Preface

It seems to me that the notion of convex function is just as fundamental as positive function or increasing function. If I am not mistaken in this, the notion ought to find its place in elementary expositions of the theory of real functions.

J. L. W. V. Jensen

Convexity is a simple and natural notion which can be traced back to Archimedes (circa 250 B.C.), in connection with his famous estimate of the value of π (by using inscribed and circumscribed regular polygons). He noticed the important fact that the perimeter of a convex figure is smaller than the perimeter of any other convex figure surrounding it.

As a matter of fact, we experience convexity all the time and in many ways. The most prosaic example is our upright position, which is secured as long as the vertical projection of our center of gravity lies inside the convex envelope of our feet. Also, convexity has a great impact on our everyday life through numerous applications in industry, business, medicine, and art. So do the problems of optimum allocation of resources and equilibrium of noncooperative games.

The theory of convex functions is part of the general subject of convexity, since a convex function is one whose epigraph is a convex set. Nonetheless it is an important theory per se, which touches almost all branches of mathematics. Graphical analysis is one of the first topics in mathematics which requires the concept of convexity. Calculus gives us a powerful tool in recognizing convexity, the second-derivative test. Miraculously, this has a natural generalization for the several variables case, the Hessian test. Motivated by some deep problems in optimization and control theory, convex function theory has been extended to the framework of infinite dimensional Banach spaces (and even further).

The recognition of the subject of convex functions as one that deserves to be studied in its own right is generally ascribed to J. L. W. V. Jensen [114], [115]. However he was not the first to deal with such functions. Among his predecessors we should recall here Ch. Hermite [102], O. Hölder [106] and O. Stolz [233]. During the twentieth century, there was intense research activity and significant results were obtained in geometric functional analysis, mathematical economics, convex analysis, and nonlinear optimization. A classic book by G. H. Hardy, J. E. Littlewood and G. Pólya [99] played a large role in the popularization of the subject of convex functions.

Roughly speaking, there are two basic properties of convex functions that make them so widely used in theoretical and applied mathematics:

- *The maximum is attained at a boundary point.*
- *Any local minimum is a global one. Moreover, a strictly convex function admits at most one minimum.*

The modern viewpoint on convex functions entails a powerful and elegant interaction between analysis and geometry. In a memorable paper dedicated to the Brunn–Minkowski inequality, R. J. Gardner [88, p. 358], described this reality in beautiful phrases: [convexity] "appears like an octopus, tentacles reaching far and wide, its shape and color changing as it roams from one area to the next. It is quite clear that research opportunities abound."

Over the years a number of notable books dedicated to the theory and applications of convex functions appeared. We mention here: L. Hörmander [108], M. A. Krasnosel'skii and Ya. B. Rutickii [132], J. E. Pečarić, F. Proschan and Y. C. Tong [196], R. R. Phelps [199], [200] and A. W. Roberts and D. E. Varberg [212]. The references at the end of this book include many other fine books dedicated to one aspect or another of the theory.

The title of the book by L. Hörmander, *Notions of Convexity,* is very suggestive for the present state of art. In fact, nowadays the study of convex functions has evolved into a larger theory about functions which are adapted to other geometries of the domain and/or obey other laws of comparison of means. Examples are log-convex functions, multiplicatively convex functions, subharmonic functions, and functions which are convex with respect to a subgroup of the linear group.

Our book aims to be a thorough introduction to contemporary convex function theory. It covers a large variety of subjects, from the one real variable case to the infinite dimensional case, including Jensen's inequality and its ramifications, the Hardy–Littlewood–Pólya theory of majorization, the theory of gamma and beta functions, the Borell–Brascamp–Lieb form of the Prékopa–Leindler inequality (as well as the connection with isoperimetric inequalities), Alexandrov's well-known result on the second differentiability of convex functions, the highlights of Choquet's theory, a brief account on the recent solution to Horn's conjecture, and many more. It is certainly a book where inequalities play a central role but in no case a book on inequalities. Many results are new, and the whole book reflects our own experiences, both in teaching and research.

This book may serve many purposes, ranging from a one-semester graduate course on Convex Functions and Applications to additional bibliographic material. In a course for first year graduate students, we used the following route:

- *Background*: Sections 1.1–1.3, 1.5, 1.7, 1.8, 1.10.
- *The beta and gamma functions*: Section 2.2.

- *Convex functions of several variables*: Sections 3.1–3.12.
- *The variational approach of partial differential equations*: Appendix C.

The necessary background is advanced calculus and linear algebra. This can be covered from many sources, for example, from *Analysis I* and *II* by S. Lang [137], [138]. A thorough presentation of the fundamentals of measure theory is also available in L. C. Evans and R. F. Gariepy [74]. For further reading we recommend the classical texts by F. H. Clarke [56] and I. Ekeland and R. Temam [70].

Our book is not meant to be read from cover to cover. For example, Section 1.9, which deals with the Hermite–Hadamard inequality, offers a good starting point for Choquet's theory. Then the reader may continue with Chapter 4, where this theory is presented in a slightly more general form, to allow the presence of certain signed measures. We recommend this chapter to be studied in parallel with the *Lectures on Choquet's theory* by R. R. Phelps [200]. For the reader's convenience, we collected in Appendix A all the necessary material on the separation of convex sets in locally convex Hausdorff spaces (as well as a proof of the Krein–Milman theorem).

Appendix B may be seen both as an illustration of convex function theory and an introduction to an important topic in real algebraic geometry: the theory of semi-algebraic sets.

Sections 3.11 and 3.12 offer all necessary background on a further study of convex geometric analysis, a fast-growing topic which relates many important branches of mathematics.

To help the reader in understanding the theory presented, each section ends with exercises (accompanied by hints). Also, each chapter ends with comments covering supplementary material and historical information. The primary sources we have relied upon for this book are listed in the references.

In order to avoid any confusion relative to our notation, a symbol index was added for the convenience of the reader. Notice that our book deals only with *real* linear spaces and all Borel measures under attention are assumed to be *regular*.

We wish to thank all our colleagues and friends who read and commented on various versions and parts of the manuscript: Madalina Deaconu, Andaluzia Matei, Sorin Micu, Florin Popovici, Mircea Preda, Thomas Strömberg, Andrei Vernescu, Peter Wall, Anna Wedestig and Tudor Zamfirescu.

We also acknowledge the financial support of Wenner–Gren Foundations (Grant 25 12 2002), which made possible the cooperation of the two authors.

In order to keep in touch with our readers, a web page for this book will be made available at http://www.inf.ucv.ro/∼niculescu/Convex_Functions.html

Craiova and Luleå
September 2004

Constantin P. Niculescu
Lars-Erik Persson

2

Comparative Convexity on Intervals

This chapter is devoted to a succinct presentation of several classes of functions acting on intervals, which satisfy inequalities of the form

$$f(M(x,y)) \leq N(f(x), f(y)),$$

for a suitable pair of means M and N. Leaving out the case of usual convex functions (when M and N coincide with the arithmetic mean), the most important classes that arise in applications are:

- the class of log-convex functions (M is the arithmetic mean and N is the geometric mean)
- the class of multiplicatively convex functions (M and N are both geometric means)
- the class of M_p-convex functions (M is the arithmetic mean and N is the power mean of order p).

They all provide important applications to many areas of mathematics.

2.1 Algebraic Versions of Convexity

The usual definition of a convex function (of one real variable) depends on the structure of $\mathbb{R}$ as an ordered vector space. As $\mathbb{R}$ is actually an ordered field, it is natural to investigate what happens when addition is replaced by multiplication and the arithmetic mean is replaced by the geometric mean.

The characteristic property of the subintervals I of $\mathbb{R}$ is

$$x, y \in I \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in I$$

so, in order to draw a parallel in the multiplicative case, we must restrict to the subintervals J of $(0,\infty)$ and use instead the following fact:

$$x, y \in J \text{ and } \lambda \in [0,1] \implies x^{1-\lambda}y^{\lambda} \in J.$$

Depending on which type of mean, arithmetic (A) or geometric (G), we consider on the domain and on the range, we shall encounter one of the following four classes of functions:

- (A, A)-*convex functions*, the usual convex functions;
- (A, G)-*convex functions*;
- (G, A)-*convex functions*;
- (G, G)-*convex functions*.

More precisely, the (A, G)-*convex functions* (usually known as log-*convex functions*) are those functions $f\colon I \to (0, \infty)$ for which

$$x, y \in I \text{ and } \lambda \in [0, 1] \implies f((1-\lambda)x + \lambda y) \leq f(x)^{1-\lambda} f(y)^{\lambda}, \tag{AG}$$

that is, for which $\log f$ is convex. If a function $f\colon I \to \mathbb{R}$ is log-convex, then it is also convex. In fact, according to the AM–GM inequality,

$$f((1-\lambda)x + \lambda y) \leq f(x)^{1-\lambda} f(y)^{\lambda} \leq (1-\lambda)f(x) + \lambda f(y).$$

The converse does not work. For example, the function $\mathrm{e}^x - 1$ is convex and log-concave.

One of the most notable examples of a log-convex function is Euler's *gamma function*,

$$\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}\,dt, \quad x > 0.$$

The place of Γ in the landscape of log-convex functions is the subject of the next section.

The class of all (G, A)-*convex functions* consists of all real-valued functions f (defined on subintervals I of $(0, \infty)$) for which

$$x, y \in I \text{ and } \lambda \in [0, 1] \implies f(x^{1-\lambda} y^{\lambda}) \leq (1-\lambda)f(x) + \lambda f(y). \tag{GA}$$

In the context of twice-differentiable functions $f\colon I \to \mathbb{R}$, (G, A)-convexity means $x^2 f'' + x f' \geq 0$.

The (G, G)-*convex functions* (called *multiplicatively convex functions* in what follows) are those functions $f\colon I \to J$ (acting on subintervals of $(0, \infty)$) such that

$$x, y \in I \text{ and } \lambda \in [0, 1] \implies f(x^{1-\lambda} y^{\lambda}) \leq f(x)^{1-\lambda} f(y)^{\lambda}. \tag{GG}$$

Equivalently, f is multiplicatively convex if and only if $\log f(x)$ is a convex function of $\log x$. This fact will be shown in Lemma 2.3.1 below. Due to the arithmetic-geometric mean inequality, all multiplicatively convex functions (and also all nondecreasing convex functions) are (G, A)-convex functions.

The theory of multiplicatively convex functions is similar to that of classical convex functions. In fact, they differ from each other only by a change of variable and a change of function:

Lemma 2.1.1 *Suppose that I is a subinterval of $(0,\infty)$ and $f\colon I\to(0,\infty)$ is a multiplicatively convex function on I. Then*

$$F=\log\circ f\circ\exp\colon\ \log(I)\to\mathbb{R}$$

is a convex function. Conversely, if J is an interval and $F\colon J\to\mathbb{R}$ is a convex function, then

$$f=\exp\circ F\circ\log\colon\ \exp(J)\to(0,\infty)$$

is a multiplicatively convex function.

The proof is straightforward. Lemma 2.1.1 can be adapted easily to other situations and allows us to deduce new inequalities from old ones. This idea is central to Section 2.3 below.

Exercises

1. (Some geometrical consequences of log-convexity)

(i) A convex quadrilateral $ABCD$ is inscribed in the unit circle. Its sides satisfy the inequality $AB\cdot BC\cdot CD\cdot DA\geq 4$. Prove that $ABCD$ is a square.

(ii) Suppose that A, B, C are the angles of a triangle, expressed in radians. Prove that

$$\sin A\sin B\sin C<\Big(\frac{3\sqrt{3}}{2\pi}\Big)^3 ABC<\Big(\frac{\sqrt{3}}{2}\Big)^3,$$

unless $A=B=C$.

[*Hint*: Note that the sine function is log-concave, while $x/\sin x$ is log-convex on $(0,\pi)$.]

2. Let (X,Σ,μ) be a measure space and let $f\colon X\to\mathbb{C}$ be a measurable function, which is in $L^t(\mu)$ for t in a subinterval I of $(0,\infty)$. Infer from the Cauchy–Buniakovski–Schwarz inequality that the function $t\to\log\int_X|f|^t\,d\mu$ is convex on I.
Remark. The result of this exercise is equivalent to *Lyapunov's inequality* [148]: *If $a\geq b\geq c$, then*

$$\Big(\int_X|f|^b\,d\mu\Big)^{a-c}\leq\Big(\int_X|f|^c\,d\mu\Big)^{a-b}\Big(\int_X|f|^a\,d\mu\Big)^{b-c}$$

(provided the integrability aspects are fixed). Equality holds if and only if one of the following conditions hold:

(i) f is constant on some subset of Ω and 0 elsewhere;
(ii) $a=b$;
(iii) $b=c$;

(iv) $c(2a-b)=ab$.

3. (P. Montel [171]) Let I be an interval. Prove that the following assertions are equivalent for every function $f\colon I\to(0,\infty)$:

(i) f is log-convex;
(ii) the function $x\to \mathrm{e}^{\alpha x}f(x)$ is convex on I for all $\alpha\in\mathbb{R}$;
(iii) the function $x\to[f(x)]^{\alpha}$ is convex on I for all $\alpha>0$.

[*Hint*: For (iii) $\Rightarrow$ (i), note that $([f(x)]^{\alpha}-1)/\alpha$ is convex for all $\alpha>0$ and $\log f(x)=\lim_{\alpha\to 0+}([f(x)]^{\alpha}-1)/\alpha$. Then apply Corollary 1.3.8.]

4. Prove that the sum of two log-convex functions is also log-convex.
[*Hint*: Note that this assertion is equivalent to the following inequality for positive numbers: $a^{\alpha}b^{\beta}+c^{\alpha}d^{\beta}\leq(a+c)^{\alpha}(b+d)^{\beta}$.]

5. (S. Simic [227]) Let $(a_n)_n$ be a sequence of positive numbers. Prove that the following assertions are equivalent:

(i) $(a_n)_n$ is log-convex (that is, $a_{n-1}a_{n+1}\geq a_n^2$ for all $n\geq 1$);
(ii) for each $x\geq 0$, the sequence $P_n(x)=\sum_{k=0}^{n}a_k\binom{n}{k}x^{n-k}$ ($n\in\mathbb{N}$) is log-convex.

6. A function $f\colon(0,\infty)\to\mathbb{R}$ is called *completely monotonic* if f has derivatives of all orders and satisfies $(-1)^n f^{(n)}(x)\geq 0$ for all $x>0$ and $n\in\mathbb{N}$. In particular, completely monotonic functions are decreasing and convex.

(i) Prove that
$$(-1)^{nk}(f^{(k)}(x))^n\leq(-1)^{nk}(f^{(n)}(x))^k(f(x))^{n-k}$$
for all $x>0$ and all integers n, k with $n\geq k\geq 0$. Infer that any completely monotonic function is actually log-convex.

(ii) Prove that the function
$$V_q(x)=\frac{\exp(x^2)}{\Gamma(q+1)}\int_x^{\infty}\mathrm{e}^{-t^2}(t^2-x^2)^q\,dt$$
is completely monotonic on $(0,\infty)$ if $q\in(-1,0]$.

2.2 The Gamma and Beta Functions

The gamma function $\Gamma\colon(0,\infty)\to\mathbb{R}$ is defined by the relation
$$\Gamma(x)=\int_0^{\infty}t^{x-1}\mathrm{e}^{-t}\,dt\quad\text{for } x>0.$$

Theorem 2.2.1 *The gamma function has the following properties:*

(i) $\Gamma(x+1)=x\Gamma(x)$ *for all* $x>0$*;*
(ii) $\Gamma(1)=1$*;*

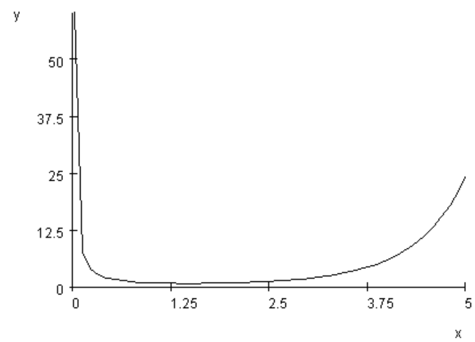

Fig. 2.1. The graph of Γ.

(iii) Γ *is* log-*convex.*

Proof. (i) Using integration by parts we get

$$\Gamma(x+1) = \int_0^\infty t^x \mathrm{e}^{-t}\,dt = \left[-t^x\mathrm{e}^{-t}\right]\Big|_{t=0}^{\infty} + x\int_0^\infty t^{x-1}\mathrm{e}^{-t}\,dt = x\Gamma(x)$$

for all $x > 0$.

The property (ii) is obvious.

(iii) Let $x, y > 0$ and let $\lambda, \mu \geq 0$ with $\lambda + \mu = 1$. Then, by the Rogers–Hölder inequality, we have

$$\begin{aligned}\Gamma(\lambda x + \mu y) &= \int_0^\infty t^{\lambda x+\mu y-1}\mathrm{e}^{-t}\,dt = \int_0^\infty (t^{x-1}\mathrm{e}^{-t})^\lambda (t^{y-1}\mathrm{e}^{-t})^\mu\,dt \\ &\leq \left(\int_0^\infty t^{x-1}\mathrm{e}^{-t}\,dt\right)^\lambda \left(\int_0^\infty t^{y-1}\mathrm{e}^{-t}\,dt\right)^\mu = \Gamma^\lambda(x)\Gamma^\mu(y)\end{aligned}$$

which proves that Γ is log-convex. □

Corollary 2.2.2 $\Gamma(n+1) = n!$ *for all* $n \in \mathbb{N}$.

Corollary 2.2.3 *The gamma function is convex and* $x\Gamma(x)$ *approaches* 1 *as* $x \to 0+$.

C. F. Gauss first noted that Γ attains its minimum at $x = 1.461632145\ldots$.

The gamma function is the unique log-convex extension of the factorial function:

Theorem 2.2.4 (H. Bohr and J. Mollerup [32], [10]) *Suppose the function* $f\colon (0,\infty) \to \mathbb{R}$ *satisfies the following three conditions:*

(i) $f(x+1) = xf(x)$ *for all* $x > 0$*;*
(ii) $f(1) = 1$*;*
(iii) f *is* log*-convex.*

Then $f = \Gamma$.

Proof. By induction, from (i) and (ii) we infer that $f(n+1) = n!$ for all $n \in \mathbb{N}$.

Now, let $x \in (0,1]$ and $n \in \mathbb{N}^\star$. Then by (iii) and (i),

$$\begin{aligned} f(n+1+x) &= f((1-x)(n+1) + x(n+2)) \\ &\leq [f(n+1)]^{1-x} \cdot [f(n+2)]^x \\ &= [f(n+1)]^{1-x} \cdot (n+1)^x \cdot [f(n+1)]^x \\ &= (n+1)^x \cdot f(n+1) \\ &= (n+1)^x \cdot n! \end{aligned}$$

and

$$\begin{aligned} n! = f(n+1) &= f(x(n+x) + (1-x)(n+1+x)) \\ &\leq [f(n+x)]^x \cdot [f(n+1+x)]^{1-x} \\ &= (n+x)^{-x} \cdot [f(n+1+x)]^x \cdot [f(n+1+x)]^{1-x} \\ &= (n+x)^{-x} \cdot f(n+1+x). \end{aligned}$$

Thus, since $f(n+1+x) = (n+x)(n-1+x)\cdots xf(x)$, we obtain

$$\left(1+\frac{x}{n}\right)^x \leq \frac{(n+x)(n-1+x)\cdots xf(x)}{n!\,n^x} \leq \left(1+\frac{1}{n}\right)^x,$$

which yields

$$f(x) = \lim_{n\to\infty} \frac{n!\,n^x}{(n+x)(n-1+x)\cdots x} \quad \text{for } x \in (0,1].$$

We shall show that the above formula is valid for all $x > 0$ so that f is uniquely determined by the conditions (i), (ii) and (iii). Since Γ satisfies all these three conditions, we must have $f = \Gamma$.

To end the proof, suppose that $x > 0$ and choose an integer number m such that $0 < x - m \leq 1$. According to (i) and what we have just proved, we get

$$\begin{aligned} f(x) &= (x-1)\cdots(x-m)f(x-m) \\ &= (x-1)\cdots(x-m) \cdot \lim_{n\to\infty} \frac{n!\,n^{x-m}}{(n+x-m)(n-1+x-m)\cdots(x-m)} \\ &= \lim_{n\to\infty}\Big(\frac{n!\,n^x}{(n+x)(n-1+x)\cdots x} \\ &\qquad\qquad \cdot \frac{(n+x)(n+x-1)\cdots(n+x-(m-1))}{n^m}\Big) \end{aligned}$$

$$= \lim_{n\to\infty} \frac{n!\, n^x}{(n+x)(n-1+x)\cdots x}$$
$$\cdot \lim_{n\to\infty} \left(\left(1+\frac{x}{n}\right)\left(1+\frac{x-1}{n}\right)\cdots\left(1+\frac{x-m+1}{n}\right)\right)$$
$$= \lim_{n\to\infty} \frac{n!\, n^x}{(n+x)(n-1+x)\cdots x}.$$

□

Corollary 2.2.5 $\Gamma(x) = \lim_{n\to\infty} \frac{n!\, n^x}{(n+x)(n-1+x)\cdots x}$ *for all* $x > 0$.

Before establishing a fundamental identity linking the gamma and sine functions, we need to express $\sin x$ as an infinite product:

Theorem 2.2.6 (L. Euler) *For all real numbers* x*,*

$$\sin x = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2\pi^2}\right).$$

Proof. De Moivre's formula shows that $\sin(2n+1)\theta$ is a polynomial of degree $2n+1$ in $\sin\theta$ (for each $n \in \mathbb{N}$, arbitrarily fixed). This polynomial has roots $\pm\sin(k\pi/(2n+1))$ for $k = 0, \ldots, n$. It follows that

$$\sin(2n+1)\theta = (2n+1)\sin\theta \prod_{k=1}^{n} \left(1 - \frac{\sin^2\theta}{\sin^2\frac{k\pi}{2n+1}}\right).$$

Suppose that $x > 0$ and fix arbitrarily two integers m and n such that $x < m < n$. The last identity shows that

$$\frac{\sin x}{(2n+1)\sin\frac{x}{2n+1}} = \prod_{k=1}^{n} \left(1 - \frac{\sin^2\frac{x}{2n+1}}{\sin^2\frac{k\pi}{2n+1}}\right).$$

Denote by a_k the k-th factor in this last product. Since $2\theta/\pi < \sin\theta < \theta$ when $0 < \theta < \pi/2$, we find that

$$0 < 1 - \frac{x^2}{4k^2} < a_k < 1 \quad \text{for } m < k \le n,$$

which yields

$$1 > a_{m+1}\cdots a_n > \prod_{k=1}^{n}\left(1 - \frac{x^2}{4k^2}\right) > 1 - \frac{x^2}{4}\sum_{k=m+1}^{n}\frac{1}{k^2} > 1 - \frac{x^2}{4m}.$$

Hence

$$\frac{\sin x}{(2n+1)\sin\frac{x}{2n+1}}$$

lies between

$$\Big(1-\frac{x^2}{4m}\Big)\prod_{k=1}^{n}\Big(1-\frac{\sin^2\frac{x}{2n+1}}{\sin^2\frac{k\pi}{2n+1}}\Big)\quad\text{and}\quad\prod_{k=1}^{n}\Big(1-\frac{\sin^2\frac{x}{2n+1}}{\sin^2\frac{k\pi}{2n+1}}\Big)$$

and so, letting $n\to\infty$, we deduce that $\sin x/x$ lies between

$$\Big(1-\frac{x^2}{4m}\Big)\prod_{k=1}^{\infty}\Big(1-\frac{x^2}{k^2\pi^2}\Big)\quad\text{and}\quad\prod_{k=1}^{\infty}\Big(1-\frac{x^2}{k^2\pi^2}\Big).$$

The proof ends by letting $m\to\infty$. □

Theorem 2.2.7 *For all real x with $0<x<1$,*

$$\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin\pi x}.$$

Proof. In fact, by Corollary 2.2.5 and Theorem 2.2.6 above we infer that

$$\begin{aligned}\Gamma(x)\Gamma(1-x)&=\lim_{n\to\infty}\frac{n!\,n^x n!\,n^{1-x}}{(n+x)\cdots x\,(n+1-x)\cdots(1-x)}\\&=\frac{1}{x\prod_{k=1}^{\infty}(1-x^2/k^2)}=\frac{\pi}{\sin\pi x}.\end{aligned}$$

□

Corollary 2.2.8 $\Gamma(1/2)=\sqrt{\pi}$.

A variant of the last corollary is the formula

$$\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathrm{e}^{-t^2/2}\,dt=1$$

which appears in many places in mathematics, statistics and natural sciences. Another beautiful consequence of Theorem 2.2.4 is the following:

Theorem 2.2.9 (The Gauss–Legendre duplication formula)

$$\Gamma\Big(\frac{x}{2}\Big)\Gamma\Big(\frac{x+1}{2}\Big)=\frac{\sqrt{\pi}}{2^{x-1}}\Gamma(x)\quad \textit{for all } x>0.$$

Proof. Notice that the function

$$f(x)=\frac{2^{x-1}}{\sqrt{\pi}}\Gamma\Big(\frac{x}{2}\Big)\Gamma\Big(\frac{x+1}{2}\Big)\quad x>0,$$

verifies the conditions (i)–(iii) in Theorem 2.2.4 and thus equals Γ. □

We will prove Stirling's formula, which is an important tool in analytic number theory. We shall need the following lemma:

Lemma 2.2.10 *The sequence* $(a_n)_n$*, whose* n*-th term is*

$$a_n = \log n! - \Big(n + \frac{1}{2}\Big) \log n + n,$$

converges.

Proof. We shall show that the sequence is decreasing and bounded below. In fact,

$$a_n - a_{n+1} = \Big(n + \frac{1}{2}\Big) \log\Big(1 + \frac{1}{n}\Big) - 1 \geq 0$$

since by the Hermite–Hadamard inequality applied to the convex function $1/x$ on $[n, n+1]$ we have

$$\log\Big(1 + \frac{1}{n}\Big) = \int_n^{n+1} \frac{dx}{x} \geq \frac{1}{n + 1/2}.$$

A similar argument (applied to the concave function $\log x$ on $[u, v]$) yields

$$\int_u^v \log x \, dx \leq (v - u) \log \frac{u + v}{2},$$

so that (taking into account the monotonicity of the log function) we get

$$\begin{aligned}\int_1^n \log x \, dx &= \int_1^{1+1/2} \log x \, dx + \int_{1+1/2}^{2+1/2} \log x \, dx + \cdots + \int_{n-1/2}^{n} \log x \, dx \\ &\leq \frac{1}{2} \log \frac{3}{2} + \log 2 + \cdots + \log(n-1) + \frac{1}{2} \log n \\ &< \frac{1}{2} + \log n! - \frac{1}{2} \log n.\end{aligned}$$

Since

$$\int_1^n \log x \, dx = n \log n - n + 1,$$

we conclude that

$$a_n = \log n! - \Big(n + \frac{1}{2}\Big) \log n + n > \frac{1}{2}.$$

The result now follows. ◻

Theorem 2.2.11 (Stirling's formula) $n! \sim \sqrt{2\pi}\, n^{n+1/2} \mathrm{e}^{-n}$.

Proof. Under the notation of the previous lemma, put

$$b_n = \mathrm{e}^{a_n} = \frac{n!}{n^{n+1/2} \mathrm{e}^{-n}} \quad \text{for } n = 1, 2, \ldots .$$

Then the sequence $(b_n)_n$ converges to some $b > 0$. Thus

$$\frac{b_n^2}{b_{2n}} = \frac{2^{2n+1/2}(n!)^2}{n^{1/2}(2n)!} \to \frac{b^2}{b} = b \quad \text{as } n \to \infty.$$

For $n = 1, 2, \ldots$, let $c_n = \dfrac{n!\, n^{1/2}}{(n + \frac{1}{2}) \cdots \frac{3}{2} \cdot \frac{1}{2}}$. Then by Corollary 2.2.5, $(c_n)_n$ converges to $\Gamma(1/2) = \sqrt{\pi}$ as $n \to \infty$. Hence

$$\frac{b_n^2}{b_{2n}} = c_n \Big(1 + \frac{1}{2n}\Big)\sqrt{2} \to \sqrt{2\pi} \quad \text{as } n \to \infty,$$

which yields $b = \sqrt{2\pi}$. Consequently,

$$b_n = \frac{n!}{n^{n+1/2}\mathrm{e}^{-n}} \to \sqrt{2\pi} \quad \text{as} \quad n \to \infty$$

and the proof is now complete. □

Closely related to the gamma function is the *beta function* B, which is the real function of two variables defined by the formula

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt \quad \text{for } x, y > 0.$$

Theorem 2.2.12 *The beta function has the following properties:*

(i) $B(x, y) = B(y, x)$ *and* $B(x + 1, y) = \frac{x}{x+y} B(x, y)$*;*
(ii) $B(x, y)$ *is a* log*-convex function of* x *for each fixed* y*;*
(iii) $B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x + y)}$.

Proof. (i) The first formula is clear. For the second,

$$\begin{aligned} B(x+1, y) &= \int_0^1 t^x (1-t)^{y-1}\, dt \\ &= \int_0^1 (1-t)^{x+y-1} \Big(\frac{t}{1-t}\Big)^x dt \\ &= \Big[\frac{-(1-t)^{x+y}}{x+y}\Big(\frac{t}{1-t}\Big)^x\Big]_{t=0}^{t=1} + \int_0^1 \frac{x}{x+y}\, t^{x-1}(1-t)^{y-1}\, dt \\ &= \frac{x}{x+y}\, B(x, y). \end{aligned}$$

(ii) Let $a, b, y > 0$ and let $\lambda, \mu \geq 0$ with $\lambda + \mu = 1$. By the Rogers–Hölder inequality,

$$\begin{aligned} B(\lambda a + \mu b, y) &= \int_0^1 (t^{a-1}(1-t)^{y-1})^\lambda (t^{b-1}(1-t)^{y-1})^\mu \, dt \\ &\leq \left(\int_0^1 t^{a-1}(1-t)^{y-1}\, dt\right)^\lambda \left(\int_0^1 t^{b-1}(1-t)^{y-1}\, dt\right)^\mu \\ &= B^\lambda(a,y) \cdot B^\mu(a,y). \end{aligned}$$

(iii) Let $y > 0$ be arbitrarily fixed and consider the function

$$\varphi_y(x) = \frac{\Gamma(x+y)B(x,y)}{\Gamma(y)}, \quad x > 0.$$

Then φ_y is a product of log-convex functions and so it is itself log-convex. Also,

$$\begin{aligned} \varphi_y(x+1) &= \frac{\Gamma(x+y+1)B(x+1,y)}{\Gamma(y)} \\ &= \frac{[(x+y)\Gamma(x+y)][x/(x+y)]B(x,y)}{\Gamma(y)} = x\varphi_y(x) \end{aligned}$$

for all $x > 0$ and

$$\varphi_y(1) = \frac{\Gamma(1+y)B(1,y)}{\Gamma(y)} = y\int_0^1 (1-t)^{y-1}\, dt = 1.$$

Thus $\varphi_y = \Gamma$ by Theorem 2.2.4, and the assertion (iii) is now clear. □

Exercises

1. Prove that $\Gamma(n+\frac{1}{2}) = \frac{(2n)!\,\sqrt{\pi}}{n!\,4^n}$ for $n \in \mathbb{N}$.

2. The integrals

$$I_n = \int_0^{\pi/2} \sin^n t\, dt \quad (\text{for } n \in \mathbb{N})$$

can be computed easily via the recurrence formula $nI_n = (n-1)I_{n-2}$ (where $n \geq 2$). Integrate the inequalities $\sin^{2n+1} x \leq \sin^{2n} x \leq \sin^{2n-1} x$ over $[0, \pi/2]$ to infer *Wallis' formula*,

$$\frac{\pi}{2} = \lim_{n\to\infty} \Big[\frac{2\cdot 2}{1\cdot 3}\cdot\frac{4\cdot 4}{3\cdot 5}\cdots\frac{2n\cdot 2n}{(2n-1)\cdot(2n+1)}\Big].$$

Remark. An alternative proof of this formula follows from Corollary 2.2.5, by noticing that $\pi/2 = (\Gamma(1/2))^2/2$.

3. Establish the formula

$$B(x,y) = 2\int_0^{\pi/2} \sin^{2x-1} t \cdot \cos^{2y-1} t \, dt \quad \text{for } x, y > 0,$$

and infer from it that

$$\int_0^{\pi/2} \sin^{2n} t \, dt = \frac{(2n)!\,\pi}{2^{2n+1}(n!)^2} \quad \text{for } n \in \mathbb{N}.$$

4. Use Corollary 2.2.5 to prove Weierstrass' formula,

$$\Gamma(x) = \frac{\mathrm{e}^{-\gamma x}}{x} \prod_{n=1}^{\infty} \Big(1 + \frac{x}{n}\Big)^{-1} \mathrm{e}^{x/n},$$

where $\gamma = \lim_{n\to\infty} (1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n) = 0.57722\ldots$ is Euler's constant.

5. (The Raabe integral) Prove that

$$\Gamma\Big(\frac{1}{p}\Big)\Gamma\Big(\frac{2}{p}\Big)\cdots\Gamma\Big(\frac{p-1}{p}\Big) = \frac{(2\pi)^{p-1/2}}{p^{1/2}} \quad \text{for all } p \in \mathbb{N}^{\star}.$$

Then infer the integral formula

$$\int_x^{x+1} \log\Gamma(t)\, dt = x(\log x - 1) + \frac{1}{2}\log 2\pi \quad \text{for all } x \geq 0.$$

[*Hint*: Notice that $\int_x^{x+1} \log\Gamma(t)\, dt - x(\log x - 1)$ is constant. The value at $x = 0$ can be computed by using Riemann sums.]

6. (L. Euler) Prove the formula

$$\int_0^{\infty} \frac{t^{x-1}}{1+t}\, dt = \frac{\pi}{\sin \pi x} \quad \text{for } 0 < x < 1.$$

[*Hint*: Put $t = u/(1-u)$ and apply Theorem 2.2.12 (iii).]

7. (An alternative proof of the log-convexity of Γ) Prove the formula

$$\frac{d^2}{dx^2} \log\Gamma(x) = \sum_{n=0}^{\infty} \frac{1}{(x+n)^2} \quad \text{for } x > 0.$$

8. (F. John's approach of the Bohr–Mollerup theorem) Let g be a real-valued concave function on $(0,\infty)$ such that $g(x)/x \to 0$ as $x \to \infty$.

(i) Prove that the difference equation

$$f(x+1) - f(x) = g(x)$$

has one and only one convex solution $f\colon (0,\infty) \to \mathbb{R}$ with $f(1) = 0$, and this solution is given by the formula

$$f(x) = -g(x) + x \cdot \lim_{n\to\infty} \Big(g(n) - \sum_{k=1}^{n-1} \frac{g(x+k) - g(k)}{x} \Big).$$

(ii) (A Stirling type formula) Prove the existence of the limit

$$c = \lim_{x\to\infty} \Big(f(x) + g(x) - \int_{1/2}^{x+1/2} g(t)\,dt \Big).$$

Remark. The Bohr–Mollerup theorem concerns the case where $g = \log$ and $f = \log\Gamma$.

9. (E. Artin [10]) Let U be an open convex subset of $\mathbb{R}^n$ and let μ be a Borel measure on an interval I. Consider the integral transform

$$F(x) = \int_I K(x,t)\,d\mu(t),$$

where the kernel $K(x,t)\colon U \times I \to [0,\infty)$ satisfies the following two conditions:

(i) $K(x,t)$ is μ-integrable in t for each fixed x;
(ii) $K(x,t)$ is log-convex in x for each fixed t.

Prove that F is log-convex on U.
[*Hint*: Apply the Rogers–Hölder inequality, noticing that

$$K((1-\lambda)x + \lambda y, t) \le (K(x,t))^{1-\lambda}(K(y,t))^{\lambda}. \quad]$$

Remark. The *Laplace transform* of a function $f \in L^1(0,\infty)$ is given by the formula $(\mathcal{L}f)(x) = \int_0^\infty f(t)\mathrm{e}^{-tx}\,dt$. By Exercise 9, the Laplace transform of any nonnegative function is log-convex. In the same way one can show that the moment $\mu_\alpha = \int_0^\infty t^\alpha f(t)\,dt$, of any random variable with probability density f, is a log-convex function in α (on each subinterval of $[0,\infty)$ where it is finite).

2.3 Generalities on Multiplicatively Convex Functions

The class of multiplicatively convex functions can be easily described as being constituted by those functions f (acting on subintervals of $(0,\infty)$) such that $\log f(x)$ is a convex function of $\log x$:

Lemma 2.3.1 *Suppose that* $f\colon I \to (0,\infty)$ *is a function defined on a subinterval of* $(0,\infty)$. *Then* f *is multiplicatively convex if and only if*

$$\begin{vmatrix} 1 & \log x_1 & \log f(x_1) \\ 1 & \log x_2 & \log f(x_2) \\ 1 & \log x_3 & \log f(x_3) \end{vmatrix} \geq 0$$

for all $x_1 \leq x_2 \leq x_3$ *in* I; *equivalently, if and only if*

$$f(x_1)^{\log x_3} f(x_2)^{\log x_1} f(x_3)^{\log x_2} \geq f(x_1)^{\log x_2} f(x_2)^{\log x_3} f(x_3)^{\log x_1}$$

for all $x_1 \leq x_2 \leq x_3$ *in* I.

This is nothing but the translation (via Lemma 2.1.1) of the result of Lemma 1.3.2.

In the same spirit, we can show that every multiplicatively convex function $f\colon I \to (0,\infty)$ has finite lateral derivatives at each interior point of I (and the set of all points where f is not differentiable is at most countable). As a consequence, every multiplicatively convex function is continuous in the interior of its domain of definition. Under the presence of continuity, the multiplicative convexity can be restated in terms of geometric mean:

Theorem 2.3.2 *Suppose that* I *is a subinterval of* $(0,\infty)$. *A continuous function* $f\colon I \to (0,\infty)$ *is multiplicatively convex if and only if*

$$x, y \in I \text{ implies } f(\sqrt{xy}) \leq \sqrt{f(x)f(y)}.$$

Proof. The necessity is clear. The sufficiency part follows from the connection between the multiplicative convexity and the usual convexity (as noted in Lemma 2.1.1) and the fact that midpoint convexity is equivalent to convexity in the presence of continuity. See Theorem 1.1.3. □

Theorem 2.3.2 reveals the essence of multiplicative convexity as being the *convexity according to the geometric mean*; in fact, under the presence of continuity, the multiplicatively convex functions are precisely those functions $f\colon I \to (0,\infty)$ for which

$$x_1, \dots, x_n \in I \text{ implies } f(\sqrt[n]{x_1 \cdots x_n}) \leq \sqrt[n]{f(x_1) \cdots f(x_n)}.$$

In this respect, it is natural to call a function $f\colon I \to (0,\infty)$ *multiplicatively concave* if $1/f$ is multiplicatively convex, and *multiplicatively affine* if f is of the form Cx^{α} for some $C > 0$ and some $\alpha \in \mathbb{R}$.

A refinement of the notion of multiplicative convexity is that of *strict multiplicative convexity*, which in the context of continuity will mean

$$f(\sqrt[n]{x_1 \cdots x_n}) < \sqrt[n]{f(x_1) \cdots f(x_n)}$$

unless $x_1 = \cdots = x_n$. Clearly, Lemma 2.1.1 (which relates the multiplicatively convex functions and the usual convex functions) has a "strict" counterpart.

A large class of strictly multiplicatively convex functions is indicated by the following result:

Proposition 2.3.3 (G. H. Hardy, J. E. Littlewood and G. Pólya [99, Theorem 177, p. 125]) *Every polynomial $P(x)$ with nonnegative coefficients is a multiplicatively convex function on $(0,\infty)$. More generally, every real analytic function $f(x)=\sum_{n=0}^{\infty} c_n x^n$ with nonnegative coefficients is a multiplicatively convex function on $(0,R)$, where R denotes the radius of convergence.*

Moreover, except for the case of functions Cx^n (with $C>0$ and $n\in\mathbb{N}$), the above examples exhibit strictly multiplicatively convex functions (which are also increasing and strictly convex). In particular,

- exp, sinh and cosh on $(0,\infty)$;
- tan, sec, csc and $\frac{1}{x}-\cot x$ on $(0,\pi/2)$;
- arcsin on $(0,1]$;
- $-\log(1-x)$ and $\frac{1+x}{1-x}$ on $(0,1)$.

See the table of series in I. S. Gradshteyn and I. M. Ryzhik [89].

Proof. By continuity, it suffices to prove only the first assertion. Suppose that $P(x)=\sum_{n=0}^{N} c_n x^n$. According to Theorem 2.3.2, we have to prove that

$$x,y>0 \text{ implies } (P(\sqrt{xy}))^2 \le P(x)P(y),$$

or, equivalently,

$$x,y>0 \text{ implies } (P(xy))^2 \le P(x^2)P(y^2).$$

The later implication is an easy consequence of Cauchy–Buniakovski–Schwarz inequality. □

The following result collects a series of useful remarks for proving the multiplicative convexity of concrete functions:

Lemma 2.3.4

(i) *If a function is* log*-convex and increasing, then it is strictly multiplicatively convex.*
(ii) *If a function f is multiplicatively convex, then the function $1/f$ is multiplicatively concave* (*and vice versa*).
(iii) *If a function f is multiplicatively convex, increasing and one-to-one, then its inverse is multiplicatively concave* (*and vice versa*).
(iv) *If a function f is multiplicatively convex, so is $x^\alpha[f(x)]^\beta$ (for all $\alpha\in\mathbb{R}$ and all $\beta>0$).*
(v) *If f is continuous, and one of the functions $f(x)^x$ and $f(\mathrm{e}^{1/\log x})$ is multiplicatively convex, then so is the other.*

In many cases the inequalities based on multiplicative convexity are better than the direct application of the usual inequalities of convexity (or yield complementary information). This includes the multiplicative analogue of the Hardy–Littlewood–Pólya inequality of majorization:

Proposition 2.3.5 *Suppose that* $x_1 \geq x_2 \geq \cdots \geq x_n$ *and* $y_1 \geq y_2 \geq \cdots \geq y_n$ *are two families of numbers in a subinterval* I *of* $(0, \infty)$ *such that*

$$\begin{aligned} x_1 &\geq y_1 \\ x_1 x_2 &\geq y_1 y_2 \\ &\vdots \\ x_1 x_2 \cdots x_{n-1} &\geq y_1 y_2 \cdots y_{n-1} \\ x_1 x_2 \cdots x_n &= y_1 y_2 \cdots y_n. \end{aligned}$$

Then

$$f(x_1) f(x_2) \cdots f(x_n) \geq f(y_1) f(y_2) \cdots f(y_n)$$

for every multiplicatively convex function $f \colon I \to (0, \infty)$.

A result due to H. Weyl [245] (see also [155]) gives us the basic example of a pair of sequences satisfying the hypothesis of Proposition 2.3.5: *Consider a matrix* $A \in \mathrm{M}_n(\mathbb{C})$ *having the eigenvalues* $\lambda_1, \ldots, \lambda_n$ *and the singular numbers* $s_1, \ldots, s_n$*, and assume that they are rearranged such that* $|\lambda_1| \geq \cdots \geq |\lambda_n|$*, and* $s_1 \geq \cdots \geq s_n$*. Then:*

$$\Big| \prod_{k=1}^{m} \lambda_k \Big| \leq \prod_{k=1}^{m} s_k \quad \text{for } m = 1, \ldots, n-1 \quad \text{and} \quad \Big| \prod_{k=1}^{n} \lambda_k \Big| = \prod_{k=1}^{n} s_k.$$

Recall that the *singular numbers* of a matrix A are precisely the eigenvalues of its modulus, $|A| = (A^{\star} A)^{1/2}$; the spectral mapping theorem assures that $s_k = |\lambda_k|$ when A is Hermitian. The fact that all examples come this way was noted by A. Horn; see [155] for details.

According to the discussion above the following result holds:

Proposition 2.3.6 *Let* $A \in \mathrm{M}_n(\mathbb{C})$ *be any matrix having the eigenvalues* $\lambda_1, \ldots, \lambda_n$ *and the singular numbers* $s_1, \ldots, s_n$*, listed such that* $|\lambda_1| \geq \cdots \geq |\lambda_n|$ *and* $s_1 \geq \cdots \geq s_n$*. Then*

$$\prod_{k=1}^{n} f(s_k) \geq \prod_{k=1}^{n} f(|\lambda_k|)$$

for every multiplicatively convex function f *which is continuous on* $[0, \infty)$.

In general it is not true that $|\lambda_k| \leq s_k$ for all k. A counterexample is given by the matrix

$$\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$$

whose eigenvalues are $\lambda_1 = 2 > \lambda_2 = -2$ and the singular numbers are $s_1 = 4 > s_2 = 1$.

Exercises

1. (C. H. Kimberling [126]) Suppose that P is a polynomial with nonnegative coefficients. Prove that

$$(P(1))^{n-1}P(x_1\cdots x_n)\ge P(x_1)\cdots P(x_n)$$

provided that all x_k are either in $[0,1]$ or in $[1,\infty)$. This fact complements Proposition 2.3.3.

2. (The multiplicative analogue of Popoviciu's inequality) Suppose there is given a multiplicatively convex function $f\colon I\to(0,\infty)$. Infer from Theorem 2.3.5 that

$$f(x)\,f(y)\,f(z)\,f^3(\sqrt[3]{xyz})\ge f^2(\sqrt{xy})f^2(\sqrt{yz})f^2(\sqrt{zx})$$

for all $x,y,z\in I$. Moreover, for the strictly multiplicatively convex functions the equality occurs only when $x=y=z$.

3. Recall that the inverse sine function is strictly multiplicatively convex on $(0,1]$ and infer the following two inequalities in a triangle ΔABC:

$$\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2}<\Big(\sin\Big(\frac{1}{2}\sqrt[3]{ABC}\Big)\Big)^3<\frac{1}{8}$$
$$\sin A\sin B\sin C<(\sin\sqrt[3]{ABC})^3<\frac{3\sqrt{3}}{8}$$

unless $A=B=C$.

4. (P. Montel [171]) Let $I\subset(0,\infty)$ be an interval and suppose that f is a continuous and positive function on I. Prove that f is multiplicatively convex if and only if

$$2f(x)\le k^{\alpha}f(kx)+k^{-\alpha}f(x/k)$$

for all $\alpha\in\mathbb{R}$, $x\in I$, and $k>0$, such that kx and x/k both belong to I.

5. (The multiplicative mean) According to Lemma 2.1.1, the multiplicative analog of the arithmetic mean is

$$\begin{aligned}M_*(f)&=\exp\Big(\frac{1}{\log b-\log a}\int_{\log a}^{\log b}\log f(\mathrm{e}^t)\,dt\Big)\\&=\exp\Big(\frac{1}{\log b-\log a}\int_a^b\log f(t)\frac{dt}{t}\Big),\end{aligned}$$

that is, the geometric mean of f with respect to the measure dt/t. Notice that

$$\begin{gathered}M_*(1)=1\\ \inf f\le f\le\sup f\Rightarrow\inf f\le M_*(f)\le\sup f\\ M_*(fg)=M_*(f)M_*(g).\end{gathered}$$

(i) Let $f\colon [a,b] \to (0,\infty)$ be a continuous function defined on a subinterval of $(0,\infty)$ and let $\varphi\colon J \to (0,\infty)$ be a multiplicatively convex continuous function defined on an interval J which includes the image of f. Prove that

$$\varphi(M_*(f)) \le M_*(\varphi \circ f),$$

which is the multiplicative analogue of Jensen's inequality.

(ii) Suppose that $0 < a < b$ and let $f\colon [a,b] \to (0,\infty)$ be a multiplicatively convex continuous function. Prove the following analogue of Hermite–Hadamard inequality,

$$f(\sqrt{ab}) \le M_*(f) \le \sqrt{f(a)f(b)};$$

the left-hand side inequality is strict unless f is multiplicatively affine, while the right-hand side inequality is strict unless f is multiplicatively affine on each of the subintervals $[a,\sqrt{ab}]$ and $[\sqrt{ab},b]$. These inequalities can be improved following an idea similar to that of Remark 1.9.3:

$$\begin{aligned} f(a^{1/2}b^{1/2}) &\le (f(a^{3/4}b^{1/4})f(a^{1/4}b^{3/4}))^{1/2} \le M_*(f) \\ &\le (f(a^{1/2}b^{1/2}))^{1/2} f(a)^{1/4} f(b)^{1/4} \\ &\le (f(a)f(b))^{1/2}. \end{aligned}$$

(iii) Notice that $M_*(f) = \exp(\frac{b-a}{\log b - \log a})$ for $f = \exp|_{[a,b]}$ $(0 < a < b)$. Then, infer from (ii) the inequalities:

$$\frac{a^{3/4}b^{1/4} + a^{1/4}b^{3/4}}{2} < \frac{b-a}{\log b - \log a} < \frac{1}{2}\Big(\frac{a+b}{2} + \sqrt{ab}\Big)$$

$$\exp\Big(\frac{b-a}{\log b - \log a}\Big) < \frac{e^b - e^a}{b-a}.$$

6. Let $f\colon I \to (0,\infty)$ be a function which is multiplicatively convex or multiplicatively concave and let $a > 0$.

(i) Prove that

$$\Big(\prod_{k=1}^{n} f(a^{k/n})\Big)^{1/n} > \Big(\prod_{k=1}^{n+1} f(a^{k/(n+1)})\Big)^{1/(n+1)} > M_*(f)$$

for all $n = 1, 2, 3, \ldots$ in each of the following two cases:

- $I = [1,a]$ (with $a > 1$) and f is increasing;
- $I = [a,1]$ (with $0 < a < 1$) and f is decreasing.

(ii) Prove that the above inequalities will be reversed in each of the following two cases:

- $I = [1,a]$ (with $a > 1$) and f is decreasing;
- $I = [a,1]$ (with $0 < a < 1$) and f is increasing.

(iii) Illustrate the assertions (i) and (ii) in the case of the functions $1+\log x$ and $\exp x$, for $x \ge 1$, and $\sin(\pi x/2)$ and $\cos(\pi x/2)$, for $x \in (0,1]$.

2.4 Multiplicative Convexity of Special Functions

We start this section by noticing that the indefinite integral of a multiplicatively convex function has the same nature:

Proposition 2.4.1 (P. Montel [171]) *Let* $f\colon [0,a) \to [0,\infty)$ *be a continuous function which is multiplicatively convex on* $(0,a)$. *Then*

$$F(x) = \int_0^x f(t)\,dt$$

is also continuous on $[0,a)$ *and multiplicatively convex on* $(0,a)$.

Proof. Due to the continuity of F, it suffices to show that

$$(F(\sqrt{xy}))^2 \le F(x)F(y) \quad \text{for all } x, y \in [0,a),$$

which is a consequence of the corresponding inequality at the level of integral sums,

$$\Big[\frac{\sqrt{xy}}{n}\sum_{k=0}^{n-1} f\Big(k\,\frac{\sqrt{xy}}{n}\Big)\Big]^2 \le \Big[\frac{x}{n}\sum_{k=0}^{n-1} f\Big(k\,\frac{x}{n}\Big)\Big]\Big[\frac{y}{n}\sum_{k=0}^{n-1} f\Big(k\,\frac{y}{n}\Big)\Big],$$

that is, of the inequality

$$\Big[\sum_{k=0}^{n-1} f\Big(k\,\frac{\sqrt{xy}}{n}\Big)\Big]^2 \le \Big[\sum_{k=0}^{n-1} f\Big(k\,\frac{x}{n}\Big)\Big]\Big[\sum_{k=0}^{n-1} f\Big(k\,\frac{y}{n}\Big)\Big].$$

To see that the later inequality holds, first notice that

$$\Big[f\Big(k\,\frac{\sqrt{xy}}{n}\Big)\Big]^2 \le \Big[f\Big(k\,\frac{x}{n}\Big)\Big]\Big[f\Big(k\,\frac{y}{n}\Big)\Big]$$

and then apply the Cauchy–Buniakovski–Schwarz inequality. □

According to Proposition 2.4.1, the *logarithmic integral*,

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}, \quad x \ge 2,$$

is multiplicatively convex. This function is important in number theory. For example, if $\pi(x)$ counts the number of primes p such that $2 \le p \le x$, then an equivalent formulation of the Riemann hypothesis is the existence of a function $C\colon (0,\infty) \to (0,\infty)$ such that

$$|\pi(x) - \mathrm{Li}(x)| \le C(\varepsilon)x^{1/2+\varepsilon} \quad \text{for all } x \ge 2 \text{ and all } \varepsilon > 0.$$

Since the function tan is continuous on $[0, \pi/2)$ and strictly multiplicatively convex on $(0, \pi/2)$, a repeated application of Proposition 2.4.1 shows that the *Lobacevski's function*

$$\mathrm{L}(x) = -\int_0^x \log \cos t \, dt$$

is strictly multiplicatively convex on $(0, \pi/2)$.

Starting with $t/(\sin t)$, (which is strictly multiplicatively convex on $(0, \pi/2]$) and then switching to $(\sin t)/t$, a similar argument leads us to the fact that the *integral sine* function,

$$\mathrm{Si}(x) = \int_0^x \frac{\sin t}{t} \, dt,$$

is strictly multiplicatively concave on $(0, \pi/2]$.

Another striking fact is the following:

Proposition 2.4.2 Γ *is a strictly multiplicatively convex function on* $[1, \infty)$.

Proof. In fact, $\log \Gamma(1 + x)$ is strictly convex and increasing on $(1, \infty)$. Moreover, an increasing strictly convex function of a strictly convex function is strictly convex. Hence, $F(x) = \log \Gamma(1 + \mathrm{e}^x)$ is strictly convex on $(0, \infty)$ and thus $\Gamma(1 + x) = \exp F(\log x)$ is strictly multiplicatively convex on $[1, \infty)$. As $\Gamma(1 + x) = x\Gamma(x)$, we conclude that Γ itself is strictly multiplicatively convex on $[1, \infty)$. □

According to Proposition 2.4.2,

$$\Gamma^3(\sqrt[3]{xyz}) < \Gamma(x)\Gamma(y)\Gamma(z) \quad \text{for all } x, y, z \geq 1$$

except the case where $x = y = z$.

On the other hand, by the multiplicative version of Popoviciu's inequality (Exercise 2, Section 2.3), we infer that

$$\Gamma(x)\Gamma(y)\Gamma(z)\Gamma^3(\sqrt[3]{xyz}) \geq \Gamma^2(\sqrt{xy})\Gamma^2(\sqrt{yz})\Gamma^2(\sqrt{zx})$$

for all $x, y, z \geq 1$; the equality occurs only for $x = y = z$.

Another application of Proposition 2.4.2 is the fact that the function $\Gamma(2x+1)/\Gamma(x+1)$ is strictly multiplicatively convex on $[1, \infty)$. This can be seen by using the Gauss–Legendre duplication formula given by Theorem 2.2.9.

Exercises

1. (D. Gronau and J. Matkowski [90]) Prove the following converse to Proposition 2.4.2: If $f\colon (0, \infty) \to (0, \infty)$ verifies the functional equation

$$f(x + 1) = xf(x),$$

the normalization condition $f(1) = 1$, and f is multiplicatively convex on an interval (a, ∞), for some $a > 0$, then $f = \Gamma$.

2. Let $f\colon I \to (0,\infty)$ be a differentiable function defined on a subinterval I of $(0,\infty)$. Prove that the following assertions are equivalent:

(i) f is multiplicatively convex;
(ii) the function $xf'(x)/f(x)$ is nondecreasing;
(iii) f verifies the inequality

$$\frac{f(x)}{f(y)} \geq \left(\frac{x}{y}\right)^{yf'(y)/f(y)} \quad \text{for all } x, y \in I.$$

A similar statement works for the multiplicatively concave functions. Illustrate this fact by considering the restriction of $\sin(\cos x)$ to $(0,\pi/2)$.

3. The *psi function* (also known as the *digamma function*) is defined by

$$\operatorname{Psi}(x) = \frac{d}{dx}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}, \quad x > 0$$

and it can be represented as

$$\operatorname{Psi}(x) = -\gamma - \int_0^1 \frac{t^{x-1}-1}{1-t}\,dt,$$

where γ is Euler's constant. See [9], [89].

(i) Prove that the function Psi satisfies the functional equation

$$\psi(x+1) = \psi(x) + \frac{1}{x}.$$

(ii) Infer from Proposition 2.4.2 and the preceding exercise the inequality

$$\frac{\Gamma(x)}{\Gamma(y)} \geq \left(\frac{x}{y}\right)^{y\operatorname{Psi}(y)} \quad \text{for all } x, y \geq 1.$$

4. Let $f\colon I \to (0,\infty)$ be a twice differentiable function defined on a subinterval I of $(0,\infty)$. Prove that f is multiplicatively convex if and only if it verifies the differential inequality

$$x[f(x)f''(x) - f'^2(x)] + f(x)f'(x) \geq 0 \quad \text{for all } x > 0.$$

Infer that the integral sine function is multiplicatively concave.

2.5 An Estimate of the AM–GM Inequality

Suppose that I is a subinterval of $(0,\infty)$ and that $f\colon I \to (0,\infty)$ is a twice differentiable function. According to Lemma 2.1.1, the values of the parameter $\alpha \in \mathbb{R}$ for which the function

$$\varphi(x) = f(x) \cdot x^{(-\alpha/2)\log x}$$

is multiplicatively convex on I are precisely those for which the function

$$\Phi(x) = \log\varphi(\mathrm{e}^x) = \log f(\mathrm{e}^x) - \frac{\alpha x^2}{2}$$

is convex on $\log(I)$. Since the convexity of Φ is equivalent to $\Phi'' \geq 0$, we infer that φ is multiplicatively convex if and only if $\alpha \leq \alpha(f)$, where

$$\begin{aligned}\alpha(f) &= \inf_{x\in\log(I)} \frac{d^2}{dx^2}\log f(\mathrm{e}^x)\\ &= \inf_{x\in I}\frac{x^2\big(f(x)f''(x) - (f'(x))^2\big) + xf(x)f'(x)}{f(x)^2}.\end{aligned}$$

By considering also the upper bound

$$\beta(f) = \sup_{x\in\log(I)} \frac{d^2}{dx^2}\log f(\mathrm{e}^x),$$

we arrive at the following result:

Lemma 2.5.1 *Under the above hypotheses, we have*

$$\begin{aligned}\exp\Big(\frac{\alpha(f)}{2n^2}\sum_{1\leq j<k\leq n}(\log x_j - \log x_k)^2\Big) &\leq \Big(\prod_{k=1}^n f(x_k)\Big)^{1/n}\Big/ f\Big(\Big(\prod_{k=1}^n x_k\Big)^{1/n}\Big)\\ &\leq \exp\Big(\frac{\beta(f)}{2n^2}\sum_{1\leq j<k\leq n}(\log x_j - \log x_k)^2\Big)\end{aligned}$$

for all $x_1,\ldots,x_n \in I$.

Particularly, for $f(x) = \mathrm{e}^x$, $x \in [A,B]$ (where $0 < A \leq B$), we have $\alpha(f) = A$ and $\beta(f) = B$, and we are led to the following improvement upon the AM–GM inequality:

Lemma 2.5.2 *Suppose that* $0 < A \leq B$ *and* $n \in \mathbb{N}^\star$. *Then*

$$\begin{aligned}\frac{A}{2n^2}\sum_{1\leq j<k\leq n}(\log x_j - \log x_k)^2 &\leq \frac{1}{n}\sum_{k=1}^n x_k - \Big(\prod_{k=1}^n x_k\Big)^{1/n}\\ &\leq \frac{B}{2n^2}\sum_{1\leq j<k\leq n}(\log x_j - \log x_k)^2\end{aligned}$$

for all $x_1,\ldots,x_n \in [A,B]$.

Since

$$\frac{1}{2n^2} \sum_{1 \le j < k \le n} (\log x_j - \log x_k)^2$$

represents the *variance* of the random variable whose distribution is

$$\begin{pmatrix} \log x_1 & \log x_2 & \dots & \log x_n \\ 1/n & 1/n & \dots & 1/n \end{pmatrix},$$

Lemma 2.5.2 reveals the probabilistic character of the AM–GM inequality. Using the usual device to approximate the integrable functions by step functions, we can derive from Lemma 2.5.2 the following more general result:

Theorem 2.5.3 *Let* (Ω, Σ, P) *be a probability space and let* X *be a random variable on this space, taking values in the interval* $[A, B]$*, where* $0 < A \le B$*. Then*

$$A \le \frac{\mathcal{E}(X) - e^{\mathcal{E}(\log X)}}{\operatorname{var}(\log X)} \le B.$$

Here $\mathcal{E}(Z) = \int_X Z(\omega)\, dP(\omega)$ represents the *mathematical expectation* of the random variable Z, and $\operatorname{var}(Z) = \mathcal{E}\big((Z - \mathcal{E}(Z))^2\big)$ the *variance* of Z.

Exercises

1. (H. Kober; see [166, p. 81]) Suppose that $x_1, \dots, x_n$ are distinct positive numbers, and $\lambda_1, \dots, \lambda_n$ are positive numbers such that $\lambda_1 + \dots + \lambda_n = 1$. Prove that

$$\frac{A(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n) - G(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n)}{\sum_{i<j} (\sqrt{x_i} - \sqrt{x_j})^2}$$

lies between $\inf_i \lambda_i/(n-1)$ and $\sup_i \lambda_i$.

2. (P. H. Diananda; see [166, p. 83]) Under the same hypothesis as in the precedent exercise, prove that

$$\frac{A(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n) - G(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n)}{\sum_{i<j} \lambda_i \lambda_j (\sqrt{x_i} - \sqrt{x_j})^2}$$

lies between $1/(1 - \inf_i \lambda_i)$ and $1/\inf_i \lambda_i$.

3. Suppose that $x_1, \dots, x_n$ and $\lambda_1, \dots, \lambda_n$ are positive numbers for which $\lambda_1 + \dots + \lambda_n = 1$. Put $A_n = A(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n)$ and $G_n = G(x_1, \dots, x_n\,; \lambda_1, \dots, \lambda_n)$.

(i) Compute the integral

$$J(x, y) = \int_0^\infty \frac{t\, dt}{(1+t)(x+yt)^2}.$$

(ii) Infer that $A_n/G_n = \exp(\sum_{k=1}^n \lambda_k (x_k - A_n)^2 J(x_k, A_n))$.

2.6 (M, N)-Convex Functions

The four algebraic variants of convexity we considered in the preceding sections can be embedded into a more general framework, by taking two regular means M and N (on the intervals I and J respectively) and calling a function $f\colon I \to J$ to be (M, N)-*midpoint convex* if it satisfies

$$f(M(x, y)) \leq N(f(x), f(y))$$

for all $x, y \in I$. As noticed in the Introduction, if f is continuous, this yields the (M, N)-convexity of f, that is,

$$f(M(x, y\,; 1 - \lambda, \lambda)) \leq N(f(x), f(y)\,; 1 - \lambda, \lambda)$$

for all $x, y \in I$ and all $\lambda \in [0, 1]$. The sundry notions such as (M, N)-strict convexity and (M, N)-concavity can be introduced in a natural way.

Many important results, such as the left-hand side of the Hermite–Hadamard inequality and the Jensen inequality, extend to this framework. See Theorems A, B and C in the Introduction.

Other results, like Lemma 2.1.1, can be extended only in the context of quasi-arithmetic means:

Lemma 2.6.1 (J. Aczél [2]) *If φ and ψ are two continuous and strictly monotonic functions (on intervals I and J respectively) and ψ is increasing, then a function $f\colon I \to J$ is $(M_{[\varphi]}, M_{[\psi]})$-convex if and only if $\psi \circ f \circ \varphi^{-1}$ is convex on $\varphi(I)$ in the usual sense.*

Proof. In fact, f is $(M_{[\varphi]}, M_{[\psi]})$-convex if and only if

$$\psi\big(f\big(\varphi^{-1}((1 - \lambda)u + \lambda v)\big)\big) \leq (1 - \lambda)\psi\big(f(\varphi^{-1}(u))\big) + \lambda\psi\big(f(\varphi^{-1}(v))\big)$$

for all $u, v \in \varphi(I)$ and $\lambda \in [0, 1]$. $\square$

A nice illustration of Lemma 2.6.1 was recently given by D. Borwein, J. Borwein, G. Fee and R. Girgensohn [35], who proved that the volume $V_n(p)$ of the ellipsoid $\{x \in \mathbb{R}^n \mid \|x\|_{L^p} \leq 1\}$ is (H, G)-strictly concave as a function of p:

Theorem 2.6.2 *Given $\alpha > 1$, the function $V_\alpha(p) = 2^\alpha \frac{\Gamma(1+1/p)^\alpha}{\Gamma(1+\alpha/p)}$ verifies the inequality*

$$V_\alpha^{1-\lambda}(p) V_\alpha^{\lambda}(q) < V_\alpha\left(\frac{1}{\frac{1-\lambda}{p} + \frac{\lambda}{q}}\right),$$

for all $p, q > 0$, $p \neq q$ and all $\lambda \in (0, 1)$.

Proof. According to Lemma 2.6.1 it suffices to prove that the function

$$U_\alpha(x) = -\log(V_\alpha(1/x)/2^\alpha) = \log\Gamma(1+\alpha x) - \alpha\log\Gamma(1+x)$$

is strictly convex on $(0,\infty)$ for every $\alpha > 1$. Using the psi function,

$$\mathrm{Psi}(x) = \frac{d}{dx}\,\log\Gamma(x),$$

we have

$$U''_\alpha(x) = \alpha^2\frac{d}{dx}\,\mathrm{Psi}(1+\alpha x) - \alpha\frac{d}{dx}\,\mathrm{Psi}(1+x).$$

Then $U''_\alpha(x) > 0$ on $(0,\infty)$ means $(x/\alpha)U''_\alpha(x) > 0$ on $(0,\infty)$, and the latter holds if the function $x \to x\frac{d}{dx}\,\mathrm{Psi}(1+x)$ is strictly increasing. Or, according to [9], [89],

$$\frac{d}{dx}\,\mathrm{Psi}(1+x) = \int_0^\infty \frac{u\mathrm{e}^{ux}}{\mathrm{e}^u - 1}\,du,$$

and an easy computation shows that

$$\frac{d}{dx}\Big(x\frac{d}{dx}\,\mathrm{Psi}(1+x)\Big) = \int_0^\infty \frac{u[(u-1)\mathrm{e}^u + 1]\mathrm{e}^{ux}}{(\mathrm{e}^u-1)^2}\,du > 0.$$

The result now follows. □

As stated in [35, p. 634], the volume function $V_n(p)$ is neither convex nor concave for $n \geq 3$.

In the next chapter we shall encounter the class of M_p-convex functions $(-\infty \leq p \leq \infty)$. A function $f\colon I \to \mathbb{R}$ is said to be M_p-*convex* if

$$f((1-\lambda)x + \lambda y) \leq M_p(f(x), f(y)\,; 1-\lambda, \lambda)$$

for all $x, y \in I$ and all $\lambda \in [0,1]$ (that is, f is (A, M_p)-convex). In order to avoid trivial situations, the theory of M_p-convex functions is usually restricted to nonnegative functions when $p \in \mathbb{R}$, $p \neq 1$.

The case $p = 1$ corresponds to the usual convex functions, while for $p = 0$ we retrieve the log-convex functions. The case $p = \infty$ is that of *quasiconvex* functions, that is, of functions $f\colon I \to \mathbb{R}$ such that

$$f((1-\lambda)x + \lambda y) \leq \sup\{f(x), f(y)\}$$

for all $x, y \in I$ and all $\lambda \in [0,1]$. Clearly, a function $f\colon I \to \mathbb{R}$ is quasiconvex if and only if its sublevel sets $\{x \mid f(x) \leq \alpha\}$ are convex for all $\alpha \in \mathbb{R}$.

If $p > 0$ (or $p < 0$), a function f is M_p-convex if and only if f^p is convex (or concave, respectively). According to Exercise 8, Section 1.1,

$$M_p(x, y\,; 1-\lambda, \lambda) \leq M_q(x, y\,; 1-\lambda, \lambda) \quad \text{for } -\infty \leq p \leq q \leq \infty,$$

which shows that every M_p-convex function is also M_q-convex for all $q \geq p$.

Exercises

1. Suppose that I and J are nondegenerate intervals and $p, q, r \in \mathbb{R}$, $p < q$. Prove that for every function $f\colon I \to J$ the following two implications hold true:

- If f is (M_q, M_r)-convex and increasing, then it is also (M_p, M_r)-convex;
- If f is (M_p, M_r)-convex and decreasing, then it is also (M_q, M_r)-convex.

Conclude that the function $V_\alpha(p)$ of Theorem 2.6.2 is also (A, G)-concave and (H, A)-concave.

2. Suppose that M and N are two regular means (respectively on the intervals I and J) and the function $N(\cdot, 1)$ is concave. Prove that:

(i) for every two (M, N)-convex functions $f, g\colon I \to J$, the function $f+g$ is (M, N)-convex;

(ii) for every (M, N)-convex function $f\colon I \to J$ and $\alpha > 0$, the function αf is (M, N)-convex.

3. Suppose that $f\colon I \to \mathbb{R}$ is a continuous function which is differentiable on $\operatorname{int} I$. Prove that f is quasiconvex if and only if for each $x, y \in \operatorname{int} I$,

$$f(y) \leq f(x) \text{ implies } f'(x)(y - x) \leq 0.$$

4. (K. Knopp and B. Jessen; see [99, p. 66]) Suppose that φ and ψ are two continuous functions defined in an interval I such that φ is strictly monotonic and ψ is increasing.

(i) Prove that

$$M_{[\varphi]}(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n) = M_{[\psi]}(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_n)$$

for every family $x_1, \ldots, x_n$ of elements of I and every family $\lambda_1, \ldots, \lambda_n$ of nonnegative numbers with $\sum_{k=1}^n \lambda_k = 1$ ($n \in \mathbb{N}^\star$) if and only if $\psi \circ \varphi^{-1}$ is affine, that is, $\psi = \alpha\varphi + \beta$ for some constants α and β, with $\alpha \neq 0$.

(ii) Infer that any power mean M_p is a mean $M_{[\varphi]}$, where $\varphi(x) = \log x$, if $p = 0$, and $\varphi(x) = (x^p - 1)/p$, if $p \neq 0$.

5. (M. Nagumo, B. de Finetti and B. Jessen; see [99, p. 68]) Let φ be a continuous increasing function on $(0, \infty)$ such that the quasi-arithmetic mean $M_{[\varphi]}$ is positively homogeneous. Prove that $M_{[\varphi]}$ is one of the power means.
[*Hint*: By Exercise 4 (i), we can replace φ by $\varphi - \varphi(1)$, so we may assume that $\varphi(1) = 0$. The same argument yields two functions α and β such that $\varphi(cx) = \alpha(c)\varphi(x) + \beta(c)$ for all $x > 0$, $c > 0$. The condition $\varphi(1) = 0$ shows that $\beta = \varphi$, so for reasons of symmetry,

$$\varphi(cx) = \alpha(c)\varphi(x) + \varphi(c) = \alpha(x)\varphi(c) + \varphi(x).$$

Letting fixed $c \neq 1$, we obtain that α is of the form $\alpha(x) = 1 + k\varphi(x)$ for some constant k. Then φ verifies the functional equation

$$\varphi(xy) = k\varphi(x)\varphi(y) + \varphi(x) + \varphi(y)$$

for all $x > 0$, $y > 0$. When $k = 0$ we find that $\varphi(x) = C \log x$ for some constant C, so $M_{[\varphi]} = M_0$. When $k \neq 0$ we notice that $\chi = k\varphi + 1$ verifies $\chi(xy) = \chi(x)\chi(y)$ for all $x > 0$, $y > 0$. This leads to $\varphi(x) = (x^p - 1)/k$, for some $p \neq 0$, hence $M_{[\varphi]} = M_p$.]

6. (Convexity with respect to Stolarsky's means) One can prove that the exponential function is (L, L)-convex. See Exercise 5 (iii), Section 2.3. Prove that this function is also (I, I)-convex. What can be said about the logarithmic function? Here L and I are respectively the logarithmic mean and the identric mean.

7. (Few affine functions with respect to the logarithmic mean; see [157]) Prove that the only (L, L)-affine functions $f\colon (0, \infty) \to (0, \infty)$ are the constant functions and the linear functions $f(x) = cx$, for $c > 0$. Infer that the logarithmic mean is not a power mean.

2.7 Relative Convexity

The comparison of quasi-arithmetic means is related to convexity via the following result:

Lemma 2.7.1 *Suppose that $\varphi, \psi\colon I \to \mathbb{R}$ are two strictly monotonic continuous functions. If φ is increasing, then*

$$M_{[\psi]} \leq M_{[\varphi]}$$

if and only if $\varphi \circ \psi^{-1}$ is convex.

Lemma 2.7.1 has important consequences. For example, it yields Clarkson's inequalities (which in turn extend the parallelogram law). The following approach (in the spirit of Orlicz spaces) is due to J. Lamperti [136]:

Theorem 2.7.2 *Suppose that $\Phi\colon [0, \infty) \to \mathbb{R}$, is an increasing and continuous function with $\Phi(0) = 0$ and $\Phi(\sqrt{x})$ convex. Consider a σ-finite measure space (X, Σ, μ) and denote by $L^{\Phi}(X)$ the set of all equivalence classes of all μ-measurable real-valued functions f such that*

$$I_{\Phi}(f) = \int_X \Phi(|f(x)|)\, d\mu < \infty.$$

If $f + g$ and $f - g$ belong to $L^{\Phi}(X)$, then

$$I_{\Phi}(f + g) + I_{\Phi}(f - g) \geq 2I_{\Phi}(f) + 2I_{\Phi}(g). \tag{2.1}$$

If $\Phi(\sqrt{x})$ *is concave and* f *and* g *belong to* $L^{\Phi}(X)$*, then the reverse inequality is true. If the convexity or concavity of* $\Phi(\sqrt{x})$ *is strict, equality holds in* (2.1) *if and only if* $fg = 0$ *almost everywhere.*

Corollary 2.7.3 (Clarkson's inequalities [57]) *If* $2 \le p < \infty$*, and* f *and* g *belong to* $L^p(\mu)$*, then*

$$\|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p \ge 2\|f\|_{L^p}^p + 2\|g\|_{L^p}^p.$$

If $0 < p \le 2$*, then the reverse inequality holds. In either case, if* $p \ne 2$*, equality occurs if and only if* $fg = 0$ *almost everywhere.*

Clarkson's inequalities easily imply the uniform convexity of the spaces $L^p(\mu)$ for $1 < p < \infty$ (see Exercise 2). J. Lamperti applied Corollary 2.7.3 to give the general form of the linear isometries $T\colon L^p(\mu) \to L^p(\mu)$, for $p > 0$, $p \ne 2$.

Clarkson's inequalities are improved on by Hanner's inequalities. See Exercise 7, Section 3.6.

Proof of Theorem 2.7.2. It suffices to prove the following result: *Suppose that* $\Phi\colon [0,\infty) \to \mathbb{R}$ *is a continuous increasing function with* $\Phi(0) = 0$ *and* $\Phi(\sqrt{t})$ *convex. Then*

$$\Phi(|z+w|) + \Phi(|z-w|) \ge 2\Phi(|z|) + 2\Phi(|w|), \tag{2.2}$$

for all $z, w \in \mathbb{C}$*, while if* $\Phi(\sqrt{t})$ *is concave the reverse inequality is true.*

Provided the convexity or concavity is strict, equality holds if and only if $zw = 0$.

In fact, since $\Phi(\sqrt{t})$ is convex, we infer from Lemma 2.7.1 and the parallelogram law the inequality

$$\begin{aligned}\Phi^{-1}\Big\{\frac{\Phi(|z+w|) + \Phi(|z-w|)}{2}\Big\} &\ge \Big\{\frac{|z+w|^2 + |z-w|^2}{2}\Big\}^{1/2} \\ &= (|z|^2 + |w|^2)^{1/2}.\end{aligned} \tag{2.3}$$

On the other hand, the convexity of $\Phi(\sqrt{t})$ and the fact that $\Phi(0) = 0$ yield that $\Phi(\sqrt{t})/t$ is nondecreasing, that is, $t^2/\Phi(t)$ is nonincreasing (respectively decreasing if the convexity is strict). See Theorem 1.3.1. Taking into account the result of Exercise 1, we infer

$$\Phi^{-1}\{\Phi(|z|) + \Phi(|w|)\} \le (|z|^2 + |w|^2)^{1/2}, \tag{2.4}$$

and thus (2.2) follows from (2.3), (2.4) and the fact that Φ is increasing. When $\Phi(\sqrt{t})$ is strictly convex, we also obtain from Exercise 1 the fact that (2.4) (and thus (2.2)) is strict unless z or w is zero. □

Lemma 2.7.1 leads us naturally to consider the following concept of relative convexity:

Definition 2.7.4 Suppose that f and g are two real-valued functions defined on the same set X, and g is not a constant function. Then f is said to be *convex relative to* g (abbreviated, $g \lhd f$) if

$$\begin{vmatrix} 1 & g(x) & f(x) \\ 1 & g(y) & f(y) \\ 1 & g(z) & f(z) \end{vmatrix} \geq 0,$$

whenever $x, y, z \in X$ with $g(x) \leq g(y) \leq g(z)$.

When X is an interval and g is continuous and increasing, a small computation shows that the condition $g \lhd f$ is equivalent with the convexity of $f \circ g^{-1}$ (on the interval $J = g(I)$).

Examples 2.7.5
Under appropriate assumptions on the domain and the range of the function f, the following statements hold true:

(i) f is convex if and only if $\mathrm{id} \lhd f$;
(ii) f is log-convex if and only if $\mathrm{id} \lhd \log f$;
(iii) f is (G, G)-convex if and only if $\log \lhd \log f$;
(iv) f is (G, A)-convex if and only if $\log \lhd f$.

A more exotic illustration of the concept of relative convexity is the following fact:

$$f \lhd f^{\alpha} \quad \text{for all } f\colon X \to \mathbb{R}_{+} \text{ and all } \alpha \geq 1.$$

For example, $\sin \lhd \sin^2$ on $[0, \pi]$, and $|x| \lhd x^2$ on $\mathbb{R}$.

In the context of C^1-differentiable functions, f is convex with respect to an increasing function g if f'/g' is nondecreasing; in the context of C^2-differentiable functions, f is convex with respect to g if and only if $f''/f' \geq g''/g'$ (provided these ratios exist).

It is important to notice that relative convexity is part of comparative convexity. For this we need the integral analogue of quasi-arithmetic mean,

$$M_{[\varphi]}\Big(\mathrm{id}_{[s,t]}; \frac{1}{t-s}\, dx\Big) = \varphi^{-1}\bigg(\frac{1}{t-s}\int_s^t \varphi(x)\, dx\bigg).$$

In fact, if $g \lhd f$, then

$$\begin{aligned} f\Big(M_{[g]}\Big(\mathrm{id}_{[a,b]}; \frac{1}{b-a}\, dx\Big)\Big) &= f\bigg(g^{-1}\bigg(\frac{1}{b-a}\int_a^b g(x)\, dx\bigg)\bigg) \\ &\leq \frac{1}{b-a}\int_a^b f(x)\, dx = M_1(f|_{[a,b]}) \end{aligned}$$

for all $a < b$ in the domain of f and g.

From the above discussion we can infer the following remark due to H. Alzer [7]: Suppose that f is an increasing continuous function (acting on subintervals of $(0, \infty)$) and $1/f^{-1}$ is convex. Then $1/x \lhd f$. As

$M_{[1/x]}(\mathrm{id}_{[a,b]}\,;\frac{1}{b-a}dx)$ coincides with the logarithmic mean $L(a,b)$, it follows that

$$f(L(a,b)) \leq \frac{1}{b-a}\int_a^b f(x)\,dx = M_1(f|_{[a,b]}).$$

We end this section by extending the Hardy–Littlewood–Pólya inequality to the context of relative convexity. Our approach is based on two technical lemmas.

Lemma 2.7.6 *If $f, g\colon X \to \mathbb{R}$ are two functions such that $g \lhd f$, then*

$$g(x) = g(y) \text{ implies } f(x) = f(y).$$

Proof. Since g is not constant, then there must be a $z \in X$ such that $g(x) = g(y) \neq g(z)$. One of the following two cases may occur:

Case 1: $g(x) = g(y) < g(z)$. This yields

$$0 \leq \begin{vmatrix} 1 & g(x) & f(x) \\ 1 & g(x) & f(y) \\ 1 & g(z) & f(z) \end{vmatrix} = (g(z) - g(x))(f(x) - f(y))$$

and thus $f(x) \geq f(y)$. A similar argument gives us the reverse inequality, $f(x) \leq f(y)$.

Case 2: $g(z) < g(x) = g(y)$. This case can be treated in a similar way. □

Lemma 2.7.7 (The generalization of Galvani's Lemma) *If $g \lhd f$ and x, u, v are points of X such that $g(x) \notin \{g(u), g(v)\}$ and $g(u) \leq g(v)$, then*

$$\frac{f(v) - f(x)}{g(v) - g(x)} \geq \frac{f(u) - f(x)}{g(u) - g(x)}.$$

Proof. In fact, the following three cases may occur:

Case 1: $g(x) < g(u) \leq g(v)$. Then

$$\begin{aligned} 0 &\leq \begin{vmatrix} 1 & g(x) & f(x) \\ 1 & g(u) & f(u) \\ 1 & g(v) & f(v) \end{vmatrix} \\ &= (g(u) - g(x))(f(v) - f(x)) - (g(v) - g(x))(f(u) - f(x)) \end{aligned}$$

and the conclusion of Lemma 2.7.7 is clear.

Case 2: $g(u) \leq g(v) < g(x)$. This case can be treated in the same way.

Case 3: $g(u) < g(x) < g(v)$. According to the discussion above we have

$$\begin{aligned} \frac{f(u) - f(x)}{g(u) - g(x)} &= \frac{f(x) - f(u)}{g(x) - g(u)} \leq \frac{f(v) - f(u)}{g(v) - g(u)} \\ &= \frac{f(u) - f(v)}{g(u) - g(v)} \leq \frac{f(x) - f(v)}{g(x) - g(v)} = \frac{f(v) - f(x)}{g(v) - g(x)} \end{aligned}$$

and the proof is now complete. □

Theorem 2.7.8 (The generalization of the Hardy–Littlewood–Pólya inequality) *Let $f, g\colon X \to \mathbb{R}$ be two functions such that $g \lhd f$ and consider points $x_1, \ldots, x_n, y_1, \ldots, y_n \in X$ and weights $p_1, \ldots, p_n \in \mathbb{R}$ such that:*

(i) $g(x_1) \geq \cdots \geq g(x_n)$ *and* $g(y_1) \geq \cdots \geq g(y_n)$;

(ii) $\sum_{k=1}^{r} p_k g(x_k) \leq \sum_{k=1}^{r} p_k g(y_k)$ *for every* $r = 1, \ldots, n$;

(iii) $\sum_{k=1}^{n} p_k g(x_k) = \sum_{k=1}^{n} p_k g(y_k)$.

Then

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

Proof. By mathematical induction. The case $n = 1$ is clear. Assuming the conclusion of Theorem 2.7.8 valid for all families of length $n - 1$, let us pass to the families of length n. The case where $g(x_k) = g(y_k)$ for some index k can be settled easily by our hypothesis and Lemma 2.7.6. Therefore we may restrict ourselves to the case where $g(x_k) \neq g(y_k)$ for all indices k. By Abel's summation formula,

$$\sum_{k=1}^{n} p_k f(y_k) - \sum_{k=1}^{n} p_k f(x_k) \tag{2.5}$$

equals

$$\frac{f(y_n) - f(x_n)}{g(y_n) - g(x_n)}\Big(\sum_{i=1}^{n} p_i g(y_i) - \sum_{i=1}^{n} p_i g(x_i)\Big)$$
$$+ \sum_{k=1}^{n-1}\Big(\frac{f(y_k) - f(x_k)}{g(y_k) - g(x_k)} - \frac{f(y_{k+1}) - f(x_{k+1})}{g(y_{k+1}) - g(x_{k+1})}\Big)\Big(\sum_{i=1}^{k} p_i g(y_i) - \sum_{i=1}^{k} p_i g(x_i)\Big)$$

which, by (iii), reduces to

$$\sum_{k=1}^{n-1}\Big(\frac{f(y_k) - f(x_k)}{g(y_k) - g(x_k)} - \frac{f(y_{k+1}) - f(x_{k+1})}{g(y_{k+1}) - g(x_{k+1})}\Big)\Big(\sum_{i=1}^{k} p_i g(y_i) - \sum_{i=1}^{k} p_i g(x_i)\Big).$$

According to (ii), the proof will be complete if we show that

$$\frac{f(y_{k+1}) - f(x_{k+1})}{g(y_{k+1}) - g(x_{k+1})} \leq \frac{f(y_k) - f(x_k)}{g(y_k) - g(x_k)}$$

for all indices k.

In fact, if $g(x_k) = g(x_{k+1})$ or $g(y_k) = g(y_{k+1})$ for some index k, this follows from (i) and Lemmas 2.7.6 and 2.7.7.

When $g(x_k) > g(x_{k+1})$ and $g(y_k) > g(y_{k+1})$ the following two cases may occur:

Case 1: $g(x_k) \neq g(y_{k+1})$. By a twice application of Lemma 2.7.7 we get

$$\begin{aligned}\frac{f(y_{k+1})-f(x_{k+1})}{g(y_{k+1})-g(x_{k+1})} &= \frac{f(x_{k+1})-f(y_{k+1})}{g(x_{k+1})-g(y_{k+1})} \leq \frac{f(x_k)-f(y_{k+1})}{g(x_k)-g(y_{k+1})} \\ &= \frac{f(y_{k+1})-f(x_k)}{g(y_{k+1})-g(x_k)} \leq \frac{f(y_k)-f(x_k)}{g(y_k)-g(x_k)}.\end{aligned}$$

Case 2: $g(x_k) = g(y_{k+1})$. In this case, $g(x_{k+1}) < g(x_k) = g(y_{k+1}) < g(y_k)$, and Lemmas 2.7.6 and 2.7.7 lead us to

$$\begin{aligned}\frac{f(y_{k+1})-f(x_{k+1})}{g(y_{k+1})-g(x_{k+1})} &= \frac{f(x_k)-f(x_{k+1})}{g(x_k)-g(x_{k+1})} \\ &= \frac{f(x_{k+1})-f(x_k)}{g(x_{k+1})-g(x_k)} \leq \frac{f(y_k)-f(x_k)}{g(y_k)-g(x_k)}.\end{aligned}$$

Consequently, (2.5) is a sum of nonnegative terms, and the proof is complete. □

The classical Hardy–Littlewood–Pólya inequality corresponds to the case where X is an interval, g is the identity, and $p_k = 1$ for all k. In this case, the hypothesis (i) can be replaced by the following one:

(i′) $g(x_1) \geq \cdots \geq g(x_n)$,

see Theorem 1.5.4. When X is an interval, g is the identity, and $p_1, \ldots, p_n$ are arbitrary weights, then the result of Theorem 2.7.8 is known as *Fuchs' inequality* [83]. Clearly, Fuchs' inequality implies Corollary 1.4.3 above.

In a similar way, we can extend another important result in majorization theory, the Tomić–Weyl theorem. See Exercise 5.

Exercises

1. (R. Cooper; see [99, p. 84]) Suppose that $\varphi, \psi \colon I \to (0,\infty)$ are two continuous bijective functions. If φ and ψ vary in the same direction and φ/ψ is nonincreasing, then

$$\psi^{-1}\Big(\sum_{k=1}^{n} \psi(x_k)\Big) \leq \varphi^{-1}\Big(\sum_{k=1}^{n} \varphi(x_k)\Big)$$

for every finite family $x_1, \ldots, x_n$ of elements of I.
[*Hint*: If $h(x)/x$ is nonincreasing for $x > 0$, then $h(\sum_{k=1}^{n} x_k) \leq \sum_{k=1}^{n} h(x_k)$ for every finite family $x_1, \ldots, x_n$ of positive numbers. See Section 1.3, Exercise 8.]

2. Infer from Clarkson's inequalities the *uniform convexity* of the spaces $L^p(\mu)$, for $1 < p < \infty$, that is, if x and y are in the unit ball of $L^p(\mu)$, then

$$\inf\Big\{1 - \Big\|\frac{x+y}{2}\Big\| \;\Big|\; \|x-y\| \geq \varepsilon\Big\} > 0 \quad \text{for all } \varepsilon \in (0,2].$$

3. Suppose that $F, g\colon I \to J$ are two continuous functions and g is strictly monotone. Prove that $g \lhd F$ if and only if for every $\alpha \geq 0$ and every $[a, b] \subset I$ the function $F - \alpha g$ attains its maximum either at a or at b.
Remark. This result can be used to prove sharpened versions of the maximum principle for elliptic partial differential operators. See [242].

4. Suppose that $f\colon [0, \pi/2] \to \mathbb{R}$ is a function such that

$$(f(y) - f(z)) \cos x + (f(z) - f(x)) \cos y + (f(x) - f(y)) \cos z \geq 0$$

for all $x \geq y \geq z$ in $[0, \pi/2]$. Prove that

$$f\Big(\frac{\pi}{7}\Big) - f\Big(\frac{2\pi}{7}\Big) + f\Big(\frac{3\pi}{7}\Big) \leq f(0) - f\Big(\frac{\pi}{3}\Big) + f\Big(\frac{\pi}{2}\Big).$$

5. (An extension of the Tomić–Weyl theorem) Suppose that $f, g\colon X \to \mathbb{R}$ are two synchronous functions with $g \lhd f$. Consider points $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ in X and real weights $p_1, \ldots, p_n$ such that:
(i) $g(x_1) \geq \cdots \geq g(x_n)$ and $g(y_1) \geq \cdots \geq g(y_n)$;
(ii) $\sum_{k=1}^{m} p_k g(x_k) \leq \sum_{k=1}^{m} p_k g(y_k)$ for all $m = 1, \ldots, n$.
Prove that

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

2.8 Comments

The idea of transforming a nonconvex function into a convex one by a change of variable has a long history. As far as we know, the class of all multiplicatively convex functions was first considered by P. Montel [171] in a beautiful paper discussing the possible analogues of convex functions in n variables. He motivates his study with the following two classical results:

Hadamard's Three Circles Theorem *Let f be an analytical function in the annulus $a < |z| < b$. Then $\log M(r)$ is a convex function of $\log r$, where*

$$M(r) = \sup_{|z|=r} |f(z)|.$$

G. H. Hardy's Mean Value Theorem *Let f be an analytical function in the annulus $a < |z| < b$ and let $p \in [1, \infty)$. Then $\log M_p(r)$ is a convex function of $\log r$, where*

$$M_p(r) = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^p \, d\theta\right)^{1/p}.$$

As $\lim_{p\to\infty} M_p(r) = M(r)$, Hardy's aforementioned result implies Hadamard's. It is well known that Hadamard's result is instrumental in deriving the Riesz–Thorin interpolation theorem (see [99]).

The presentation of the class of multiplicatively convex functions (as was done in Sections 2.3 and 2.4) follows C. P. Niculescu [176]. The multiplicative mean (see [178] and Section 2.3, Exercises 5 and 6) provides the right analogue of the arithmetic mean in a fully multiplicative theory of convexity.

The theory of Euler's functions gamma and beta follows the same steps as in E. Artin [10] and R. Webster [243].

As noted by T. Trif [238], the result of Proposition 2.4.2 can be improved: the gamma function is strictly multiplicatively concave on $(0,\alpha]$ and strictly multiplicatively convex on $[\alpha,\infty)$, where $\alpha \approx 0.21609$ is the unique positive solution of the equation $\operatorname{Psi}(x) + x\frac{d}{dx}\operatorname{Psi}(x) = 0$. This fact has a full generalization in the context of (M_p, M_p)-convexity.

The quantum analogue of the gamma function, the *q-gamma function* Γ_q of F. H. Jackson, is defined by

$$\Gamma_q(x) = \frac{(q\,;q)_\infty}{(q^x\,;q)_\infty}\,(1-q)^{1-x} \quad \text{for } x>0 \ \ (0<q<1),$$

where $(a\,;q)_\infty = \prod_{k=0}^{\infty}(1-aq^k)$. For it, the Bohr–Mollerup theorem has the following form: Γ_q *is the only solution of the functional equation*

$$\Gamma_q(x+1) = \frac{1-q^x}{1-q}\,\Gamma_q(x)$$

which is log-*convex and satisfies* $\Gamma_q(1) = 1$ (see [9]). Γ_q is multiplicatively convex at least on $(2,\infty)$ (see D. Gronau and J. Matkowski [91]).

The well-known inequalities in a triangle ΔABC, such as

$$\sin A + \sin B + \sin C \le 3\sqrt{3}/2 \quad \text{and} \quad \sin A \sin B \sin C \le 3\sqrt{3}/8,$$

can be traced back to an old paper by G. Berkhan [21], from 1907.

R. A. Satnoianu [221] observed that the functions which are convex, multiplicatively convex and increasing are the source of Erdős–Mordell type inequalities in a triangle. Examples of such functions are numerous. See Proposition 2.3.3.

The estimate given in Theorem 2.5.3 for the AM–GM inequality was mentioned in [176].

The general notion of mean was clarified by B. de Finetti [80].

The idea to consider the general notion of (M,N)-convex function (associated to a pair of means) can be traced back to G. Aumann [13]. Important contributions came from J. Aczél [2], [3], J. Matkowski [157], J. Matkowski and J. Rätz [158], [159]. The canonical extension of a mean, as well as Theorems A, B and C in the Introduction, are due to C. P. Niculescu [183].

The result of Exercise 5, Section 2.6, concerning the characterization of the power means among the quasi-arithmetic means, was recently extended by J. Matkowski [157] to the context of strict and homogeneous means which verify some nondegeneracy conditions.

The comparability Lemma 2.7.1 is due to B. Jessen (see [99, p. 75]). The concept of relative convexity can be also traced back to Jessen (see [99, Theorem 92, p. 75]). Later, it was developed by G. T. Cargo [47], N. Elezović and J. Pečarić [71] and many others. The generalization of the classical inequalities of Hardy–Littlewood–Pólya, Fuchs and Tomić–Weyl to the framework of relative convexity follows closely the paper [189] by C. P. Niculescu and F. Popovici.

Recently, M. Bessenyei and Z. Páles [26] have considered a more general concept of relative convexity, which goes back to a result of G. Pólya; see [99, Theorem 123, p. 98]. Given a pair (ω_1, ω_2) of continuous functions on an interval I, such that

$$\begin{vmatrix} \omega_1(x) & \omega_1(y) \\ \omega_2(x) & \omega_2(y) \end{vmatrix} \neq 0 \quad \text{for all } x < y, \tag{2.6}$$

a function $f\colon I \to \mathbb{R}$ is said to be (ω_1, ω_2)-*convex* if

$$\begin{vmatrix} f(x) & f(y) & f(z) \\ \omega_1(x) & \omega_1(y) & \omega_1(z) \\ \omega_2(x) & \omega_2(y) & \omega_2(z) \end{vmatrix} \geq 0$$

for all $x < y < z$ in I. It is proved that the (ω_1, ω_2)-convexity implies the continuity of f at the interior points of I, as well as the integrability on compact subintervals of I.

If I is an open interval, $\omega_1 > 0$ and the determinant in formula (2.6) is positive, then f is (ω_1, ω_2)-convex if and only if the function $f/\omega_1 \circ (\omega_2/\omega_1)^{-1}$ is convex in the usual sense. Under these restrictions, M. Bessenyei and Z. Páles proved a Hermite–Hadamard type inequality. Note that this case of (ω_1, ω_2)-convexity falls under the incidence of relative convexity.

There is much information available nowadays concerning the Clarkson type inequalities, and several applications have been described. Here we just mention that even the general Edmunds–Triebel logarithmic spaces satisfy Clarkson's inequalities: see [191], where some applications and relations to several previous results and references are also presented.

A classical result due to P. Jordan and J. von Neumann asserts that the parallelogram law characterizes Hilbert spaces among Banach spaces. See M. M. Day [64, pp. 151–153]. There are two important generalizations of the parallelogram law (both simple consequences of the inner-product structure).

The Leibniz–Lagrange identity. *Suppose there is given a system of weighted points* $(x_1, m_1), \ldots, (x_r, m_r)$ *in an inner-product space* H*, whose barycenter position is*

$$x_G = \sum_{k=1}^{r} m_k x_k \Big/ \sum_{k=1}^{r} m_k.$$

Then for all points $x \in H$ we have the equalities

$$\sum_{k=1}^{r} m_k \|x - x_k\|^2 = \Big(\sum_{k=1}^{r} m_k\Big) \|x - x_G\|^2 + \sum_{k=1}^{r} m_k \|x_G - x_k\|^2$$
$$= \Big(\sum_{k=1}^{r} m_k\Big) \|x - x_G\|^2 + \frac{1}{\sum_{k=1}^{r} m_k} \cdot \sum_{i<j} m_i m_j \|x_i - x_j\|^2.$$

This identity is at the origin of many well-known formulas concerning the distances between some special points in a triangle. For example, in the case where x_1, x_2, x_3 are the vertices of a triangle and m_1, m_2, m_3 are proportional to the length sides a, b, c, then x_G is precisely the center I of the inscribed circle. The above identity gives us (for $x = O$, the center of the circumscribed circle) the celebrated *formula of Euler*,

$$OI^2 = R(R - 2r).$$

More information can be found at `www.neiu.edu/~mathclub/Seminar Notes/Some Mathematical Consequences of the Law of the Lever`.

E. Hlawka's identity. *We have*

$$\|x\|^2 + \|y\|^2 + \|z\|^2 + \|x + y + z\|^2 = \|x + y\|^2 + \|y + z\|^2 + \|z + x\|^2,$$

for all x, y, z in an inner-product space H.

This yields *Hlawka's inequality*: In any inner-product space H, for all $x, y, z \in H$ we have

$$\|x + y + z\| + \|x\| + \|y\| + \|z\| - \|x + y\| - \|y + z\| - \|z + x\| \geq 0.$$

In fact, based on Hlawka's identity, the left-hand side equals

$$(\|x\| + \|y\| - \|x + y\|)\Big(1 - \frac{\|x\| + \|y\| + \|x + y\|}{\|x\| + \|y\| + \|z\| + \|x + y + z\|}\Big)$$
$$+ (\|y\| + \|z\| - \|y + z\|)\Big(1 - \frac{\|y\| + \|z\| + \|y + z\|}{\|x\| + \|y\| + \|z\| + \|x + y + z\|}\Big)$$
$$+ (\|z\| + \|x\| - \|z + x\|)\Big(1 - \frac{\|z\| + \|x\| + \|z + x\|}{\|x\| + \|y\| + \|z\| + \|x + y + z\|}\Big)$$

which is a combination of nonnegative terms.

Hlawka's inequality is not characteristic to Euclidean spaces! In fact, it was extended by J. Lindenstrauss and A. Pełczyński [146] to all Banach spaces E whose finite dimensional subspaces can be embedded (linearly and isometrically) in suitable spaces $L^p([0, 1])$, with $1 \leq p \leq 2$. On the other hand, Hlawka's inequality does not work for all Banach spaces. A counterexample is provided by $\mathbb{C}^2$, endowed with the sup norm, and the vectors $x = (1, -1)$, $y = (i, i)$, $z = (-i, 1)$.

A large generalization of Hlawka's inequality, based on ergodic theory, was given by M. Rădulescu and S. Rădulescu [210].