A CLASS OF NEGATIVE-DEFINITE FUNCTIONS

C. S. HERZ

Let G be a finite-dimensional real vector space. A *proper negative-definite function* defined on G is a complex-valued function ψ with the property that for each $t > 0$

$$\exp\{-t\psi(x)\} = \int_{G'} \exp(i\xi x) P(t, d\xi)$$

where $P(t, \cdot)$ is a (Radon) probability measure on the dual space, G'.

We shall be concerned with real, homogeneous, negative-definite functions, i.e., those negative-definite functions for which there exists a positive constant α such that for each scalar λ and each vector $x \in G$ we have

$$\psi(\lambda x) = |\lambda|^\alpha \psi(x).$$

The associated probability measures here correspond to the *symmetric stable laws* of Paul Lévy [6].

It is very easy to see that one must have $\alpha \leq 2$. We shall restrict our attention to the range $1 \leq \alpha \leq 2$ and treat $\psi^{1/\alpha}$ instead of ψ. Let us make a definition and alter the notation slightly.

DEFINITION. Suppose $1 \leq p \leq 2$. A continuous non-negative function ψ defined on G is an L^p-norm if

(i) $\psi(\lambda x) = |\lambda| \psi(x)$ for each real λ and each $x \in G$, and

(ii) ψ^p is a proper negative-definite function.

The concern of this paper is what are the L^p-norms on G. In one sense the question has been answered by Lévy; see Theorem 1 below. It does not appear to me, however, that the connection between normed vector spaces and symmetric stable laws is obvious from Lévy's presentation, and I think the connection is an illuminating one. The central idea of the present article is that the terminology "L^p-norm" is apt.

It is well known (cf. [2]) that if $0 < \beta \leq 1$ then ψ^β is a proper negative-definite function whenever ψ is. A fortiori, an L^p-norm is an L^r-norm for $1 \leq r \leq p$. The simplest example of an L^2-norm is a function ψ of the form $\psi(x) = |\xi x|$ where $\xi \in G'$. Since sums and positive multiples of negative-definite functions are negative-definite, the L^p-norms on G form a cone which contains, in particular, functions of the type

---

$$\psi(x) = \left\{ \sum_{i=1}^{n} | \xi_i x |^p \right\}^{1/p}$$

where $\xi_1, \cdots, \xi_n$ is any finite set of elements of G'. An illusory greater generality is obtained by writing

$$\psi(x) = \left\{ \sum_{i=1}^{n} \sigma_i | \xi_i x |^p \right\}^{1/p}$$

where $\sigma_1, \cdots, \sigma_n$ are non-negative numbers, but this form suggests the generalization to

$$\psi(x) = \left\{ \int_{G'} | \xi x |^p \sigma(d\xi) \right\}^{1/p}.$$

There is no point in integrating over all of G'; on account of the homogeneity we can assume that σ is supported on the unit sphere S of G'. (Since we have not assumed any metric structure in G', "unit sphere" may be interpreted as a compact set which meets each ray from the origin exactly once.) There is no need to go beyond this generalization. Lévy [6, §63] proved

THEOREM 1. *A necessary and sufficient condition that ψ be an L^p-norm on G is that there exist a non-negative measure σ on the unit sphere S in G' such that*

$$\psi(x) = \left\{ \int_{S} | \xi x |^p \sigma(d\xi) \right\}^{1/p}.$$

Historically, real proper negative-definite functions arose in geometry. ψ is such a function if and only if G, with the metric $d(x, y) = \psi^{1/2}(x-y)$, has an isometric embedding in a real hilbert space (cf. [7]), and the references there and, for a more extensive theory including the treatment of complex negative-definite functions (cf. [5]). Theorem 1 may be viewed as asserting that the L^p-norms correspond to linear isometric embeddings in L^p spaces. We shall now see that the geometrical viewpoint still has something to offer. By a "norm" on G we shall understand a function ψ satisfying the definition of an L^p-norm with (ii) replaced by (ii') $\psi(x+y) \leqq \psi(x) + \psi(y)$. (We shall allow the possibility that $\psi(x) = 0$ for some $x \neq 0$.) It follows from Theorem 1 that an L^p-norm is a norm. We are now in a position to ask what norms on G are L^p-norms. Of course, when dim $G = 1$ every norm is an L^2-norm; but in higher dimensions it is not always easy to decide, even using Lévy's results, when a given norm is an L^p-norm. It helps to look at the problem in terms of Minkowski geometry.

Consider the set $K = \{x \in G : \psi(x) \leqq 1\}$. This is a convex body which is symmetrical with respect to the origin, and ψ is just the Minkowski gauge functional of K, i.e., $\psi(x) = \inf \lambda$ where $\lambda > 0$ and $\lambda^{-1} x \in K$. The condition that ψ be an L^p-norm should be reflected in the geometry of K, the unit ball in the ψ-norm. The least restrictive condition is that ψ should be an L^1-norm. We can prove by completely elementary means (without using Lévy's representation, Theorem 1),

THEOREM 2. *Every norm on a 2-dimensional vector space is negative-definite, i.e., it is an L^1-norm.*[1]

PROOF. Let ψ be a norm and K the associated unit ball. We may as well assume that K is bounded, i.e., $\psi(x) > 0$ for $x \neq 0$, for otherwise the situation is reducible to a lower-dimensional one. In this case $\int_G \exp\{-t\psi(x)\} dx < \infty$ for each $t > 0$, and we shall deduce this theorem by showing that

$$I(t, \xi) = \int_G \exp\{-t\psi(x) - i\xi x\} dx$$

is non-negative for each $t > 0$ and each $\xi \in G'$ whence we have

$$\exp\{-t\psi(x)\} = (2\pi)^{-n} \int_{G'} \exp(i\xi x) I(t, \xi) d\xi \quad \text{where} \quad n = \dim G.$$

We may start to work with n arbitrary although the theorem is rather trivial when $n = 1$ and is false in general for $n > 2$. Let ω be the differential form

$$\omega = \psi^{-n}(\xi_1 d\xi_2 \wedge \cdots \wedge d\xi_n \pm \cdots), \; n \text{ terms,}$$

where $\xi_1, \cdots, \xi_n$ is a system of linear forms with determinant 1, and ω is invariant under even permutations of the ξ_j's. It is easy to see that $dx = \psi^{n-1} d\psi \wedge \omega$ and hence

$$I(t, \xi) = (n - 1)! \int_\Gamma (t + i\xi)^{-n} \omega,$$

where Γ is the boundary of K. In case $n = 2$, it suffices to establish the theorem when K is a polygon, and for a given polygon, K, it is

---

[1] *Added in proof.* This theorem and the counterexample which follow have recently appeared in Thomas S. Ferguson, *A representation of the symmetric bivariate Cauchy distribution*, Ann. Math. Statist. 33 (1962), 1256–1266. Although Ferguson's work and this paper are independent of each other and use rather different language, the two proofs of this theorem are fundamentally identical.

enough to show that $I(t, \xi) \geq 0$ when ξ is not orthogonal to any edge of K. We take ξ as one coordinate; let η be any other so that $\omega = \xi d\eta - \eta d\xi$ on Γ. The part of Γ lying in $\eta \geq 0$ is described by a system of vertices: $(a_0, b_0), (a_1, b_1), \cdots, (a_s, b_s)$; and edges: $\eta = b_k + \alpha_k(\xi - a_k)$ for $a_k \leq \xi \leq a_{k+1}$. All the vertices (a_j, b_j) except possibly (a_0, b_0) and (a_s, b_s) are vertices of Γ. On account of the symmetry, $-b_0 = b_s$ and $-a_0 = a_s = a = \sup_{x \in K} \xi x$. In each case $a_j < a_{j+1}$ and $\alpha_j > \alpha_{j+1}$ since K is convex. We also notice that $b_{j+1} - b_j = \alpha_j(a_{j+1} - a_j)$. The contribution to $I(t, \xi)$ of the jth edge is evidently

$$i(b_j - \alpha_j a_j)(t + ia_j)^{-1} - i(b_{j+1} - \alpha_j a_{j+1})(t + ia_{j+1})^{-1}.$$

Thus the whole contribution along with the part of Γ where $\eta \geq 0$ is

$$i\alpha_0 a_0(t + ia_0)^{-1} + i\alpha_{s-1} a_s(t + ia_s)^{-1} - i \sum_{j=1}^{s-1} (\alpha_j - \alpha_{j-1}) a_j (t + ia_j)^{-1}.$$

$I(t, \xi)$ is twice the real part of this, i.e., twice

$$\alpha_0 a_0^2 (t^2 + a_0^2)^{-1} - \alpha_{s-1} a_s^2 (t^2 + a_s^2)^{-1} + \sum_{j=1}^{s-1} (\alpha_j - \alpha_{j-1}) a_j^2 (t^2 + a_j^2)^{-1}.$$

Each $\alpha_j - \alpha_{j-1}$ is negative; therefore, since $u^2(t^2 + u^2)^{-1}$ is an increasing function of u, we only make matters worse by replacing each $a_j^2(t^2 + a_j^2)^{-1}$ with $a^2(t^2 + a^2)^{-1}$. We now have

$$I(t, \xi) \geq 2a^2(t^2 + a^2)^{-1} \left\{ \alpha_0 - \alpha_{s-1} + \sum_{j=1}^{s-1} (\alpha_j - \alpha_{j-1}) \right\} = 0.$$

We have thus shown that $I(t, \xi) \geq 0$ when K is a polygon and ξ is not orthogonal to any edge of K. By the continuity in ξ we may conclude $I(t, \xi) \geq 0$ for all ξ. Thus ψ is negative-definite when K is a polygon. Each convex set in the plane is approximable by polygons in such a way that the corresponding Minkowski functionals converge pointwise; this completes the proof in all cases.

The technique of calculation used above is usually excessively complicated when $n > 2$. We can, however, use it to show that in case $n = 3$, if K is a cube then ψ is not negative-definite.

COUNTEREXAMPLE. Suppose ξ, η, ζ is a set of three linearly independent linear forms on G. Put $\psi(x) = \max |\xi x|, |\eta x|, |\zeta x|$. Then ψ is not negative-definite.

PROOF. Without loss of generality we may take $\dim G = 3$ and ξ, η, ζ with determinant 1. We shall calculate explicitly that $I(t, \lambda\xi) = 16t(t^2 + \lambda^2)^{-3}(3t^2 - \lambda^2)$ which is negative for $\lambda^2 > 3t^2$. Γ, the boundary of $K = \{x : \psi(x) \leq 1\}$, has three pairs of opposite faces. We recall that

$$I(t, \lambda\xi) = 2 \int_{\Gamma} (t + i\lambda\xi)^{-3}\omega.$$

The face on which $\xi x = 1$ has area 4, and thus it contributes $8(t+i\lambda)^{-3}$ to $I(t, \lambda\xi)$. Thus the faces on which $|\xi x| = 1$ contribute $16(t^3 - 3t\lambda^2)(t^2 + \lambda^2)^{-3}$. The contribution to $I(t, \lambda\xi)$ of each of the other four faces is $2\int_{-1}^{1}d\eta\int_{-1}^{1}(t+i\lambda\xi)^{-3}d\xi = 8t(t^2+\lambda^2)^{-2}$. Thus $I(t, \lambda\xi) = 16(t^3 - 3t\lambda^2)(t^2+\lambda^2)^{-3} + 32t(t^2+\lambda^2)^{-2}$, as claimed.

As it turns out, the theorem we have proved is really a consequence of Lévy's Theorem. His condition that ψ be an L^1-norm is

$$\psi(x) = \int_{S} |\theta x|\, \sigma\, (d\theta)$$

where there is no harm in assuming that

$$\int_{S} \theta x\, \sigma\, (d\theta) \equiv 0.$$

This last condition is necessary and sufficient for the non-negative measure σ to be the spherical representative of the surface area measure on a convex set C in G' (cf. [3, §60]). Here we want S to be a genuine unit sphere in some euclidean metric on G'. Let $|x|$ be the dual norm on G. For $x \in G$ with $|x| = 1$, we see that $\psi(x)$ is simply twice the area of the projection of C on a plane perpendicular to the direction x. Thus, changing C by a scale factor to eliminate a 2, the L^1-norm ψ is the Quermass function of a convex set C in G'. (See [3, §§30–32] for this and what follows.) Now let K' be the polar of K, i.e., the set of points $\xi \in G'$ such that $\xi x \leqq 1$ for all $x \in K$. ψ is the support function of K'. Putting things together, we see that K' is the *projection-body* of C; it is the convex set whose supporting planes are at a distance from the origin equal to the area of the projection of C on the plane in question. Therefore, we have deduced

COROLLARY TO THEOREM 1. *A necessary and sufficient condition that ψ be an L^1-norm is that it be the support function of a compact set K' which is a projection-body.*

Now in the 2-dimensional case, any norm ψ is the support function of a symmetric convex set K', and it is easy to see that such a set is the projection-body of a $\pi/2$-rotation of itself contracted by half. Thus Theorem 2 is a corollary of the corollary.

The case of L^1-norms in the plane, covered by Theorem 2, is the simplest nontrivial one. The general question of finding an effective

characterization of L^p-norms on G where $1 < p < 2$ and dim $G > 1$, or even where $p = 1$ and dim $G > 2$, seems to be quite difficult. Therefore, we shall now confine our attention to a special class of functions.

DEFINITION. Suppose $1 \leqq q \leqq \infty$. A function ψ defined on G is an l^q-norm if there exist $\xi_1, \cdots, \xi_n \in G'$ such that span $\{\xi_1, \cdots, \xi_n\} = G'$ and for each $x \in G$

$$\psi(x) = \left\{ \sum | \xi_i x |^q \right\}^{1/q}.$$

As we observed earlier, an l^q-norm is an L^p-norm whenever $1 \leqq p \leqq q \leqq 2$; and it is trivial that each L^2-norm is an l^2-norm. Aside from these statements we have to be content with some negative facts. What will be decisive is a fundamental property of L^p-spaces, [4].

LEMMA. *A necessary condition that ψ be an L^p-norm is that for each pair $x, y \in G$ we have*

$$\left\{ \psi^{p'}\left(\frac{x+y}{2} \right) + \psi^{p'}\left(\frac{x-y}{2} \right) \right\}^{1/p'} \leqq \left\{ \frac{\psi^p(x) + \psi^p(y)}{2} \right\}^{1/p}$$

where $p' = p(p-1)^{-1}$.

From this lemma we may deduce by means of simple estimates

THEOREM 3. *Let ψ be an l^q-norm on G where* dim $G > 1$. *A necessary condition that ψ be an L^p-norm is that $p \leqq q \leqq p'$.*

The geometrical interpretation of an l^∞-norm is clear: a norm ψ is an l^∞-norm if and only if its unit sphere, $\{x \in G : \psi(x) = 1\}$, is a polyhedron. (A fortiori, l^1-norms are l^∞-norms.) Theorem 3 shows that when dim $G > 1$, an l^∞-norm can never be an L^p-norm for $p > 1$. Thus, in contrast to the technique used to prove Theorem 2, approximation by polyhedra is useless for determining what symmetrical convex sets are associated with L^p-norms for $p > 1$.

The material of this paper is closely related to some of the topics in Chapters XI and XII of Banach's book [1]. It seems likely that the ideas carry over to infinite-dimensional vector spaces (see [5] for negative-definite functions on such spaces) but some details are still lacking.

BIBLIOGRAPHY

1. S. Banach, *Théorie des opérations linéaires*, Monogr. Mat., Vol. I, Warsaw, 1932.
2. S. Bochner, *Stable laws of probability and completely monotone functions*, Duke Math. J. **3** (1937), 726–728.
3. T. Bonnesen and W. Fenchel, *Theorie der Konvexen Körper*, Springer, Berlin, 1934.

4. J. A. Clarkson, *Uniformly convex spaces*, Trans. Amer. Math. Soc. 40 (1936), 396–414.

5. C. S. Herz, *Negative definite functions* (to appear).

6. P. Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.

7. I. J. Schoenberg, *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc. 44 (1938), 522–536.

CORNELL UNIVERSITY

PSEUDO-UNIFORM CONVEXITY IN H^1

D. J. NEWMAN

Although the familiar Banach space, H^1, is not uniformly convex, we will display two important properties which H^1 shares with the uniformly convex spaces.

Loosely speaking, these two properties are

1. "Weak convergence+convergence of norms⇒convergence" and

2. "A certain kind of minimizing sequence automatically converges."

We delay the precise statements and take this opportunity to point out that these two results were originally conjectured by A. Shields and H. S. Shapiro, both of whom also supplied valuable ideas and suggestions leading to the solutions.

Definitions and notation. Functions of the class H^1 are denoted by f, g, $\cdots$. The H^p norm of f ($1 \leq p \leq \infty$) is written $\|f\|_p$ and we abbreviate $\|f\|_1$ to $\|f\|$. By "weak" convergence of f_n to f, written $f_n \to^w f$, we mean that $f_n(Z)$ converges uniformly to $f(Z)$ in every compact subset of $|Z| < 1$. Ordinary convergence, written $f_n \to f$ means of course that $\|f_n - f\| \to 0$.

Let $F(\theta)$ be any function of class L^1 $(0, 2\pi)$, and consider the problem of minimizing $\int_0^{2\pi} |F(\theta) - f(e^{i\theta})| d\theta$ over all $f \in H^1$. It is a theorem of Rogosinski and Shapiro (see [1]) that this problem has a unique solution for each F, and this solution we denote by f_F. We also denote $|F(\theta) - f_F(e^{i\theta})| = P_F(\theta)$ and $\int_0^{2\pi} P_F(\theta) d\theta = d_F$.

With these definitions, then, we are able to give our precise statements.

THEOREM 1. *If $f_n \to^w f$ and $\|f_n\| \to \|f\|$ then $f_n \to f$.*