A note on l^p-norms

Ehrhard Behrends, Olga Katkova, Anna Vishnyakova

Abstract. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be nonnegative numbers. If the l^p-norm of $(x_1, \ldots, x_n)$ coincides with that of $(y_1, \ldots, y_n)$ for n different values of $p > 0$, then the x_i are a permutation of the y_i. This generalizes a well-known algebraic result which concerns the exponents $p = 1, \ldots, n$.

Some consequences as well as some generalizations to infinite sequences and to functions are also discussed.

AMS-classification: 46B25;

keywords: l^p-norm, Dirichlet series, sign change, Müntz-Szasz theorem.

1. The problem

Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be real or complex numbers such that $\sum x_i^k = \sum y_i^k$ for $k = 1, \ldots, n$. Then there is a permutation π such that $x_i = y_{\pi(i)}$ for all i. This is a well-known result, it is contained in many books (see e. g. [4], §33).

The idea of the proof is to show that, as a consequence of the assumption, the polynomials $\prod_i (x - x_i)$ and $\prod_i (x - y_i)$ have the same coefficients and thus the same roots. This argument implies that it is in fact not necessary that we work with numbers, the x_i, y_i can be elements of any commutative ring without divisors of zero. On the other hand, it is crucial for the idea to work that the exponents k range from 1 to n.

The starting point of the present investigations was the desire to find *an analytic proof* of the result, it was to be hoped that in this way one could get rid of the special family of exponents. An appropriate analytic variant surely had to be of the following form: if $\sum x_i^p$ coincides with $\sum y_i^p$ for "sufficiently many" exponents p, then the y_i are a permutation of the x_i. However, a moment's reflection shows that this is not to be expected: if all p are even numbers, then the sums cannot contain information concerning the sign of the numbers under consideration. Also, in order to be able to work with non-integer exponents, one has to restrict oneself to nonnegative x_i, y_i, otherwise the expressions x_i^p, y_i^p have no well-defined meaning.

So we arrived at the following

Problem: *Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be nonnegative real numbers such that $\sum x_i^p = \sum y_i^p$ holds for n different exponents $p > 0$. Does it follow that the x_i are a permutation of the y_i?*

> The natural *translation into the language of functional analysis* reads as follows: Do n different l^p-norms on $\mathbb{R}^n$ suffice to characterize vectors with nonnegative components up to rearrangement?

We will show in this note that *the answer is in the positive.* The next section contains some preparations, in section 3 we present the proof, generalizations and some consequences, and finally, in section 4, we indicate an open problem.

2. Zeroes and sign changes

Let $a_1, \ldots, a_r$ and $\lambda_1, \ldots, \lambda_r$ be real numbers such that all a_ρ are nonzero and the λ_ρ are strictly positive and increasing. For $x \in \mathbb{R}$ we define

$$f(x) := f_{a_1,\ldots,a_r,\lambda_1,\ldots,\lambda_r}(x) := a_1 \exp(-\lambda_1 x) + \cdots + a_r \exp(-\lambda_r x),$$

and we want to bound the number of positive zeroes of this Dirichlet series.

Crucial for our investigations will be the following

Lemma 2.1 *There are at most as many $x > 0$ such that $f(x) = 0$ as there are sign changes of the sequence $a_1, a_1 + a_2, \cdots, a_1 + \cdots + a_r$.*

> Recall that, for a real sequence $b_1, \ldots, b_r$, the *number of sign changes* is defined as follows: if all b_ρ are nonzero, then the number of sign changes is the cardinality of the indices $\rho \in \{1, \ldots, r-1\}$ such that $b_\rho b_{\rho+1} < 0$; for arbitrary sequences first cancel the b_ρ which are zero and then apply the first rule.

Proof. The result is not new, it is contained e. g. in part 5, chapter 1 of [2] where it is attributed to Laguerre (see [1]). Nevertheless, for the sake of completeness, we include the argument.

Let $\Phi : [0,\infty[\to \mathbb{R}$ be a bounded piecewise continuous function which is not identically zero. We claim that the number of zeroes of its Laplace transform $g_\phi : [0,\infty[\to \mathbb{R}$,

$$g_\Phi(x) := \int_0^\infty e^{-xt}\Phi(t)dt,$$

can be bounded by the number s of sign changes of Φ. (By definition, Φ has s sign changes if there are $0 < t_1 < t_2 < \cdots < t_s$ such that the function $(t-t_1)\cdots(t-t_s)\Phi(t)$ doesn't change sign.)

This is proved by induction on s. The case $s = 0$ is obvious, we assume the assertion to be established for a fixed s, and we consider a Φ having $s+1$ sign changes. Then $\Psi(t) := (t-t_0)\Phi(t)$ will have only s sign changes for a suitably chosen $t_0 > 0$. Our hypothesis implies that g_Ψ vanishes at not more than s points. And since $e^{t_0 x}g_\Psi(x)$ is the derivative of $-e^{t_0 x}g_\Phi(x)$, it follows with Rolle's theorem from elementary calculus that g_Φ cannot have more than $s+1$ zeroes.

The lemma is an immediate consequence, we simply consider

$$\Phi(t) := \sum a_i \chi_{[\lambda_i,\infty[},$$

where $\chi_{[\lambda_i,\infty[}$ denotes the characteristic function associated with the interval $[\lambda_i,\infty[$. Then sign changes of Φ correspond to sign changes in the sequence $a_1, a_1+a_2, \cdots, a_1+\cdots+a_r$, and $g_\phi(x) = f(x)/x$ for all x. □

3. The solution of the problem and some consequences

Here is our main result:

Theorem 3.1 *Let $g_1,\ldots,g_n$ be strictly positive numbers and $0 < p_1 < \cdots < p_n$. Then, for real x_i, y_i such that $0 \le x_1 \le \cdots \le x_n$ and $0 \le y_1 \le \cdots \le y_n$, the set of equations*

$$\sum_i g_i x_i^{p_k} = \sum_i g_i y_i^{p_k}, k = 1,\ldots,n$$

implies that $x_i = y_i$ for $i = 1,\ldots,n$.

Proof. Let us assume that, without loss of generality, the x_i, y_i are smaller than one and that $x_1 > 0$; $l < n$ denotes the smallest index i such that $y_i > 0$. We put $\alpha_i := \log x_i, \beta_i := \log y_i$,

$$f(x) := g_1 \exp(\alpha_1 x) + \cdots + g_n \exp(\alpha_n x) - g_l \exp(\beta_l x) - \cdots - g_n \exp(\beta_n x)$$

for $x > 0$. This function is identically zero iff $l = 1$ and $x_i = y_i$ holds for all i, and thus our assertion is equivalent with the statement that f has at most $n - 1$ zeroes if it is not the zero function.

In order to apply lemma 2.1 we assume that $f \neq 0$, and we rewrite f as

$$f(x) = a_1 \exp(-\lambda_1 x) + \cdots + a_{2n-l+1} \exp(-\lambda_{2n-l+1} x)$$

with positive *increasing* λ_i. In the sequence $a_1, \ldots, a_{2n-l+1}$ the $g_n, \ldots, g_1$ as well as the $-g_n, \ldots, -g_l$ are contained as subsequences, the concrete form will depend on the relative order of the x_i, y_i (for example, if $x_n > y_n$, then $\lambda_1 = -\alpha_n$ and $a_1 = g_n$).
In view of lemma 2.1 we have to show:

> Regardless of how one mixes the sequences $g_n, \ldots, g_1$ and $-g_n, \ldots, -g_l$ to get a sequence $a_1, \ldots, a_{2n+l-1}$, the number of sign changes in $a_1, a_1 + a_2, \ldots, a_1 + \cdots + a_{2n+l-1}$ is bounded by $n - 1$.

The proof is by induction on n. The statement clearly holds with $n = 1$, and if it is true for some $n - 1$ we argue as follows. Suppose that the sequence $a_1, \ldots$ begins with g_n and that $-g_n$ occurs at the r'th position. Denote by $a'_1, \ldots, a'_{2n-l-1}$ the sequence which consists of the elements of $a_1, \ldots$ without a_1 and a_r. By hypothesis, there are at most $n - 2$ sign changes in $a'_1, a'_1 + a'_2, \ldots$, and this implies that $a_1, a_1 + a_2, \ldots$ has at most $n - 2 + 1$ sign changes. (The additional sign change for the original sequence can occur after the r'th position.)
If $a_1 = -g_n$, we argue similarly. □

Consequences:

1. What we have called a problem in section 1 is in fact a true statement. One only has to consider the case $g_1 = \cdots = g_n = 1$ in the theorem.

2. The analysis of the proof shows that one can assert more. If one knows for some $0 < x_1 \leq \cdots \leq x_n$ the relative order with respect to certain $0 < y_1 \leq \cdots \leq y_n$ (and nothing more), there is an $m \leq n$ such that the following holds: Whenever, for any positive $g_1, \cdots, g_n$, $\sum g_i x_i^p = \sum g_i y_i^p$

holds for m different values of p, then $x_i = y_i$ for all i. The number $m = n$ corresponds to the general case of no particular information, but m can be considerably smaller. (An extreme case is the trivial situation where all y_j dominate all x_i, then – for arbitrary g – even $m = 1$ suffices.)

3. The theorem provides an analytical proof of the original algebraic assertion, at least for the case of real numbers. (If the x_i and the y_i are given in increasing order, then $\sum_i x_i^k = \sum_i y_i^k, k = 1, \ldots, n$ implies that

$$\sum_i (x_i + l)^k = \sum_i (y_i + l)^k, k = 1, \ldots, r$$

for $l = 1, 2, \ldots$; this can easily be proved by induction on l. Now it suffices to choose l so large that all $x_i + l, y_i + l$ are positive and to apply the theorem.)

4. For different strictly positive x, y we denote by $\mu_{x,y}$ the measure $\delta_x - \delta_y$ (with δ_x=the Dirac measure associated with x). Given n such measures $\mu_{x_1,y_1}, \ldots, \mu_{x_n,y_n}$ we can consider them as linear functionals on the space X of real valued functions on $\{x_1, y_1, \ldots, x_n, y_n\}$ (we assume that this set contains $2n$ points).

Now let $\mu := g_1\mu_{x_1,y_1} + \cdots + g_n\mu_{x_n,y_n}$ be any measure in the cone generated by the μ_{x_i,y_i}. The annihilator V of μ is a $(2n-1)$-dimensional subspace of the $2n$-dimensional space X. By our theorem, V can only contain "few" functions of the type $x \mapsto x^p$, namely at most $n - 1$ of them. Thus measures of this type are in a sense "anti-orthogonal" to functions of the type $x \mapsto x^p$.

Let's now turn to some **generalizations**, at first we pass from finite to infinite sequences. Here the natural setting is as follows:

Problem: *Let $x_1 \geq x_2 \geq \cdots \geq 0$ and $y_1 \geq y_2 \geq \cdots \geq 0$ be sequences. Does it follow from $\sum x_i^p = \sum y_i^p$ for "sufficiently many" exponents p that $x_i = y_i$ for all i?*

This is not to be expected if the sums are infinite, and therefore we restrict ourselves to sequences which lie in certain l^p-spaces with $p > 0$. It is clear that finitely many p will not suffice, here is a positive result:

Proposition 3.2 *Suppose that $(x_i), (y_i) \in l^{p_0}$ for some $p_0 > 0$ and that $x_1 \geq x_2 \geq \cdots \geq 0, y_1 \geq y_2 \geq \cdots \geq 0$. Then, if*

$$\sum g_i x_i^{p_r} = \sum g_i y_i^{p_r} \tag{3.1}$$

holds for any bounded sequence g_i of strictly positive numbers and infinitely many different p_r such that $\inf p_r > p_0$, then $x_i = y_i$ for all i.

Proof. We consider two cases.
Case 1: The (p_r) have a finite accumulation point q.
Let $f : \{\text{Re } z > p_0\} \to \mathbb{C}$ be the function $f(z) = \sum g_i x_i^z - \sum g_i y_i^z$. f is analytic on its domain, and the zeroes have q as an accumulation point. Thus f is identically zero, and it follows easily that then $x_i = y_i$ for all i.
Case 2: The (p_r) have ∞ as an accumulation point.
For simplicity let's assume that $p_r \to \infty$. We suppose that the assertion does *not* hold. Without loss of generality we have $x_1 > y_1$, and after suitable normalization we also may assume that $x_1 = 1$. If we now let p_r tend to infinitiy in (3.1), then the left hand side tends to g_1 whereas the right hand side tends to 0; this contradiction establishes the result. □

A *second natural generalization* deals with functions instead of sequences. For simplicity we will only consider continuous functions $f, h : [0,1] \to [0,\infty[$, and we want to conclude from $\int_0^1 f^p(x)dx = \int_0^1 h^p(x)dx$ for "many" p that f and h are the same functions up to rearrangement. A sufficient condition can easily be found with the help of the Müntz-Szasz theorem (this observation is due to D. Werner):

Proposition 3.3 *Let f, h be as before and assume additionally that they both are increasing and that g is a strictly positive continuous density. Then $f = h$ provided that $\int_0^1 f^{p_r}(x)g(x)dx = \int_0^1 h^{p_r}(x)g(x)dx$ for some sequence $0 < p_1 < p_2 < \cdots$ of exponents which satisfy $\sum 1/p_r = \infty$.*

Proof. We can assume that $\int_0^1 g(x)dx = 1$, and we will regard f and h as $[a,b]$-valued random variables defined on the probability space induced by g on $[0,1]$. Denote by μ resp. ν the induced measures on $[a,b]$. Then

$$\int_a^b t^{p_r} d\mu(t) = \int_a^b t^{p_r} d\nu(t)$$

for all r. But the functions $t \mapsto t^{p_r}$ have a dense linear span in $C[a,b]$ by the Müntz-Szasz theorem (see [3], theorem 15.26), and therefore μ coincides with ν. From this it can easily be concluded that $f = h$. □

4. An open problem

In this note we have concentrated ourselves on the analytic aspects of the problem, and since arbitrary positive exponents p should be taken into account the restriction to nonnegative x_i, y_i was natural. We have already remarked that also in the case of integer exponents some restrictions are necessary if one passes to more general x_i, y_i, but by what number theoretical properties could they be described? More precisely:

> Let $m_1 < \cdots < m_n$ be positive integers and A a subset of the complex plane. We say that $(m_1, \ldots, m_n)$ are *A-admissible* if
>
> $$\sum_i x_i^{m_k} = \sum_i y_i^{m_k}$$
>
> for $k = 1, \ldots, n$ implies that the x_i are a permutation of the y_i; this has to hold for arbitrary $x_1, \ldots, x_n, y_1, y_n \in A$.
>
> What are the A-admissible $(m_1, \ldots, m_n)$ for $A = \mathbb{Z}, \mathbb{R}$, and $\mathbb{C}$?

For the case $A = [0, \infty[$ theorem 3.1 provides a complete answer (then *all* $(m_1, \ldots, m_n)$ are admissible), for other A we only have partial results. In fact, the situation is even more complicated. To illustrate what is meant consider the case $n = 2$. Here $(1, 3)$ surely is not $\mathbb{C}$-admissible since $x_1 + x_2 = x_1^3 + x_2^3 = 0$ is satisfied as soon as $x_1 + x_2 = 0$. On the other hand, if $x_1 + x_2$ does *not* vanish, then $x_1^2 + x_2^2$ can be expressed by $x_1 + x_2$ and $x_1^3 + x_2^3$:

$$x_1^2 + x_2^2 = (x_1 + x_2)^2 - 2\frac{(x_1 + x_2)^3 - (x_1^3 + x_2^3)}{3(x_1 + x_2)}.$$

Consequently, by the result from our introduction, x_1, x_2 will be a permutation of y_1, y_2 if $x_1 + x_2 = y_1 + y_2$ and $x_1^3 + x_2^3 = y_1^3 + y_2^3$ hold. This motivates to call the $(m_1, \ldots, m_n)$ *nearly A-admissible* if there is a Borel set $\Delta \subset \mathbb{C}^n$ of measure zero such that the desired conclusion holds under the additional assumption that the vector $(\sum_i x_i^{m_k})_{k=1,\ldots,n}$ does not belong to Δ. We have just shown that $(1, 3)$ is nearly $\mathbb{C}$-admissible, and with little more effort it can be demonstrated that there are no further examples in the case $n = 2$ (up to $(m_1, m_2) = (1, 2)$, of course).

Our approach indicates how to treat the problem in general. Fix n and denote by s_k the elementary symmetrical polynomial $x_1^k + \cdots + x_n^k$ for $k = 1, \ldots$. Then the idea is to express the $s_1, \ldots, s_n$ by the $s_{m_1}, \ldots, s_{m_n}$. According to [5] this has been studied already in the twenties by the japanese mathematician Kakeya who has obtained the following surprisingly simple characterization: all $s_1, \ldots, s_n$ are functions of the $s_{m_1}, \ldots, s_{m_n}$ iff the complement of $\{m_1, \ldots, m_n\}$ with respect to $\mathbb{N}$ is an additive semigroup: with r, l different from all m_i also $r + l$ has this property.

REFERENCES

[1] E. N. LAGUERRE. *Sur la règle de Descartes.* C. R. des sciences **XCII** (1881).

[2] G. PÓLYA, G. SZEGÖ. *Aufgaben und Lehrsätze aus der Analysis II.* Springer Verlag, Berlin 1964.

[3] W. RUDIN. *Real and Complex Analysis.* Mc Graw-Hill, 1974 (2nd edition).

[4] B. L. VAN DER WAERDEN. *Algebra.* Springer Verlag, Berlin 1966.

[5] D. ZAGIER. *(personal communication).*

I. Mathematisches Institut, Freie Universität Berlin, Arnimallee 2–6,
D-14 195 Berlin, Germany; e-mail: behrends@math.fu-berlin.de

Dept. of Mechanics and Mathematics, Kharkov State University, 4 Svobody sq.,
310077 Kharkov, Ukraine; e-mail: katkova@ilt.kharkov.ua vishnyakova@ilt.kharkov.ua