An Inequality for Hyperbolic Polynomials

LARS GÅRDING

Communicated by S. S. CHERN

1. Introduction. Let us equip real n -space R^n with the Lorentz metric

$$xy = x_1y_1 - x_2y_2 - \cdots - x_ny_n,$$

and let C be the forward light-cone

$$x_1 > 0, \quad xx > 0.$$

Then we have the following well-known elementary inequality:

$$xy \geq (xx)^{\frac{1}{2}}(yy)^{\frac{1}{2}}, \quad x, y \in C.$$

Equality holds if and only if x and y are proportional. The geometrical meaning of this is that the hyperboloid sheet $xx = c > 0$, $x_1 > 0$, lies on the positive side of any of its tangent planes, *i.e.* the sheet is convex. The inequality is a special case of an inequality for homogeneous hyperbolic polynomials which is used in the preceding paper by S. S. CHERN. This general inequality is an immediate consequence of various properties of hyperbolic polynomials proved in the second chapter of [1], but since the emphasis in that paper is on non-homogeneous polynomials and the inequality is interesting in itself, we shall give an independent proof of it. Since most relevant properties of hyperbolic polynomials enter into this proof, this article contains, in fact, a rather complete theory of homogeneous hyperbolic polynomials.

1. Hyperbolic polynomials. Let $P = P(x)$ be a homogeneous polynomial of degree $m > 0$ in n variables $x = \{x_k\}_1^n$, and let $a = \{a_k\}_1^n$ be real. We say that P is hyperbolic with respect to a or, using a convenient solecism, that P is hyperbolic a , if the equation $P(sa + x) = 0$ has m real zeros for every real x . Geometrically, this means that any straight line in R^n with the direction a meets the real hypersurface $P = 0$ in m real points. An equivalent definition is that

$$P(a) \neq 0, \quad P(sa + x) \neq 0 \quad \text{for} \quad \text{Im } s \neq 0, \quad x \text{ real.}$$

If we factorize,

$$(1) \quad P(sa + x) = P(a) \prod_1^m (s + \lambda_k(a, x)),$$

still another equivalent definition is that $P(a) \neq 0$ and that all the numbers $\{\lambda_k\}_1^m$ are real whenever x is. From this form of the definition we see that the polynomial $P(x)/P(a)$ is real when x is real, so that it has real coefficients. Thus a hyperbolic polynomial is essentially real, and we may without loss of generality restrict ourselves to real polynomials. We notice in passing that the fact that P is homogeneous implies that the numbers λ defined by (1) are homogeneous of order 1 in x . More generally,

$$(2) \quad \lambda_k(a, tx + sa) = t\lambda_k(a, x) + s$$

where t and s are any complex numbers and both sides are suitably labelled. We also have

$$(3) \quad \lambda_k(ta, x) = \frac{1}{t} \lambda_k(a, x), \quad t \neq 0.$$

Let us now give some examples of hyperbolic polynomials. If $m = 1$, then P is hyperbolic any a with $P(a) \neq 0$. Consider the case $m = 2$. Replacing P by $P/P(a)$, we may assume without loss of generality that $P(a) = 1$. Let $Q(x, y)$ be the polarized form of P . That

$$P(sa + x) = s^2 Q(a, a) + 2s Q(a, x) + Q(x, x) = 0$$

has only real zeros is then equivalent to the inequality

$$P(x) = Q(x, x) \leq Q^2(a, x).$$

This means that

$$Q^2(a, x) - P(x) = L_2^2 + L_3^2 + \dots$$

is a sum of squares or vanishes identically. Hence

$$P = L_1^2 - L_2^2 - \dots, \quad L_1(x) = Q(a, x)$$

is itself a square or else has Lorentz signature. Conversely, if P has this form, it is hyperbolic with respect to any a with $P(a) > 0$. The proof is left to the reader. The name hyperbolic comes from this connection with the quadrics.

It follows directly from the definition that if P and Q are hyperbolic a , then PQ is. Hence there are hyperbolic polynomials of any degree. Applying the following lemma, we can construct a large class of hyperbolic polynomials.

Lemma 1. *If P is hyperbolic a and $m > 1$, then*

$$Q(x) = \sum a_k \frac{\partial}{\partial x_k} P(x)$$

is hyperbolic a .

The proof is immediate. We have

$$Q(sa + x) = \frac{d}{ds} P(sa + x),$$

so that, by Rolle's theorem, the equation $Q(sa + x) = 0$ has $m - 1$ real zeros separating those of the equation $P(sa + x) = 0$. Repeated application of this lemma shows that the polynomials $\{P_i\}_1^m$ defined by

$$P(sa + x) = \sum s^{m-i} P_i(x)$$

are hyperbolic a if P is.

We finish this section by giving a few standard examples of hyperbolic polynomials.

Example 1. The polynomial

$$P = x_1^2 - x_2^2 - \cdots - x_n^2$$

is hyperbolic $(1, 0, \cdots, 0)$. Actually it is hyperbolic any a with $P(a) > 0$.

Example 2. The polynomial

$$P = x_1 \cdots x_n$$

is hyperbolic any a with $P(a) \neq 0$. Since

$$P(sa + x) = \sum_0^n s^{n-i} S_i(x),$$

where $a = (1, \cdots, 1)$ and

$$S_i = \sum x_1 \cdots x_i$$

are the elementary symmetric functions, any S_i with $j > 0$ is hyperbolic $(1, \cdots, 1)$.

Example 3. Let $n = \frac{1}{2}p(p+1)$ be a triangular number and let x be the symmetric $p \times p$ matrix (x_{ik}) . Then the polynomial

$$P(x) = \det x$$

is hyperbolic any (positive or negative) definite a since in this case the equation $P(sa + x) = 0$ has only real zeros. In this case, the numbers λ defined by (1) are the eigenvalues of x with respect to a .

Example 4. Let $n = p^2$ be a square and let x be the hermitian matrix (x_{ik}) . The real components of x can be taken to be the p^2 numbers $\operatorname{Re} x_{ik}$ and $\operatorname{Im} x_{ik}$. Then

$$P(x) = \det x$$

is hyperbolic any definite a .

In this and the preceding example, the traces $\{P_i\}_0^p$ defined by

$$P(se + x) = \sum s^{p-i} P_i(x),$$

where e is the unit matrix, are hyperbolic e .

2. The cone. It follows directly from the definition that if P is hyperbolic a , then it is hyperbolic any non-zero multiple of a . We shall see that P is also hyperbolic any b belonging to either of two opposite cones, one containing a and the other containing $-a$.

Definition. Let $C = C(P, a)$ be the set of all x such that $P(ta + x) \neq 0$ when $t \geq 0$.

It is clear that C contains every positive multiple of a and that if we put

$$h(a, x) = \min_k \lambda_k(a, x),$$

an equivalent definition is

$$h(a, x) > 0.$$

Since the left side is continuous, C is open and bounded by parts of the hyper-surface $P = 0$ and it contains that component C' of the complement of $P = 0$ which contains a . As we shall see later, $C = C'$. We give a few examples. If $m = 1$, then C is the half-space $P(a)P(x) > 0$. In example 1, C is the region $P(x) > 0, x_1 a_1 > 0$, in example 2, it is the octant $x_1 a_1 > 0, \dots, x_n a_n > 0$ and in the examples 3 and 4 it consists of all positive definite matrices if a is positive definite and all negative definite matrices if a is negative definite.

Let us turn to the general case. As we have already remarked, C is open and contains a . The equation (3) shows that

$$C(P, -a) = -C(P, a),$$

and (2) that

$$tC(P, a) = C(P, a), \quad t > 0.$$

The following theorem is fundamental.

Theorem 1. If P is hyperbolic a and $b \in C(P, a)$, then the polynomial

$$Q(s, t) = P(sa + tb + x), \quad x \text{ real},$$

has no zeros when

$$(4) \quad \operatorname{Im} s \operatorname{Im} t \geq 0, \quad (\operatorname{Im} s)^2 + (\operatorname{Im} t)^2 > 0$$

or when

$$(5) \quad t > 0, \quad s + h(a, x) \geq 0.$$

Proof. Since

$$Q = P(a)s^m + P(b)t^m + \dots,$$

where the two coefficients on the right do not vanish, the equation $Q = 0$ defines a relation R between s and t in which s has m correspondents $\{t_k\}_1^m$ and t has m correspondents $\{s_k\}_1^m$. Since the principal part of Q is

$$P(sa + tb) = P(a) \prod (s + t\lambda_k(a, b)),$$

it follows that

$$\{s_k\}_1^m = -t\{\lambda_k\}_1^m + o(t)$$

and

$$(6) \quad \{t_k\}_1^m = -s\{\lambda_k^{-1}\}_1^m + o(s)$$

as $t \rightarrow \infty$ or $s \rightarrow \infty$ respectively. Now assume that R contains a pair $(s, t) = (s_o, t_o)$ which satisfies (4). Then s_o and t_o are on the same side of the real axis and, since P is hyperbolic, we only have to consider the case when $\text{Im } t_o \neq 0$. Now move s from s_o to infinity parallel to the imaginary axis and in the direction of $\text{Im } t_o$. Then (6) shows that all of its correspondents go off to infinity in the opposite direction. Since one of the correspondents occupies the place t_o when $s = s_o$, at least one must at some time cross the imaginary axis. This means that R contains a pair (s, t) with $\text{Im } s \neq 0$ and t real, which is impossible. Hence the first part of the theorem is proved. Next suppose that R contains a pair $(s, t) = (s_o, t_o)$ which satisfies (5). Then, moving s along the real axis from $s_o \geq -h$ to $+\infty$, again using (6), we see that R contains a pair $(s, 0)$ with $s > -h$. But this is impossible, so that the theorem is proved. An easy consequence is

Theorem 2. *Let P be hyperbolic a. Then the function h is positive homogeneous and concave:*

$$(7) \quad h(a, sx) = sh(a, x), \quad s \geq 0; \quad h(a, x + y) \geq h(a, x) + h(a, y).$$

In particular, $C = C(a)$ is convex. Further, P is hyperbolic any b in C and $C(b) = C(a)$.

Proof. The first part of (7) is obvious. Let $b \in C$, i.e., $h(a, b) > 0$. Then $P(b) \neq 0$ and Theorem 1 shows that $P(tb + x) \neq 0$ if $\text{Im } t \neq 0$ and $P(sa + b + x) \neq 0$ if $s + h(a, x) \geq 0$. Hence P is hyperbolic b and $h(a, x + b) > h(a, x)$. Now (2) shows that $h(a, y - ta) = h(a, y) - t$. Hence, choosing $t < h(a, y)$, we have $h(a, x + y - ta) > h(a, x)$, i.e. $h(a, x + y) > h(a, x) + t$. Letting t approach $h(a, y)$, we get (7). If $x \in C(a)$, (7) shows that $h(a, tb + x) > 0$ for all $t \geq 0$ and hence also

$$P(b) \prod (t + \lambda_k(b, x)) = P(tb + x) = P(a) \prod \lambda_k(a, tb + x) \neq 0$$

for all $t \geq 0$. Hence $x \in C(b)$ so that $C(a) \subset C(b)$, and, by the symmetry, $C(a) = C(b)$. This finishes the proof.

Remark. That C is convex is an example of a general convexity theorem by S. BOCHNER [2]: A function $f(x + iy)$ which is analytic in a tube $K + iR^n$ [with an open base $K \subset R^n$] is also analytic in the tube based on the convex closure of K . In our case,

$$P(x + iy) = P(x) \prod (1 + i\lambda_k(x, y))$$

does not vanish in $T = C + iR^n$ but vanishes on the boundary of C . Hence

$f = 1/P$ is analytic in T but in no larger tube, so that C is convex. The first part of Theorem 1 says, in particular, that f is analytic in T if P is hyperbolic a .

Remark. It follows from (4) that

$$(8) \quad \min_k \operatorname{Im} \lambda_k(x + ib) > 0$$

or, equivalently,

$$\min_k \operatorname{Re} \lambda_k(b + ix) > 0$$

for all real x and all b in C . Using this, we can obtain an interesting complement to (7). Let us label the numbers $\lambda_k(x) = \lambda_k(a, x)$ so that

$$\lambda_1(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_m(x).$$

Then

$$(9) \quad \lambda_k(x + b) > \lambda_k(x)$$

for all x and all $b \in C$. In the examples 3 and 4 this is a well-known monotonicity property of the eigenvalues of symmetric and hermitian matrices respectively. To prove (9), develop $\varphi(s) = \lambda_k(x + sb)$ in a Puiseux series around s_0 . Since the series is real for real $t = s - s_0$, it contains no fractional powers and hence

$$\varphi(s) = \varphi(s_0) + ct + O(t^2)$$

where c is real. In particular,

$$\operatorname{Im} \varphi(s_0 + it) = ct + O(t^2).$$

But, by (8), the left side has the same sign as t and hence $c \geq 0$. This shows that φ is non-decreasing and, being algebraic and having the form $\lambda_k(b) s + o(s)$ for large s , it is not constant in any interval. Hence (9) follows. The formulas (8) and (9) in a slightly different setting and the above deduction of (9) from (8) are due to LAX [3]. See also WEINBERGER [4].

3. Edge and lineality. We have seen that $C = C(P, a)$ is a convex cone not equal to the entire space. We shall be interested in its edge EC , i.e. the set of all x such that $C + x = C$. It is clear that EC is a linear subspace of R^n . If $m = 1$, it is the plane $P = 0$, and in the examples 1 to 4 it reduces to 0. We shall see that it coincides with the lineality LP of P which is defined as the set of all x such that $P(tx + y) = P(y)$ for all t and y . It is a linear subspace of R^n and P is constant on the translates of LP . When $LP = 0$ we say that P is complete. This means that P depends on all variables. Differentiations with respect to y show that if R is any linear combination of derivatives of P , then $LP \subset LR$.

Theorem 3. *If P is hyperbolic a , then $EC = LP$ and x belongs to LP if and only if*

$$(10) \quad \lambda_1(a, x) = \cdots = \lambda_m(a, x) = 0.$$

Proof. If $x \in LP$, then, in particular, $P(ta + x) = t^m P(a)$ for all t and hence x satisfies (10). Conversely, assume that x satisfies (10). Then, by Theorem 2,

$$h(a, b + tx) = \min \lambda_k(a, b + tx) \geq h(a, b) > 0$$

for all t if $b \in C$. Hence

$$P(b) \prod (1 + t\lambda_k(b, x)) = P(b + tx) = P(a) \prod \lambda_k(a, b + tx) \neq 0$$

for all t . This means that

$$\lambda_1(b, x) = \cdots = \lambda_m(b, x) = 0$$

and hence $P(b + tx) = P(b)$ for all b in C and all t . Since C is open, this holds for all b and t and hence $x \in LP$. This shows that $x \in LP$ is equivalent to (10). Since $sC = C$ when $s > 0$, we have $x \in EC$ if and only if $C + sx = C$ for all s . Hence if $x \in EC$ then $\min \lambda_k(a, a + sx) > 0$ so that

$$P(a + sx) = P(a) \prod \lambda_k(a, a + sx) \neq 0$$

for all s . This means that x satisfies (10) so that $x \in LP$. Conversely, if $x \in LP$ then $P(ta + y + sx) = P(ta + y)$ for all s, t and y . Hence if $y \in C$ then $y + sx \in C$ for all real s , i.e. $x \in EC$. This finishes the proof.

4. The inequality. We can now prove the announced inequality. Let

$$M(x^1, \cdots, x^m), \quad x^k = \{x_j^k\}_1^n,$$

be the completely polarized form of the polynomial P . It is uniquely defined by being linear in each argument, invariant under permutations and satisfying $M(x, \cdots, x) = P(x)$. Explicitly, it can be obtained by differentiation:

$$M = \frac{1}{m!} \prod_k \left(\sum_j x_j^k \frac{\partial}{\partial x_j} \right) P(x).$$

First we shall give a more precise form of Lemma 1.

Theorem 4. If P is hyperbolic a , $m > 1$ and $b \in C(a)$, then

$$Q(x) = M(b, x, \cdots, x)$$

is hyperbolic a and $C(Q, a) \supset C(P, a)$.

Proof. Since P is hyperbolic b and $mQ = (\sum b_k \partial/\partial x_k)P$, Lemma 1 shows that Q is hyperbolic b . Further, since the zeros of

$$mQ(sb + x) = \frac{d}{ds} P(sb + x) = 0$$

separate those of $P(sb + x) = 0$, we get $C(Q, b) \supset C(P, b)$. Hence, by Theorem 2 applied to Q , the polynomial Q is hyperbolic a and the same theorem shows that

$$C(Q, a) = C(Q, b), \quad C(P, a) = C(P, b).$$

The next theorem contains the announced inequality.

Theorem 5. *Let P be hyperbolic a , let $P(a) > 0$ and $m > 1$ and let M be the completely polarized form of P . If $x^1, \dots, x^m \in C = C(P, a)$, then*

$$M(x^1, \dots, x^m) \geq P(x^1)^{1/m} \dots P(x^m)^{1/m}$$

with equality if and only if $x^1, \dots, x^m$ are pairwise proportional modulo LP .

Proof. Put

$$Q(x) = M(b, x, \dots, x).$$

We shall first prove the special case

$$(11) \quad Q(x) \geq P(b)^{1/m} P(x)^{m-1/m}, \quad x, b \in C.$$

Since P is homogeneous and $P(a) > 0$, it is no restriction to assume that $P(x) = 1$. Then

$$\prod_1^m (t + \lambda_k) = P(tx + b) = t^m + mQt^{m-1} + \dots + P(b)$$

where, by virtue of Theorem 2, the numbers $\lambda_k = \lambda_k(x, b)$ are positive. Since

$$P(b) = \lambda_1 \dots \lambda_m, \quad Q(x) = \frac{1}{m} (\lambda_1 + \dots + \lambda_m),$$

(11) follows, and equality holds if and only if

$$\lambda_1 = \dots = \lambda_m = \lambda.$$

But this means that for all t

$$P((t + \lambda)x + b - \lambda x) = P(tx + b) = (t + \lambda)^m,$$

so that, by Theorem 3, $b - \lambda x \in LP$. This proves our special case. To prove the general case, we proceed by induction. When $m = 2$, (11) is already the desired inequality. Assuming that it holds for polynomials of degree $m - 1$ and applying it to Q with $b = x^1$, which by Theorem 4 is hyperbolic a , we get

$$M \geq Q(x^2)^{1/m-1} \dots Q(x^m)^{1/m-1}$$

and, applying (11) to each factor on the right,

$$Q(x^k) \geq P(x^1)^{1/m} P(x^k)^{m-1/m}, \quad k > 1.$$

Inserting, we get the desired result. Equality holds everywhere if and only if all x^k are pairwise proportional modulo $LP \subset LQ$.

We leave to the reader to prove that the theorem implies that the hyper-surfaces $P(x) = c > 0$, $x \in C$, are convex and that they are strictly convex modulo LP .

Remark. Let K be the subset of C where $P(x) > c$. That K is convex is again an example of BOCHNER's theorem. In fact,

$$|P(x + iy)| = P(x) \prod |(1 + i\lambda_k(x, y))| \geq P(x) > c$$

in $T = K + iR^n$ and $P(x) = c$ on the boundary of K . Hence $f = 1/P - c$ is analytic in T but in no larger tube so that K is convex.

REFERENCES

- [1] L. GÅRDING, Linear hyperbolic partial differential equations with constant coefficients. *Acta Math.* **85** (1951), 1–62.
- [2] S. BOCHNER & T. W. MARTIN, *Several complex variables*. Princeton, 1948.
- [3] P. D. LAX, Differential equations, difference equations and matrix theory. *Comm. Pure and Appl. Math.* **6** (1958), 175–194.
- [4] H. F. WEINBERGER, Remarks on the preceding paper of Lax. *Op. cit.*, 195–196.

University of Lund