Modern Birkhäuser Classics

BIRKHÄUSER

Notions of Convexity

Lars Hörmander

Notions of Convexity

Lars Hörmander

Reprint of the 1994 Edition

Birkhäuser
Boston • Basel • Berlin

Lars Hörmander
Lund University
Center for Mathematical Sciences
SE-22100 Lund
Sweden

Originally published as Volume 127 in the series *Progress in Mathematics*

Cover design by Alex Gerasev.

Mathematics Subject Classification (2000): 00A05, 01A60, 03-02, 26A51, 26B25, 31-02, 31B05, 31C10, 32-02, 32F05, 32F15, 32T99, 32U05, 32W05, 35A27, 52A40

Library of Congress Control Number: 2006937427

ISBN-10: 0-8176-4584-5 e-ISBN-10: 0-8176-4585-3
ISBN-13: 978-0-8176-4584-7 e-ISBN-13: 978-0-8176-4585-4

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com (IBT)

Lars Hörmander
Department of Mathematics
University of Lund
Box 118, S-221 00 Lund
Sweden

Library of Congress Cataloging In-Publication Data

Hörmander, Lars, 1931-
Notions of convexity / Lars Hörmander.
p. cm. -- (Progress in mathematics ; v. 127)
Includes bibliographical references and indexes.
ISBN 0-8176-3799-0 (acid free).
1. Convex domains. I. Title. II. Progress in mathematics (Boston, Mass.) ; vol. 127
QA639.5.H67 1994 94-32572
515.'94--dc20 CIP

Printed on acid-free paper

Birkhäuser ®

ISBN 0-8176-3799-0
ISBN 3-7643-3799-0
Typeset by the author in $\mathcal{AMS}$-TEX.
Printed and bound by Quinn-Woodbine, Woodbine, NJ.
Printed in the U.S.A.

9 8 7 6 5 4 3 2 1

Modern Birkhäuser Classics

Many of the original research and survey monographs in pure and applied mathematics published by Birkhäuser in recent decades have been groundbreaking and have come to be regarded as foundational to the subject. Through the MBC Series, a select number of these modern classics, entirely uncorrected, are being re-released in paperback (and as eBooks) to ensure that these treasures remain accessible to new generations of students, scholars, and researchers.

Lars Hörmander

Notions of Convexity

Birkhäuser
Boston • Basel • Berlin

PREFACE

The term convexity used to describe these lectures given at the University of Lund in 1991–92 should be understood in a wide sense. Only Chapters I and II are devoted to convex sets and functions in the traditional sense of convexity. The following chapters study other kinds of convexity which occur in analysis. Most prominent is the pseudo-convexity (plurisubharmonicity) in the theory of functions of several complex variables discussed in Chapter IV. It relies on the theory of subharmonic functions in $\mathbf{R}^2$, so Chapter III is devoted to subharmonic functions in $\mathbf{R}^n$ for any n. Existence theorems for constant coefficient partial differential operators in $\mathbf{R}^n$ are related to various kinds of convexity conditions, depending on the operator. Chapter VI gives a survey of the rather incomplete results which are known on their geometrical meaning. There are also natural classes of "convex" functions related to subgroups of the linear group, which specialize to several of the notions already mentioned. They are discussed in Chapter V. The last chapter, Chapter VII, is devoted to the conditions for solvability of microdifferential equations, which can also be considered as a branch of convexity theory. The whole chapter is an exposition of a part of the thesis of J.-M. Trépreau.

Thus the main purpose is to discuss notions of convexity — for functions and for sets — which occur in the theory of partial differential equations and complex analysis. However, it is impossible to resist the temptation to present a number of beautiful related topics, such as basic inequalities in analysis and isoperimetric inequalities. In fact, this gives an opportunity to show how conversely the theory of partial differential equations contributes to convexity theory. Originally I also planned to discuss the role of convexity in linear and non-linear functional analysis, but that turned out to be impossible in the time available. Another topic which is conspicuously missing is the theory of the real and the complex Monge-Ampère equations, which should have been presented in Chapters II and IV.

Convexity theory has contacts with many areas of mathematics. However, only applications in complex analysis and the theory of linear partial differential equations are discussed here, without aiming for completeness. I hope that in spite of that the book will prove useful for readers with main interest in other directions, and that it does justice to the beauty of the subject.

To minimize the number of references relied on I have often referred to my books denoted by ALPDO and CASV (see the bibliography at the end) instead of original works. Further references can be found in these books. At the beginning of the notes no prerequisites are assumed beyond calculus and linear algebra. Measure and integration theory are required in Section 1.7 and from Chapter III on. Distribution theory has been used systematically from Chapter III when it simplifies or clarifies the presentation, even where it could be avoided. However, only the most elementary part of the first seven chapters in ALPDO are required. Some background in differential geometry is assumed in Section 2.3, and the proof of the Fenchel-Alexandrov inequality there requires some knowledge of elliptic differential operators. At the end basic Riemannian geometry is also required, and Section 6.2 assumes familiarity with pseudodifferential operators. The last section, Section 7.4, assumes some background in analytic microlocal analysis, and some knowledge of symplectic geometry is needed in Section 7.3. Only the simplest facts from functional analysis are needed except in Section 6.3 where deeper results on duality theory are used. However, these are exceptions which can be bypassed with no loss of continuity. Apart from these points the notes should be accessible to any graduate student with an interest in analysis.

As already mentioned Chapter VII is based on J.-M. Trépreau's thesis. The presentation here owes much to the patience with which he has corrected and improved earlier versions; any remaining mistakes are of course my own. I wish to thank him for all this help and for informing me about improvements that he made in a recent unpublished manuscript. In the final version they have been partially replaced by still more recent unpublished results due to A. Ancona presented in Section 1.7 and at the end of Sections 3.2 and 4.1. I am grateful for his permission to include them here.

I would also like to thank Anders Melin for his critical reading of a large part of the manuscript, and M. Andersson, M. Passare and R. Sigurdsson who agreed to the inclusion in Chapter IV of some material from an unpublished manuscript of theirs. Thanks are also due to the publishers and their referees.

Lund in June 1994

Lars Hörmander

CONTENTS

CHAPTER I

CONVEX FUNCTIONS OF ONE VARIABLE

Summary. Section 1.1 just recalls well-known elementary facts which are essential for all the following chapters. Section 1.2 is devoted to proofs of basic inequalities in analysis by convexity arguments. The Legendre transform (conjugate convex functions) is discussed in Section 1.3 in a spirit which prepares for the case of several variables in Chapter II. Section 1.4 is an interlude presenting an interesting characterization of the Γ function by the functional equation and logarithmic convexity, due to Bohr and Mollerup. We introduce representation of convex functions by means of Green's function in Section 1.5, as a preparation for the representation formulas for subharmonic functions. In Section 1.6 we discuss some weaker notions of convexity which occur in microlocal analysis. Section 1.4 and most of Section 1.6 can be bypassed with no loss of continuity. The last section, Section 1.7, studies when the minimum of a family of (convex) functions is convex. The extension to (pluri-)subharmonic functions in Chapters III and IV will be essential in Chapter VII.

1.1. Definitions and basic facts. Let I be an interval on the real line $\mathbf{R}$, which may be open or closed, finite or infinite at either end, and let f be a real valued function defined in I.

Definition 1.1.1. f is called *convex* if the graph lies below the chord between any two points, that is, for every compact interval $J \subset I$, with boundary ∂J, and every linear function L we have

(1.1.1) $$\sup_J (f-L) = \sup_{\partial J}(f-L).$$

One calls f *concave* if $-f$ is convex.

Let $\partial J = \{x_1, x_2\}$. An arbitrary point in J can then be written $\lambda_1 x_1 + \lambda_2 x_2$ where $\lambda_j \geq 0$ and $\lambda_1 + \lambda_2 = 1$. Since $L(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 L(x_1) + \lambda_2 L(x_2)$, and we can choose L and a constant a with $L + a = f$ on ∂J, it follows that (1.1.1) is equivalent to

(1.1.1)′
$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2), \quad \text{if } \lambda_1, \lambda_2 \geq 0,\ \lambda_1 + \lambda_2 = 1,\ x_1, x_2 \in I.$$

If f is both convex and concave, then there must be equality in (1.1.1)′, that is, $f = L + a$ where L is linear and a is a constant. Such a function is called *affine*; it can of course be uniquely extended to all of $\mathbf{R}$. More generally, a map f between two vector spaces is called affine if it is of the form $f = L + a$ with L linear and a constant. This is equivalent to

$$(1.1.2) \qquad f(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 f(x_1) + \lambda_2 f(x_2), \quad \text{when } \lambda_1 + \lambda_2 = 1.$$

Indeed, if $L = f - f(0)$ we obtain $L(\lambda x) = \lambda L(x)$ when $x_2 = 0$, hence $L(x_1 + x_2) = L(x_1) + L(x_2)$ follows if $\lambda_1 = \lambda_2 = \frac{1}{2}$. This means that L is linear. Conversely, if $f = L + a$ with L linear we obtain not only (1.1.2) but more generally

$$(1.1.2)' \qquad f(\sum_1^n \lambda_j x_j) = \sum_1^n \lambda_j f(x_j), \quad \text{if } \sum_1^n \lambda_j = 1.$$

The following statements are immediate consequences of (1.1.1) or (1.1.1)′:

Theorem 1.1.2. *If f_j are convex functions in I and $c_j \in \mathbf{R}$ are ≥ 0, $j = 1, \dots, n$, then $f = \sum_1^n c_j f_j$ is a convex function in I.*

Theorem 1.1.3. *Let f_α, $\alpha \in A$, be a family of convex functions in I, and let J be the set of points $x \in I$ such that $f(x) = \sup_{\alpha \in A} f_\alpha(x)$ is $< +\infty$. Then J is an interval (which may be empty) and f is a convex function in J. If f_j, $j = 1, 2, \dots$, is a sequence of convex functions and J is the set of points $x \in I$ where $F(x) = \overline{\lim}_{j \to \infty} f_j(x) < +\infty$, then J is an interval and F is a convex function in J unless $F = -\infty$ in the interior of J or J consists of a single point.*

To prove the second statement one just has to write $F(x) = \lim F_N(x)$ where $F_N(x) = \sup_{j > N} f_j(x)$ and use the obvious first part.

Exercise 1.1.1. Prove that one cannot replace sup by inf or $\overline{\lim}$ by $\underline{\lim}$ in Theorem 1.1.3.

Exercise 1.1.2. Let I and J be two compact intervals with $J \subset I$ and lengths $|I|$, $|J|$, and let f be a convex function in I. Prove that if m and M are constants such that $f \leq M$ in I and $f \geq m$ in J then

$$f \geq M - (M - m)|I|/(|J| + d(J, \partial I)) \quad \text{in } I,$$

where $d(J, \partial I)$ is the shortest distance from J to ∂I and the denominator is assumed $\neq 0$.

Theorem 1.1.4. *Let f be a real-valued function defined in an interval I, and let φ be a function defined in another interval J with values in I. Then $f \circ \varphi$ is convex for every convex f if and only if φ is affine; and $f \circ \varphi$ is convex for every convex φ if and only if f is convex and increasing.*

Proof. If $f \circ \varphi$ is convex for $f(x) = x$ and for $f(x) = -x$, then φ is both convex and concave, hence affine. Conversely, if φ is affine it is obvious that $f \circ \varphi$ inherits convexity from f. Now assume that $f \circ \varphi$ is convex for every convex φ. Taking $\varphi(x) = x$ we conclude that f must be convex. If $y_1 < y_2$ are points in I, then $\varphi(x) = y_1 + (y_2 - y_1)|x|$ is convex in $[-1, 1]$, and if $f \circ \varphi$ is convex it follows since $f \circ \varphi(\pm 1) = f(y_2)$ and $f \circ \varphi(0) = f(y_1)$ that $f(y_1) \le f(y_2)$, so f must be increasing. Conversely, assume that f is increasing and convex, and let $x_1, x_2 \in I$, $\lambda_1, \lambda_2 \ge 0$, $\lambda_1 + \lambda_2 = 1$. Then

$$f(\varphi(\lambda_1 x_1 + \lambda_2 x_2)) \le f(\lambda_1 \varphi(x_1) + \lambda_2 \varphi(x_2)) \le \lambda_1 f(\varphi(x_1)) + \lambda_2 f(\varphi(x_2)),$$

where the first inequality holds since φ is convex and f is increasing, the second since f is convex. This completes the proof.

If $x_1 < x < x_2$ then $x = \lambda_1 x_1 + \lambda_2 x_2$ for $\lambda_1 = (x_2 - x)/(x_2 - x_1)$, $\lambda_2 = (x - x_1)/(x_2 - x_1)$, so (1.1.1)$'$ means that

$$(x_2 - x_1) f(x) \le (x_2 - x) f(x_1) + (x - x_1) f(x_2), \quad \text{that is,}$$

$$(f(x) - f(x_1))/(x - x_1) \le (f(x_2) - f(x))/(x_2 - x). \tag{1.1.1$''$}$$

Hence we have:

Theorem 1.1.5. *f is convex if and only if for every $x \in I$ the difference quotient $(f(x+h) - f(x))/h$ is an increasing function of h when $x + h \in I$ and $h \ne 0$.*

Corollary 1.1.6. *If f is convex then the left derivative $f_l'(x)$ and the right derivative $f_r'(x)$ exist at every interior point of I. They are increasing functions. If $x_1 < x_2$ are in the interior of I we have*

$$f_l'(x_1) \le f_r'(x_1) \le (f(x_2) - f(x_1))/(x_2 - x_1) \le f_l'(x_2) \le f_r'(x_2). \tag{1.1.3}$$

In particular, f is Lipschitz continuous in every compact interval contained in the interior of I.

There is no need for f to be continuous at the end points of I, but $f(x)$ has a finite limit when x converges to a finite end point of I belonging to I, again by Theorem 1.1.5. Changing the definition at the end points if necessary we can therefore assume that f is continuous also there. The right (left) derivative exists then at the left (right) end point but may be

$-\infty$ $(+\infty)$. If we allow f to take the value $+\infty$ we can always make the interval I closed.

Using the continuity of f we obtain from (1.1.3) if $x_1 < x_2$ are points in I

$$\lim_{\varepsilon\to+0} f_r'(x_1+\varepsilon) \le (f(x_2)-f(x_1))/(x_2-x_1) \le \lim_{\varepsilon\to+0} f_l'(x_2-\varepsilon). \tag{1.1.3$'$}$$

If we let $x_2 \downarrow x_1$ or $x_1 \uparrow x_2$, we obtain

Theorem 1.1.7. *If f is convex in I and x is an interior point, then*

$$f_r'(x) = \lim_{\varepsilon\to+0} f_r'(x+\varepsilon) = \lim_{\varepsilon\to+0} f_l'(x+\varepsilon), \tag{1.1.4}$$

$$f_l'(x) = \lim_{\varepsilon\to+0} f_r'(x-\varepsilon) = \lim_{\varepsilon\to+0} f_l'(x-\varepsilon). \tag{1.1.5}$$

We shall therefore write $f'(x+0) = f_r'(x)$, $f'(x-0) = f_l'(x)$. The following conditions are equivalent

(1) *f_l' is continuous at x;*
(2) *f_r' is continuous at x;*
(3) *$f_r'(x) = f_l'(x)$, that is, f is differentiable at x.*

These conditions are fulfilled except at countably many points.

Proof. The last statement follows from the fact that if $x_1 < x_2$ are points in I, then

$$\sum_{x_1<x<x_2} (f_r'(x) - f_l'(x)) \le f_l'(x_2) - f_r'(x_1) < \infty.$$

Exercise 1.1.3. Let $x_1, x_2, \dots$ be different real numbers and let $a_j > 0$ be chosen so that $\sum_1^\infty a_j(1+|x_j|) < \infty$. Show that

$$f(x) = \sum_1^\infty a_j|x-x_j|$$

is a convex function and that

$$f_r'(x_j) - f_l'(x_j) = 2a_j,\ \forall j; \quad f'(x) = \sum_1^\infty a_j \operatorname{sgn}(x-x_j) \quad \text{if } x \ne x_j\ \forall j.$$

Exercise 1.1.4. Show that if f_j are non-negative continuous convex functions in a compact interval I and $f = \sum_1^\infty f_j$ converges at the end points, then f is continuous and convex in I and

$$f'(x \pm 0) = \sum_1^\infty f_j'(x \pm 0)$$

for every x in the interior of I.

Exercise 1.1.5. Show that if f_j are convex functions in the interval I and $f_j(x) \to f(x) \in \mathbf{R}$ for every $x \in I$, then f is convex in I and $f_j \to f$ uniformly on every compact interval contained in the interior of I. Show that

$$f'(x-0) \le \varliminf_{j\to\infty} f_j'(x-0) \le \varlimsup_{j\to\infty} f_j'(x+0) \le f'(x+0), \quad j \to \infty,$$

for every x in the interior of I. Give an example where there is inequality throughout.

Exercise 1.1.6. Let I be an open interval and f_j a sequence of convex functions in I having a uniform upper bound on every compact interval $J \subset I$. Prove that either $f_j \to -\infty$ uniformly on every such interval J or else there is a subsequence f_{j_k} converging uniformly on every J to a convex function.

Exercise 1.1.7. Let f be convex in the interval I and bounded above there. Show that f is decreasing (increasing) if I is infinite to the right (left); thus f is constant if $I = \mathbf{R}$.

To prove a converse of Theorem 1.1.7 we need a suitable form of the mean value theorem:

Lemma 1.1.8. *Let f be a continuous function in a closed interval $\{x; a \le x \le b\}$ such that $f_r'(x)$ exists when $a \le x < b$. If $f_r'(x) \ge C$ for all such x then $f(b) - f(a) \ge C(b-a)$. If instead $f_r'(x) \le C$ then $f(b) - f(a) \le C(b-a)$.*

Proof. It suffices to prove the first statement. If $C' < C$ then

$$F = \{x \in [a,b]; f(x) - f(a) \ge C'(x-a)\}$$

is closed because f is continuous, and $a \in F$. The supremum y of F is in F, and $y = b$ since we would otherwise have

$$f(y+h)-f(a) = f(y+h)-f(y)+f(y)-f(a) > C'h+C'(y-a) = C'(y+h-a)$$

for sufficiently small $h > 0$, contradicting the definition of y. Hence $f(b) - f(a) \geq C'(b-a)$ for every $C' < C$ which proves the lemma.

The lemma gives the inequality

$$\inf_{[a,b)} f'_r \leq (f(b) - f(a))/(b-a) \leq \sup_{[a,b)} f'_r, \tag{1.1.6}$$

and we obtain:

Theorem 1.1.9. *If f is a continuous function in the interval I such that f'_r exists at every interior point x of I and increases with x, then f is convex, and $\int_x^y f'_r(t)\,dt = f(y) - f(x)$, $x, y \in I$. (The same result is true with f'_r replaced by f'_l.)*

Proof. (1.1.1)$''$ follows at once by (1.1.6) if $x_1 < x < x_2$ are interior points of I. Since f is continuous the inequality remains valid if x_1 or x_2 is a boundary point, so the convexity follows in view of Theorem 1.1.5. The second statement follows since the right derivative of $\int_x^y f'_r(t)\,dt$ with respect to y is equal to $f'_r(y)$ by the monotonicity and right continuity.

Corollary 1.1.10. *Let f be a continuous function in I which is in C^2 in the interior of I. Then f is convex if and only if $f'' \geq 0$ there. If $f'' > 0$ one calls f* strictly convex.

Example 1.1.11. $f(x) = e^{ax}$ is a convex function on $\mathbf{R}$ for every $a \in \mathbf{R}$. If $r \geq 1$, then $f_r(x) = x^r$ is a convex function when $x \geq 0$, if $r < 0$ then f_r is convex when $x > 0$, but if $0 < r \leq 1$ then x^r is concave when $x \geq 0$. The functions $g(x) = x \log x$ and $h(x) = -\log x$ are convex when $x > 0$.

Another immediate consequence of Theorem 1.1.9 and Corollary 1.1.6 is:

Corollary 1.1.12. *Convexity is a local property: If f is defined in an interval I and every point in I is contained in an open interval $J \subset I$ such that the restriction of f to J is convex, then f is convex.*

We have defined convexity in terms of affine majorants, but there is also an equivalent definition in terms of affine minorants:

Theorem 1.1.13. *A real-valued function f defined in an interval I is convex if and only if for every x in the interior of I there is an affine linear function g with $g \leq f$ and $g(x) = f(x)$.*

Proof. Assume that f is convex. Choose $k \in [f'_l(x), f'_r(x)]$ and let $g(y) = f(x) + k(y-x)$. Since $g(x) = f(x)$ and (1.1.3) gives

$$f(y) \geq f(x) + (y-x)f'_r(x) \geq g(x), \text{ if } y \geq x;$$
$$f(y) \geq f(x) + (y-x)f'_l(x) \geq g(y), \text{ if } y \leq x,$$

the necessity is proved. Now assume that f satisfies the condition in the theorem. We must prove that (1.1.1)$'$ holds. In doing so we may assume that $x_1 \neq x_2$ and that $\lambda_1\lambda_2 > 0$, which implies that $x = \lambda_1 x_1 + \lambda_2 x_2$ is an interior point of I. If g is an affine minorant of f with $f(x) = g(x)$ then

$$\sum_1^2 \lambda_j f(x_j) \geq \sum_1^2 \lambda_j g(x_j) = g(\sum_1^2 \lambda_j x_j) = g(x) = f(x),$$

which completes the proof.

In view of (1.1.2)$'$ the second part of the proof gives a much more general result with no change other than extension of the summation from 1 to n:

Theorem 1.1.14. *Let f be convex in the interval I, and let $x_1, \dots, x_n \in I$. Then we have*

$$(1.1.1)''' \qquad f\Big(\sum_1^n \lambda_j x_j\Big) \leq \sum_1^n \lambda_j f(x_j), \quad \text{if } \lambda_1, \dots, \lambda_n \geq 0, \ \sum_1^n \lambda_j = 1.$$

If $\lambda_j > 0$ for every j, then there is equality in (1.1.1)$'''$ *if and only if f is affine in the interval* $[\min x_j, \max x_j]$.

Exercise 1.1.8. Prove (1.1.1)$'''$ directly from (1.1.1)$'$ by induction with respect to n.

(1.1.1)$'''$ is usually called *Jensen's inequality*, and so is the following more general version involving integrals instead of sums:

Exercise 1.1.9. Let f be a convex function in the interval I, let T be a compact space with a positive measure $d\mu$ such that $\int_T d\mu(t) = 1$, and let $x(t)$ be a μ integrable function on T with values in I. Prove that

$$f\Big(\int_T x(t)d\mu(t)\Big) \leq \int_T f(x(t))d\mu(t).$$

Exercise 1.1.10. Let Π be an orthogonal projection in a finite dimensional Euclidean vector space E. Show that if A is a symmetric linear operator in E then

$$\operatorname{Tr}(\Pi f(\Pi A \Pi)\Pi) \leq \operatorname{Tr}(\Pi f(A)\Pi)$$

for every convex function f (Berezin's inequality). (Recall that if B is a linear transformation in E then $\operatorname{Tr} B = \sum(Be_j, e_j)$ if e_j is any orthonormal basis in E and $(\cdot,\cdot)$ denotes the scalar product. If B is symmetric then $f(B)$ has the same eigenvectors as B with every eigenvalue λ replaced by $f(\lambda)$. — Hint: Express both sides in terms of the eigenvectors of $\Pi A \Pi$ in ΠE and of A in E.)