set

$$K_n = \left\{x : x = \frac{y+z}{2}, \text{ with } y, z \in K \text{ and } d(y,z) \geq 1/2^n\right\}.$$

Clearly, $\operatorname{Ext} K = K \setminus \bigcup_n K_n$ and an easy compactness argument shows that each K_n is closed. Consequently, $\operatorname{Ext} K = \bigcap_n \complement K_n$ is a countable intersection of open sets.

Step 2. We may choose a maximal Borel probability measure $\lambda \succ \mu$. To show that Zorn's lemma may be applied, consider a chain $\mathcal{C} = (\lambda_\alpha)_\alpha$ in

$$\mathcal{P} = \{\lambda : \lambda \succ \mu, \lambda \text{ Borel probability measure on } K\}.$$

As $(\lambda_\alpha)_\alpha$ is contained in the weak-star compact set

$$\{\lambda : \lambda \in C(K), \lambda \geq 0, \ \lambda(1) = 1\},$$

by a compactness argument we may find a subnet $(\lambda_\beta)_\beta$ which converges to a measure $\tilde{\lambda}$ in the weak-star topology. A moment's reflection shows that $\tilde{\lambda}$ is an upper bound for $\mathcal{C}$. Consequently, we may apply Zorn's lemma to choose a maximal Borel probability measure $\lambda \succ \mu$. It remains to prove that λ does the job.

Step 3. Since K is metrizable, it follows that $C(K)$ (and thus $A(K)$) is separable. This is a consequence of Urysohn's lemma in general topology. Every sequence $(h_n)_n$ of affine functions with $\|h_n\| = 1$, which is dense in the unit sphere of $A(K)$, separates the points of K in the sense that for every $x \neq y$ in K there is an h_n such that $h_n(x) \neq h_n(y)$. Consequently, the function

$$\varphi = \sum_{n=1}^{\infty} 2^{-n} h_n^2$$

is continuous and strictly convex, from which it follows that

$$\mathcal{E} = \{x : \varphi(x) = \overline{\varphi}(x)\} \subset \text{Ext } K.$$

In fact, if $x = (y+z)/2$, where y and z are distinct points of K, then the strict convexity of φ implies that

$$\varphi(x) < \frac{\varphi(y) + \varphi(z)}{2} \leq \frac{\overline{\varphi}(y) + \overline{\varphi}(z)}{2} \leq \overline{\varphi}(x).$$

Step 4. As a consequence of the maximality of λ, we shall show that

$$\lambda(\varphi) = \lambda(\overline{\varphi}). \tag{1}$$

Then $\overline{\varphi} - \varphi \geq 0$ and $\lambda(\overline{\varphi} - \varphi) = 0$, which yields $\lambda\left(\{x : \varphi(x) \neq \overline{\varphi}(x)\}\right) = 0$. Hence λ is concentrated on $\mathcal{E}$.

The proof of (1) is based on Lemma 7. Consider the sublinear functional $q\colon C(K) \to \mathbb{R}$, given by $q(f) = \lambda(\overline{f})$, and the linear functional L defined on $A(K)+\mathbb{R}\cdot\varphi$ by $L(h+\alpha\varphi) = \lambda(h)+\alpha\lambda(\overline{\varphi})$. By Lemma 7, if $\alpha \geq 0$, then $L(h+\alpha\varphi) = q(h+\alpha\varphi)$, while if $\alpha < 0$, then

$$0 = \overline{\alpha\varphi - \alpha\varphi} \leq \overline{\alpha\varphi} + \overline{(-\alpha\varphi)} = \overline{\alpha\varphi} - \alpha\overline{\varphi},$$

which yields

$$L(h+\alpha\varphi) = \lambda(h+\alpha\overline{\varphi}) \leq \lambda(\overline{h+\alpha\varphi}) = q(h+\alpha\varphi).$$

By the Hahn–Banach extension theorem, there exists a linear extension ω of L to $C(K)$ such that $\omega \leq q$. If $f \leq 0$, then $\overline{f} \leq 0$, so $\omega(f) \leq q(f) = \lambda(\overline{f}) \leq 0$. Therefore $\omega \geq 0$ and the Riesz–Kakutani representation theorem allows us to identify ω with a suitable Borel probability measure on K.

If f is in *Conv* (K), then $-f$ is concave and Lemma 7 yields

$$\omega(-f) \leq q(-f) = \lambda(\overline{-f}) = \lambda(-f)$$

that is, $\lambda \prec \omega$. Or, λ is maximal, which forces $\omega = \lambda$. Consequently,

$$\lambda(\varphi) = \omega(\varphi) = L(\varphi) = \lambda(\overline{\varphi}),$$

which ends the proof. ■

As noticed E. Bishop and K. de Leeuw, if K is non-metrizable, then $\operatorname{Ext} K$ needs not be a Borel set.

Theorem 8 (*The generalization of the Choquet-Bishop-de Leeuw theorem*). Let μ be a Steffensen-Popoviciu measure on a compact convex subset K of a locally convex Hausdorff space E. Then there exists a Borel probability λ on K such that the following two conditions are verified*:*

i) $\lambda \succ \mu$ and λ and μ have the the same barycenter;

ii) λ vanishes on every Baire subset of K which is disjoint from the set of extreme points of K.

9 Choquet's Theory for Signed Measures

Convexity inequalities involving signed linear combinations:

Theorem 9 (*The Jensen-Steffensen Inequality*). *Let* $x_1 \leq x_2 \leq \ldots \leq x_n$ *be points in* $[a,b]$ *and let* $p_1, p_2, \ldots, p_n$ *be real numbers such that the sums* $P_k = \sum_{j=1}^{k} p_j$ *verify* $0 \leq P_k \leq P_n$ *and* $P_n > 0$. *Then for every real valued convex function* f *defined on* $[a,b]$ *we have the inequality*

$$f\left(\frac{1}{P_n}\sum_{k=1}^{n} p_k\, x_k\right) \leq \frac{1}{P_n}\sum_{k=1}^{n} p_k\, f(x_k)$$

A. M. Fink [5] proved an extension of (HH) which also escapes Choquet's theory:

Theorem 10 *If* $f : [a,b] \to \mathbb{R}$ *is a continuous convex function, then*

$$f(x_\mu) \leq \frac{1}{\mu([a,b])}\int_a^b f(x)\, d\mu(x) \tag{FHH}$$

for every real Borel measure μ *on* $[a,b]$ *which is "end positive" in the sense that*

$$\mu([a,b]) > 0, \quad \int_a^t (t-x)\, d\mu(x) \geq 0, \quad \text{and} \quad \int_t^b (x-t)\, d\mu(x) \geq 0, \tag{EP}$$

for every $t \in [a,b]$.

As above, $x_\mu = \int_a^b x\, d\mu(x)/\mu([a,b])$ represents the barycenter of μ.

All Borel probability measures are "end positive". An example in the category of signed measures is provided by the restriction of $(x^2 + a)dx$ to $[-1,1]$, for every $a > -1/3$.

Definition 11 (See [14]).*A Steffensen-Popoviciu measure is any real Borel measure μ on K such that*

$$\mu(K) > 0 \text{ and } \int_K f^+(x)\,d\mu(x) \geq 0 \tag{SPM}$$

for every continuous convex function f on K.

When K is an interval $[a, b]$ and μ is a real Borel measure on $[a, b]$, with $\mu([a, b]) > 0$, the condition (SPM) coincides with the condition of end positivity (EP) mentioned above, a fact which was known to T. Popoviciu. In fact, (SPM) yields $\mu(K) > 0$ and

$$\int_K (x'(x) + t)^+ \; d\mu(x) \geq 0 \quad \text{for every } x' \in E' \text{ and every } t \in \mathbb{R}$$

and the dual of $\mathbb{R}$ consists only of homoteties $x' : x \to sx$. T. Popoviciu's argument for the other implication, (EP) $\Rightarrow$(SPM), was as follows: If $f \geq 0$ is a piecewise linear, continuous and convex function, then f can be represented as a finite combination, with non-negative coefficients, of functions of the form 1, $(x - t)^+$ and $(t - x)^+$, so that

$$\int_K f(x)\,d\mu(x) \geq 0;$$

in the general case, we have to approximate f^+ by piecewise linear continuous and convex functions. It is worth noticing that T. Popoviciu [12] was interested in a slightly different problem, precisely, when a real Radon measure on an interval $[a, b]$ is nonnegative for all n-convex functions on that interval. However, he did not mention any possible connection with the Hermite-Hadamard inequality.

The notion of barycenter can be introduced in the same way and an inspection of the argument of Lemma 5 shows that it remains valid in the general context of Steffensen-Popoviciu measures:

Proposition 12 (*The generalized Jensen-Steffensen inequality* [14]).

Suppose that μ is a Steffensen-Popoviciu measure on a compact convex set K. Then

$$f(x_\mu) \leq \frac{1}{\mu(K)} \int_K f(x)\, d\mu(x) \tag{JS}$$

for every continuous convex function $f : K \to \mathbb{R}$.

Clearly, if μ is a real Borel measure on K, with $\mu(K) > 0$, and (JS) holds for some point $x_\mu \in K$ and every continuous convex function $f : K \to \mathbb{R}$, then μ is a Steffensen-Popoviciu measure.

The next step is to enlarge the concept of majorization.

Definition 13 *Given two Steffensen-Popoviciu measures μ and λ on the compact convex set K, we say that μ is* majorized *by λ (denoted $\mu \prec \lambda$) if*

$$\frac{1}{\mu(K)} \int_K f\, d\mu \leq \frac{1}{\lambda(K)} \int_K f\, d\lambda$$

for every continuous convex function $f : K \to \mathbb{R}$.

In the discrete case, suppose that there are given real points $x_1 \leq ... \leq x_n$ and real weights $p_1, ..., p_n$. According to (EP), the discrete measure $\mu = \sum_{k=1}^n p_k\, \delta_{x_k}$ is a Steffensen-Popoviciu measure if and only if

$$\sum_{k=1}^n p_k > 0, \quad \sum_{k=1}^m p_k(x_m - x_k) \geq 0 \quad \text{and} \quad \sum_{k=m}^n p_k(x_k - x_m) \geq 0 \tag{dEP}$$

for every $m \in \{1, ..., n\}$. A special case when (dEP) holds is the following, used by Steffensen in his extension of Jensen's inequality:

$$\sum_{k=1}^{n} p_k > 0, \quad \text{and} \quad 0 \leq \sum_{k=1}^{m} p_k \leq \sum_{k=1}^{n} p_k, \quad \text{for every } m \in \{1, ..., n\}. \tag{dSt}$$

In fact, (dSt) ⇒ (dEP) by Abel's summation formula (the discrete analogue of integration by parts).

A consequence of (dSt) and Theorem 12 is the following inequality of G. Szegö: *If* $a_1 \geq a_2 \geq ... \geq a_{2m-1} \geq 0$ *and* f *is a convex function in* $[0, a_1]$, *then*

$$\sum_{k=1}^{2m-1} (-1)^{k-1} f(a_k) \geq f\left(\sum_{k=1}^{2m-1} (-1)^{k-1} a_k\right).$$

This corresponds to the measure $\mu = \sum_{k=1}^{2m-1} (-1)^{k-1} \delta_{a_k}$, whose barycenter is $x_\mu = \sum_{k=1}^{2m-1} (-1)^{k-1} a_k$.

In the case of absolutely continuous measures, $d\mu = p(x)\,dx$, the condition (EP) reads as:

$$\int_a^b p(x)\,dx > 0, \quad \int_a^t (t-x)p(x)\,dx \geq 0 \quad \text{and} \quad \int_t^b (x-t)p(x)\,dx \geq 0 \tag{cEP}$$

for every $t \in [a, b]$. In practice, the following continuous analogue of (dSt) suffices:

$$\int_a^b p(x)\,dx > 0 \quad \text{and} \quad 0 \leq \int_a^t p(x)\,dx \leq \int_a^b p(x)\,dx \text{ for every } t \in [a, b]. \tag{cSt}$$

Problem 14 *What's the physical signification of Steffensen-Popoviciu measures?*

Problem 15 *Is it possible to extend Choquet's theory to the case of* n*-convex functions?*

We shall recall here an inequality of J. F. Steffensen [25], [19], that shed some light on the case of 0-convex functions:

Theorem 16 *Let* $g : [a,b] \to \mathbb{R}$ *be an integrable function such that* $\lambda = \int_a^b g(t)\,dt \in [0, b-a]$. *Then the following two conditions are equivalent:*

i) $0 \leq \int_a^x g(t)\,dt \leq x-a$ *and* $0 \leq \int_x^b g(t)\,dt \leq b-x$, *for every* $x \in [a,b]$;

ii) $\int_a^{a+\lambda} h(t)\,dt \leq \int_a^b h(t)g(t)\,dt \leq \int_{b-\lambda}^b h(t)\,dt$, *for every nondecreasing function* $h : [a,b] \to \mathbb{R}$.

Proof. *i)* $\Rightarrow$ *ii)* In fact,

$$\int_a^b h(t)g(t)\,dt - \int_a^{a+\lambda} h(t)\,dt = \int_a^{a+\lambda} h(t)\left(g(t)-1\right)\,dt + \int_{a+\lambda}^b h(t)g(t)\,dt$$

$$= \int_a^{a+\lambda} h(t)\,d\left(\int_a^t g(s)\,ds - t + a\right) - \int_{a+\lambda}^b h(t)\,d\left(\int_t^b g(s)\,ds\right)$$

$$= \int_a^{a+\lambda} \left(\int_a^t g(s)\,ds - t + a\right)\,dh(t) + \int_{a+\lambda}^b \left(\int_t^b g(s)\,ds\right)\,dh(t)$$

which gives us the left hand inequality of ii). The other one can be obtained in a similar manner.

ii) $\Rightarrow$ *i)* Consider the functions $h = -\chi_{[a,x]}$ and $h = \chi_{[x,b]}$. ■

If $f : [a,b] \to \mathbb{R}$ is a Lipschitz function with $||f||_{Lip} = M$, then f is differentiable a.e. and $0 \leq g = \left(f' + M\right)/(2M) \leq 1$ verifies condition *i)* of Theorem 16. For this choice of g and $h = x - a$, the condition *ii)* yields Iyengar's inequality (I).

10 The case of quasi Steffensen-Popoviciu measures

we can handle the case total mass is 0 by considering small perturbations $\mu_\varepsilon = \mu + \varepsilon\delta_z$. Then

$$f(x_\mu) \cdot (\mu(K) + \varepsilon) \leq \int_K f(x)\, d\mu(x) + \varepsilon f(z)$$

for every continuous convex function f on K. Letting $\varepsilon \to 0$, we arrive at the following result:

Proposition 17 (See [14]). *Let μ be a real Borel measure on a compact convex K such that $\mu(K) = 0$ and*

$$\int_K f^+ \, d\mu(x) \geq 0$$

for every continuous convex function f on K. Then

$$\int_K f(x)\, d\mu(x) \geq 0$$

for every continuous convex function f on K.

Due to Popoviciu's characterization of convex functions (mentioned in Section 6), better results can be proved on intervals:

Proposition 18 *Let μ be a real Borel measure on $[a,b]$ such that*

$$\mu([a,b]) = 0, \ \int_a^t (t-x)\, d\mu(x) \geq 0 \quad \textit{and} \quad \int_t^b (x-t)\, d\mu(x) \geq 0$$

for every $t \in R$. Then

$$\int_a^b f(x)\, d\mu(x) \geq 0$$

for every convex function f on $[a,b]$.

As an immediate consequence we obtain L. Fuchs' extension of the majorization principle:

Theorem 19 *Let* $f : [a,b] \to \mathbb{R}$ *be a convex function. Then for every* $x_1, ..., x_n,\ y_1, ..., y_n \in [a,b]$ *and every* $p_1, ..., p_n \in \mathbb{R}$ *such that*

i) $x_1 > ... > x_n, \quad y_1 > ... > y_n$

ii) $\sum\limits_{k=1}^{r} p_k\, x_k \leq \sum\limits_{k=1}^{r} p_k\, y_k$ *for every* $r = 1, ..., n-1$

iii) $\sum\limits_{k=1}^{n} p_k\, x_k = \sum\limits_{k=1}^{n} p_k y_k$

we have the inequality

$$\sum_{k=1}^{n} p_k\, f(x_k) \leq \sum_{k=1}^{n} p_k\, f(y_k).$$

We pass now to the case of absolutely continuous measures:

Proposition 20 *Let* $p(x)$ *be a continuous or a monotonic density on an interval* $[a,b]$, *such that*

$$\int_a^b p(x)\, dx = 0 \quad \text{and} \quad \int_a^t p(x)\, dx \geq 0, \quad \int_t^b p(x)\, dx \geq 0 \tag{qcSt}$$

for every $t \in [a,b]$. *Then*

$$\int_a^b f(x)\, p(x)\, dx \geq 0$$

for every convex function f *on* $[a,b]$.

When the graph of $p(x)$ is symmetric with respect to the line $x = (a+b)/2$, it suffices to ask (qcSt) only for t in the appropriate half interval. This remark allows us to retrieve the following result due to Levin and Stečkin: *Suppose that* $g : [-a, a] \to \mathbb{R}$ *is an even function, nondecreasing on* $[0, a]$, *and* $f : [-a, a] \to \mathbb{R}$ *is a convex function. Then*

$$\frac{1}{2a}\int_{-a}^{a} f(x)g(x)\,dx \geq \left(\frac{1}{2a}\int_{-a}^{a} f(x)dx\right)\left(\frac{1}{2a}\int_{-a}^{a} g(x)\,dx\right).$$

In fact, $p(x) = g(x) - \frac{1}{2a}\int_{-a}^{a} g(x)\,dx$ is an even weight such that

$$\int_{-a}^{-t} p(x)\,dx = \int_{t}^{a} p(x)\,dx \geq 0$$

for every $t \in [0, a]$.

11 The story continues

While the left part of the Hermite-Hadamard inequality imposes the restriction to Steffensen-Popoviciu measures as a general framework for the whole inequality, the right part works outside this restriction. More precisely, on each interval $[a, b]$ there are Borel measures μ which are not Steffensen-Popoviciu yet

$$\frac{1}{\mu([a,b])}\int_a^b f(x)\,d\mu(x) \leq \frac{b-x_\mu}{b-a}\cdot f(a)+\frac{x_\mu-a}{b-a}\cdot f(b)$$

works for every convex function on $[a, b]$. In fact, as noticed A. M. Fink [5],

$$\frac{3}{2}\int_{-1}^{1} f(x)(x^2-x)dx \leq 1\cdot f(-1)+0\cdot f(1)=f(-1)$$

for every convex function $f : [-1, 1] \to \mathbb{R}$.

See Florea and Niculescu [6] for a complete solution.

12 Open Problem

Unify the results in Sections 6 and 7.

As well known, not every subharmonic function is convex.

References

[1] H. Alzer, *Ungleichungen für* $(e/a)^a/(b/e)^b$, Elemente Math., **40** (1985), 120-123.

[2] E. Artin, *The Gamma Function*, Holt, Rinehart and Winston, New York, 1964. English translation of German original, *Einführung in die Theorie der Gammafunktion*, Teubner, 1931.

[3] F. Burk, *The Geometric, Logarithmic and Arithmetic Mean Inequality*, Amer. Math. Month., **94** (1987), 527-528.

[4] Lj. Dedić, C. E. M. Pearce and J. Pečarić, *Hadamard and Dragomir-Agarwal Inequalities, Higher-Order Convexity and the Euler Formula*, J. Korean Math. Soc., **38** (2001), 1235-1243.

[5] A. M. Fink, *A best possible Hadamard inequality*, Math. Inequal. Appl., **1** (1998), 223-230.

[6] A. Florea and C. P. Niculescu, *A Hermite-Hadamard inequality for convex-concave functions*, Bull. Soc. Sci. Math. Roum., to appear.

[7] L. Galvani, *Sulle funzioni converse di una o due variabili definite in aggregate qualunque,* Rend. Circ. Mat. Palermo, **41** (1916), 103-134.

[8] J. Hadamard, *Étude sur les propriétés des fonctions entières et en particulier d'une fonction considerée par Riemann,* J. Math. Pures et Appl., **58** (1893), 171-215.

[9] G. H. Hardy, J. E. Littlewood and G. Pólya, *Some simple inequalities satisfied by convex functions*, Messenger Math., **58** (1929), 145-152.

[10] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities,* Cambridge Mathematical Library, 2nd ed., 1952, Reprinted 1988.

[11] Ch. Hermite, *Sur deux limites d'une intégrale définie*, Mathesis, **3** (1883), p. 82.

[12] K. S. K. Iyengar, *Note on an inequality*, Math. Student, **6** (1938), 75-76.

[13] D. S. Mitrinović and I. B. Lacković, *Hermite and convexity*, Aequationes Mathematicae, **28** (1985), 229-232.

[14] C. P. Niculescu, *Choquet theory for signed measures*, Math. Inequal. Appl., **5** (2002), 479-489.

[15] C. P. Niculescu, *The Hermite-Hadamard inequality for functions of a vector variable*, Math. Inequal. Appl., **5** (2002), 619-623.

[16] C. P. Niculescu, *Convexity according to means*, Math. Inequal. Appl., **6** (2003), 571-579.

[17] C. P. Niculescu and L.-E. Persson, *Old and new on the Hermite-Hadamard inequality,* Real Analysis Exchange, **29** (2003/04), no. 2, 663-686.

[18] C. P. Niculescu and L.-E. Persson, *Convex Functions and their Applications. A Contemporary Approach*, CMS Books in Mathematics vol. **23**, Springer-Verlag, New York, 2006.

[19] C. P. Niculescu and F. Popovici, *A Note on the Denjoy-Bourbaki Theorem,* Real Analysis Exchange, **29** (2003/04), no. 2, 639-646.

[20] J. Pečarić, F. Proschan and Y. L. Tong, *Convex Functions, Partial Orderings and Statistical Applications*, Academic Press Inc., 1992.

[21] R. R. Phelps, *Lectures on Choquet's Theorem*, 2nd Ed., Lecture Notes in Math., Nr. **1757** (2001).

[22] T. Popoviciu, *Sur quelques propriétés des fonctions d'une variable réelle convexes d'ordre superieur*, Mathematica (Cluj), **8** (1934), 1-85.

[23] T. Popoviciu, *Notes sur les fonctions convexes d'ordre superieur* (IX), Bull. Math. Soc. Roum. Sci., **43** (1941), 85-141.

[24] T. Popoviciu, *Les Fonctions Convexes*, Hermann, Paris, 1944.

[25] J. F. Steffensen, *On certain inequalities between mean values, and their application to actuarial problems*, Skandinavisk Aktuarietidskrift, 1918, 82-97.

[26] J. F. Steffensen, *On certain inequalities and methods of approximation*, J. Inst. Actuaries, **51** (1919), 274-297.

[27] Tung Po-Lin, *The power mean and the logarithmic mean*, Amer. Math. Monthly, **81** (1974), 879-883.

[28] R. Webster, *Convexity*, Oxford Univ. Press, Oxford·New York·Tokyo, 1994.