ON CONJUGATE CONVEX FUNCTIONS

W. FENCHEL

1. Since the classical work of Minkowski and Jensen it is well known that many of the inequalities used in analysis may be considered as consequences of the convexity of certain functions. In several of these inequalities pairs of "conjugate" functions occur, for instance pairs of powers with exponents a and α related by $1/a + 1/\alpha = 1$. A more general example is the pair of positively homogeneous convex functions defined by Minkowski and known as the distance (or gauge) function and the function of support of a convex body. The purpose of the present paper is to explain the general (by the way rather elementary) idea underlying this correspondence. Subjected to a more precise formulation the result is the following:

To each convex function $f(x_1, \ldots, x_n)$ defined in a convex region G and satisfying certain conditions of continuity there corresponds in a unique way a convex region Γ and a convex function $\phi(\xi_1, \ldots, \xi_n)$ defined in Γ and with the same properties such that

$$x_1\xi_1 + \ldots + x_n\xi_n \leqq f(x_1, \ldots, x_n) + \phi(\xi_1, \ldots, \xi_n), \tag{1}$$

for all points $(x_1, \ldots, x_n)$ in G and all points $(\xi_1, \ldots, \xi_n)$ in Γ. The inequality is exact in a sense explained below. The correspondence between G, f and Γ, ϕ is symmetric, and the functions f and ϕ are called conjugate.[1]

The hypersurfaces $y = f(x_1, \ldots, x_n)$ and $\eta = \phi(\xi_1, \ldots, \xi_n)$ correspond to each other in the polarity with respect to the paraboloid

$$2y = x_1^2 + \ldots + x_n^2 .$$

Let $F(x)$ be strictly increasing for $x \geqq 0$. Then $f(x) = \int_0^x F(x)dx$ is convex, and its conjugate function is $\phi(\xi) = \int_0^\xi \Phi(\xi)d\xi$ where $\Phi(\xi)$ is the inverse function of $F(x)$. The inequality (1) for $n = 1$ therefore yields the well-known inequality of W. H. Young[2]

$$x\xi \leqq \int_0^x F(x)dx + \int_0^\xi \Phi(\xi)d\xi .$$

(1) may thus be considered as a generalization of this inequality.

Received March 24, 1948.

[1]The case $n = 1$ has been considered by S. Mandelbrojt [3] under the assumption that the ranges G and Γ are identical with the entire axis $-\infty < x < \infty$. This, however, is incompatible with the complete reciprocity between f and ϕ which will appear from an example given below. Mandelbrojt's formulation of the theorem is thus not quite correct due to the fact that the least upper bounds occurring in it may be infinite.

[2]See e.g. [2] p. 111.

If $f(x_1, \ldots, x_n)$ is positively homogeneous of degree one, then G is the entire space $x_1, \ldots, x_n$ while Γ is closed and bounded, and $\phi(\xi_1, \ldots, \xi_n)$ is identically zero. In this case (1) expresses that $f(x_1, \ldots, x_n)$ is the function of support of the convex body Γ.[3]

2. The euclidean spaces with coordinates $x_1, \ldots, x_n$ and $x_1, \ldots, x_n, y$ will be denoted by R^n and R^{n+1} respectively, points and vectors in these spaces by x and x, y respectively. Furthermore we write

$$x' + x'' = (x'_1 + x''_1, \ldots, x'_n + x''_n), \quad \lambda x = (\lambda x_1, \ldots, \lambda x_n),$$

$$\Sigma x\xi = x_1\xi_1 + \ldots + x_n\xi_n.$$

θ will always denote a number in the interval $0 < \theta < 1$.

The point set G of R^n is supposed to be convex, i.e. if x' and x'' belong to G, the whole segment $(1 - \theta)x' + \theta x''$ belongs to G. But G need neither be closed nor open nor bounded. The interior points of segments belonging to G are shortly called the interior points of G. All other points of accumulation of G, belonging to G or not, will be called the boundary or extreme points of G.

A function $f(x)$ defined in G is called convex if

$$f\big((1 - \theta)x' + \theta x''\big) \leqq (1 - \theta)f(x') + \theta f(x'') \tag{2}$$

for any two points x' and x'' of G and all θ. It is well known that this implies that $f(x)$ is continuous at the interior points of G. For our purpose we have also to consider the behaviour of $f(x)$ at the boundary points. Let x^* be a boundary point of G. For functions of one variable $\lim\limits_{x \to x^*} f(x)$ exists or is ∞. But this is not necessarily the case for functions of several variables. If x^* belongs to G the only general conclusion to be drawn from (2) is that

$$\underline{\lim}_{x \to x^*} f(x) \leqq f(x^*); \tag{3}$$

for, from

$$f\big((1 - \theta)x + \theta x^*\big) \leqq (1 - \theta)f(x) + \theta f(x^*)$$

it follows that

$$\underline{\lim}_{x \to x^*} f(x) \leqq \underline{\lim}_{\theta \to 1} f\big((1 - \theta)x + \theta x^*\big) \leqq f(x^*),$$

and (2) remains valid if $f(x^*)$ is replaced by any other value satisfying (3).

If necessary, we now change G and f by adding to G all those boundary points x^* not yet belonging to G for which $\underline{\lim}_{x \to x^*} f(x)$ is finite and by defining f at these and at the boundary points previously belonging to G by

$$f(x^*) = \underline{\lim}_{x \to x^*} f(x). \tag{4}$$

The new G and the function f obtained in this way are obviously again convex; for, let x' and x'' be arbitrary points of the new G and $x'_{(\nu)}$ and $x''_{(\nu)}$, $\nu = 1, 2, \ldots,$

[3]See e.g. [1] p. 23-24.

sequences of interior points of G such that

$$x'_{(\nu)} \to x',\ x''_{(\nu)} \to x'',\ f(x'_{(\nu)}) \to f(x'),\ f(x''_{(\nu)}) \to f(x''),$$

then we get from

$$f\big((1-\theta)x'_{(\nu)} + \theta x''_{(\nu)}\big) \leqq (1-\theta)f(x'_{(\nu)}) + \theta f(x''_{(\nu)})$$

for $\nu \to \infty$

$$\underline{\lim}_{x \to (1-\theta)x' + \theta x''} f(x) \leqq \underline{\lim}_{\nu \to \infty} f\big((1-\theta)x'_{(\nu)} + \theta x''_{(\nu)}\big) \leqq (1-\theta)f(x') + \theta f(x''),$$

which shows that $(1-\theta)x' + \theta x''$ belongs to G and that (2) is valid, as the left-hand side is $f\big((1-\theta)x' + \theta x''\big)$.

With (3) in mind we may say that (4) expresses that the functions which will be considered in the following are convex and semi-continuous from below, and G is "closed relative to f," i.e. all boundary points at which $\underline{\lim}\, f(x)$ is finite belong to G, or in other words, at each boundary point which does not belong to G we have $\lim f(x) = \infty$.

3. The theorem to be proved may now be formulated thus:

Let G be a convex point set in R^n and $f(x)$ a function defined in G convex and semi-continuous from below and such that $\lim_{x \to x^*} f(x) = \infty$ for each boundary point x^* of G which does not belong to G. Then there exists one and only one point set Γ in R^n and one and only one function $\phi(\xi)$ defined in Γ with exactly the same properties as G and $f(x)$ such that

$$\Sigma x\xi \leqq f(x) + \phi(\xi), \tag{5}$$

where to every interior point x of G there corresponds at least one point ξ of Γ for which equality holds.

In the same way $G, f(x)$ correspond to $\Gamma, \phi(\xi)$.

We define Γ as the set of all points ξ with the property that the function $\Sigma x\xi - f(x)$ is bounded from above in G, and we define $\phi(\xi)$ in Γ as the least upper bound of this function:

$$\phi(\xi) = \underset{x \in G}{\text{l.u.b.}} \big(\Sigma x\xi - f(x)\big).$$

Then (5) is valid. The inequality $\Sigma x\xi - f(x) \leqq z$ or

$$f(x) \geqq \Sigma x\xi - z$$

means that the hyperplane $y = \Sigma x\xi - z$ in R^{n+1} with the normal vector $\xi, -1$ lies nowhere above the hypersurface $y = f(x)$, and $-z$ is the intercept of this hyperplane on the y-axis. It is a well-known fact that there exists at least one hyperplane of support of the convex hypersurface, i.e. a hyperplane which contains at least one point of the hypersurface and lies nowhere above it. This shows that Γ is not empty. Further we see that if there exists a hyperplane of support with the normal vector $\xi, -1$ and if $x^o, f(x^o)$ is a point of contact, then we have

$$\phi(\xi) = \Sigma x^o\xi - f(x^o),$$

and $-\phi(\xi)$ is the y-intercept of this hyperplane. If x^o is an arbitrary interior point of G, a hyperplane of support through $x^o, f(x^o)$ exists, and this proves the assertion on the equality sign in (5).

It is evident that Γ and $\phi(\xi)$ are convex. In fact, let ξ' and ξ'' be arbitrary points of Γ, then we have for $x\epsilon G$,

$$\Sigma x\xi' - f(x) \leqq \phi(\xi'), \ \Sigma x\xi'' - f(x) \leqq \phi(\xi''),$$

hence $$\Sigma x\big((1-\theta)\xi' + \theta\xi''\big) - f(x) \leqq (1-\theta)\phi(\xi') + \theta\phi(\xi'')$$

which shows that $(1-\theta)\xi' + \theta\xi''$ belongs to Γ and that

$$\phi\big((1-\theta)\xi' + \theta\xi''\big) \leqq (1-\theta)\phi(\xi') + \theta\phi(\xi'').$$

Let now ξ^* be a boundary point of Γ and $\xi\epsilon\Gamma$, $x\epsilon G$. Then it follows from (5) that

$$\varliminf_{\xi \to \xi^*} \phi(\xi) \geqq \Sigma x\xi^* - f(x)$$

and this shows on the one hand that $\xi^*\epsilon\Gamma$ if $\varliminf \phi(\xi)$ is finite, i.e. that Γ is closed relative to $\phi(\xi)$, and on the other hand that

$$\varliminf_{\xi \to \xi^*} \phi(\xi) \geqq \phi(\xi^*),$$

i.e. that $\phi(\xi)$ is semi-continuous from below. Hence Γ and ϕ have the same properties as G and f.

4. It remains to be proved that if we start with Γ and $\phi(\xi)$ the same procedure gives G and $f(x)$ again. We have to consider the set G^* of all points x for which $\Sigma\xi x - \phi(\xi)$ is bounded from above in Γ, together with the function

$$f^*(x) = \underset{\xi\epsilon\Gamma}{\text{l.u.b.}} \big(\Sigma\xi x - \phi(\xi)\big)$$

defined in G^*.

If $x\epsilon G$ we get from (5)

$$\Sigma\xi x - \phi(\xi) \leqq f(x) \tag{6}$$

for all $\xi\epsilon\Gamma$, hence $G \subset G^*$ and $f^*(x) \leqq f(x)$ in G. But to an interior point x of G there corresponds a ξ such that equality is valid in (6), which implies $f^*(x) \geqq f(x)$. Hence $f^*(x) = f(x)$ at the interior points of G and, as both functions are convex and semi-continuous from below, also at the boundary points of G.

Let now x^o be a point of R^n not in G. We have to prove that it does not belong to G^*, i.e. that

$$\underset{\xi\epsilon\Gamma}{\text{l.u.b.}} \big(\Sigma\xi x^o - \phi(\xi)\big) = \infty. \tag{7}$$

Since the quantity $\Sigma\xi x^o - \phi(\xi)$ is the y-coordinate of the point at which the hyperplane

$$y = \Sigma\xi x - \phi(\xi)$$

of R^{n+1} intersects the line $x = x^o$ parallel to the y-axis, we have to show that there are hyperplanes below the hypersurface $y = f(x)$ which have arbitrary large intercepts on the line $x = x^o$. Suppose first that x^o is an exterior point of G. Then there exists a hyperplane H parallel to the y-axis which separates the line $x = x^o$ from G and $y = f(x)$. Consider any hyperplane of support S of $y = f(x)$. Let S turn around the intersection of H and S so that the part lying below $y = f(x)$ moves downwards. Then the point at which S intersects the line $x = x^o$ moves upwards and tends to infinity. Suppose next that x^o is a boundary point of G but not belonging to G. Then we have $f(x) \to \infty$ for $x \to x^o$. Consider any segment belonging to G and having x^o as one of its end points. Let x' be a fixed point and x'' a variable point of the segment between

https://doi.org/10.4153/CJM-1949-007-x Published online by Cambridge University Press

x' and x^o such that $f(x'') > f(x')$. A plane of support through x'', $f(x'')$ then intersects the line $x = x^o$ at a point the y-coordinate of which is greater than $f(x'')$ and therefore tends to infinity if $x'' \to x^o$. This completes the proof of the theorem.

5. In section **1** it has been asserted that the hypersurfaces $y = f(x)$ and $\eta = \phi(\xi)$ correspond to each other in the polarity with respect to $2y = \Sigma x^2$. This is obviously true in the sense that each of the hypersurfaces is the envelope of the polar hyperplanes of the points of the other. For $y = f(x)$ may be considered as the envelope of the hyperplanes

$$y = \Sigma x\xi - \phi(\xi),$$

where $\xi \epsilon \Gamma$ is the parameter, and the poles of these hyperplanes are the points ξ, $\phi(\xi)$.

6. Suppose now that $y = f(x)$ is strictly convex, i.e. each hyperplane of support contains only one point of $y = f(x)$. Let further $\eta = \phi(\xi)$ satisfy the same condition; for $y = f(x)$ this means that there passes at most one hyperplane of support through a point of $y = f(x)$. Then $f(x)$ has continuous derivatives[4]

$$f_i(x) = \frac{\partial f}{\partial x_i}$$

and we have

$$\xi_i = f_i(x).$$

These relations establish a continuous one to one correspondence between the interior points of G and those of Γ. Solving them with respect to the x we get

$$x_i = \phi_i(\xi)$$

where, for reasons of symmetry, the ϕ_i must be the derivatives of ϕ. From this it is seen that in the case of $n = 1$ the derivatives of two conjugate convex functions are mutually inverse functions. This proves the assertion of section **1** on the inequality of Young. Furthermore we get an explicit expression for $\phi(\xi)$ if $f(x)$ is given, viz.

$$\phi(\xi) = \sum_{i=1}^{n} \xi_i\phi_i(\xi) - f\left(\phi_i(\xi)\right)$$

valid in the interior of Γ. Hence, our correspondence between f and ϕ is the Legendre transformation of the theory of differential equations.

References

[1] T. Bonnesen, W. Fenchel, *Theorie der konvexen Körper* (Berlin, 1934).

[2] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities* (Cambridge, 1934).

[3] S. Mandelbrojt, "Sur les fonctions convexes," *C.R. Acad. Sci. Paris*, vol. 209 (1939), 977-978.

The Technical University of Denmark
Copenhagen

[4]See [1] p. 23, 26. The argument used there in the case of positively homogeneous convex functions may easily be generalized to the case considered here.