On maximizers of a convolution operator in $L_p$-spaces

G. V. Kalachev and S. Yu. Sadov

Abstract. The paper is concerned with convolution operators in $\mathbb{R}^d$, whose kernels are in $L_q$, which act from $L_p$ into $L_s$, where $1/p + 1/q = 1 + 1/s$. It is shown that for $1 < q, p, s < \infty$ there exists a maximizer (a function with $L_p$-norm 1) at which the supremum of the $s$-norm of the convolution is attained. A special analysis is carried out for the cases in which one of the exponents $q, p, or s$ is 1 or $\infty$.

Bibliography: 12 titles.

Keywords: convolution, Young inequality, existence of an extremal function, tight sequence, concentration compactness.

§ 1. Introduction

Let $L_p(\mathbb{R}^d)$ be the space of measurable complex-valued functions with norm

$$\|f\|_p = \left(\int |f(x)|^p \, dx\right)^{1/p}, \quad 1 \leq p < \infty,$$

or with norm $\|f\|_\infty = \sup\{a > 0 : |f(x)| \leq a$ almost everywhere}. Throughout, $p' = (1-1/p)^{-1}$ denotes the conjugate exponent. We consider the convolution operator $K_k : f \mapsto k * f$ with kernel $k \in L_q(\mathbb{R}^d),$

$$K_k f(x) = \int k(x-y) f(y) \, dy.$$ 

We write $L_p$ in place of $L_p(\mathbb{R}^d)$ and $K$ in place of $K_k$ where no ambiguity can arise. The dimension $d$ and the kernel $k$ will be assumed to be fixed in the statements and proofs which follow.

Let $1 \leq p, q, r \leq \infty$. If $k \in L_q$ and if

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2,$$ (1.1)

then the operator $K$ acts boundedly from $L_p$ into $L_{p'}$ and its norm (to be referred to as the $(p,r)$-norm$^1$) can be estimated from above using Young’s inequality: $\|K\|_{p,r} \leq \|k\|_q$. A function $f \in L_p$ is called a maximizer of the convolution operator $K$ with respect to a pair of exponents $(p,r)$ if $\|f\|_p = 1$ and $\|k * f\|_{r'} = \|K\|_{p,r}$.

The main result of this paper is the following theorem on the existence of a maximizer.

$^1$Here we take account of the fact that $\|K\|_{p,r}$ gives an extremum to the symmetric bilinear form $\|K\|_{p,r} = \|K\|_{r,p} = \sup \int k(x+y) f(x) g(y) \, dx \, dy \mid \|f\|_p = \|g\|_r = 1$.

AMS 2010 Mathematics Subject Classification. Primary 44A35, 46E30, 49J99.

© 2019 Russian Academy of Sciences (DoM), London Mathematical Society, Turpion Ltd.
Theorem 1.1. Let $1 < p, q, r < \infty$ and let (1.1) hold. Then for any kernel $k \in L_q$ there exists a maximizer of the operator $K$ with respect to the pair of exponents $(p, r)$.

Pearson [1] proved that the operator $K_k$ has a maximizer under the assumption that the function $k$ is radially symmetric, nonnegative and decreasing away from the origin; moreover, excessive integrability is required: $k \in L_{q+\varepsilon} \cap L_{q-\varepsilon}$.

Theorem 1.1 originates from the observation that maximizers of convolution operators exist in numerical calculations aimed at evaluating the norms of the Laplace transform from $L_p$ into $L_{p'}$, $1 < p < 2$, on the positive half-axis (calculating the exact constants in the Hardy inequality [2]); this is equivalent to evaluating the corresponding norms of convolution operators with nonsymmetric kernels $k_p(x) = \exp(x/p - e^x)$ on $\mathbb{R}$. For $p = 2$, the norm is known analytically, but no maximizer exists. In turn, the evaluation of the norms of the Laplace transform is motivated by the fact that the exact constants in the Hausdorff-Young inequality for the Fourier transform (which is analogous to the above Hardy inequality) were given by Babenko [3] and Beckner [4]. Refinements of the Hardy inequality are not directly related to the main theme of our investigation and hence are not discussed here.

The problem in proving the existence of a maximizer in this and similar settings lies in the invariance of the problem under a noncompact symmetry group (the translation group on $\mathbb{R}^d$ in our case), which implies that the set of maximizers, if nonempty, is not compact. Starting from some normed maximizing sequence $(f_n)$, $\|f_n\|_p = 1$, $\|Kf_n\|_{r'} \to \|K\|$, and using the Banach-Alaoglu theorem to produce a weakly convergent subsequence $(f_n(m))$, we cannot hope to prove that it strongly converges: for example, a subsequence $(f_n(m))$ can “move away” to infinity by translations and have 0 as its weak limit.

On the other hand, if the original maximizing sequence is replaced by a sequence of shifted functions $\tilde{f}_n = f_n(\cdot - a_n)$ (which is also a maximizing sequence), by choosing the shifts $a_n$ we can hope to obtain a relatively compact sequence $(\tilde{f}_n)$ (in terms of inequalities, this amounts to establishing suitable uniform estimates) so that the weakly convergent subsequence will converge strongly and its limit will be a maximizer.

Pearson’s proof exploits the fact that in the case of a radially symmetric kernel $k$ the functions $\tilde{f}_n$ can also be made radially symmetric due to M. Riesz’s inequality for nondecreasing rearrangements of functions. In general, we can control the localization and centring of the terms of a maximizing sequence intrinsically, by employing functionals analogous to the concentration function and the Lévy dispersion function in probability theory.

Considering Theorem 1.1 in the broader context of the existence of extremal solutions to problems with noncompact symmetry group, we cannot but mention P.-L. Lions’s concentration compactness method (see [5]–[7]). Our proof does not depend on Lions’s theory nor does it exploit an analogue of the concentration compactness lemma (Lemma I.1 in [5], Lemma 5.1 in [7]); variants of ‘vanishing’ and ‘dichotomy’ are implicitly eliminated using other means. At the same time, some common elements become apparent: for example, in Lions’ theory the inequality in Lemma 3.1 is interpreted as a subadditivity property (of crucial importance) of the parameterized objective function $I_\lambda$. 
The paper is organized as follows. The proof of Theorem 1.1 is given in §§2–4. In §2 we introduce the necessary terminology and establish the equivalence of Theorem 1.1 to Proposition 2.1 (at first sight the latter seems weaker). We describe the proof of Proposition 2.1 with reference to the properties employed at concrete steps. The detailed statements and proofs of the required properties, as well as intermediate auxiliary results, are given in §3 (uniform estimates for the size of near-supports of the terms of the maximizing sequence) and §4 (compactness properties and selection of shifts).

Theorem 1.1 excludes the cases when at least one of the exponents $p, q, r$ in the Young inequality is 1 or $\infty$. The analysis of these cases is given in §5.

§2. Preliminary considerations and the scheme of the proof

The exponents $q, p$ and $r$ and the convolution kernel $k \in L_q$ are assumed to be fixed.

By the definition of the norm of the operator $K$, for any $\varepsilon > 0$ there exists a function $f_\varepsilon \in L_p$ such that $\|f_\varepsilon\|_p = 1$ and $\|K f_\varepsilon\|_{r'} > \|K\|_{p,r}(1 - \varepsilon)$. Any such function will be referred to as an $\varepsilon$-maximizer.

Definition 2.1. A sequence $(f_n)$ of functions of unit norm in $L_p$ is said to be a maximizing sequence for the operator $K$ if $\|K f_n\|_{r'} \to \|K\|_{p,r}$ as $n \to \infty$.

A few trivial but important properties are worth mentioning.

(1) If a function $f$ is an $\varepsilon$-maximizer, then $f$ is also an $\varepsilon_1$-maximizer for any $\varepsilon_1 > \varepsilon$.

(2) The set of $\varepsilon$-maximizers is shift-invariant (since convolution operators commute with shifts).

(3) Any subsequence of a maximizing sequence is itself a maximizing sequence. The next result is a clear corollary of Theorem 1.1.

Proposition 2.1. Let $1 < p, q, r < \infty$ and let (1.1) hold. Then for any kernel $k \in L_q$ there exists a maximizing sequence of functions $f_n \in L_p$ for which the $L_{r'}$-limit $h = \lim_{n \to \infty} K f_n \in L_{r'}$ exists. Moreover, $\|h\|_{r'} = \|K\|_{p,r}$.

We shall show below that in fact Theorem 1.1 and Proposition 2.1 are equivalent.

Lemma 2.1. Theorem 1.1 follows from Proposition 2.1.

Proof. Let $h$ be the function whose existence is asserted in Proposition 2.1. Let $g \in L_{r'}$ with $\|g\|_{r'} = 1$ and $(g, h) = \|K\|_{p,r}$. Then $g$ is a maximizer for the operator $K': L_r \to L_{p'}$ dual to $K$. Indeed,

$$\|K'g\|_{r'} \geq \sup_n \| (K'g, f_n) \| = \sup_n \| (g, K f_n) \| \geq |(g, h)| = \|K'\|_{r,p}. $$

The operator $K'$ is the convolution operator with kernel $\tilde{k}(x) = k(-x)$. Changing the roles of $K$ and $K'$ we obtain the result of Theorem 1.1. The lemma is proved.

It will be more convenient for us to prove Proposition 2.1. We shall recall some concepts which describe the concentration of functions and sequences of functions.

Throughout, we shall use the standard inner product and norm in the space $\mathbb{R}^d$. In Definition 2.4 we shall assume that $\{e_1, \ldots, e_d\}$ is a fixed orthonormal basis in $\mathbb{R}^d$. 
Definition 2.2. Given a function $f \in L^p$ and a unit vector $v$, the $\delta$-diameter of the function $f$ of order $p$ in the direction $v$ is defined by

$$D_{\delta,v}^p(f) = \inf_{b>a} \left\{ b-a \left| \int_{a<(x,v)<b} |f(x)|^p \, dx \geq \|f\|_p^p - \delta \right. \right\}.$$ 

In particular, if $\delta \geq \|f\|_p^p$, then $D_{\delta,v}^p(f) = 0$.

Remark 2.1. In essence, the $\delta$-diameter of Definition 2.2 is the \textit{Lévy dispersion function} (see §1.1, Appendix 4 in the Russian translation of [8]). More precisely, given a function $f \in L^p(\mathbb{R}^d)$ with $p$-norm one and a unit vector $v \in \mathbb{R}^d$, we have the distribution function (in the sense of probability theory)

$$F(t) = \int_{(v,x)<t} |f(x)|^p \, dx.$$ 

The Lévy concentration function

$$Q_F(\lambda) = \sup_{t \in \mathbb{R}} (F(t + \lambda) - F(t))$$

corresponds to this function. The inverse function is the \textit{dispersion function for the measure $dF$}; in our notation it is

$$D_{\delta,v}^p(f) = \inf_{Q_F(\lambda) \geq 1 - \delta} \lambda.$$

Definition 2.3. Let $\delta < \|f\|_p$. Then there exist $a$ and $b$ such that $b-a = D_{\delta,v}^p(f)$ and

$$\int_{a<(x,v)<b} |f(x)|^p \, dx = \|f\|_p^p - \delta.$$ 

The interval $[a, b]$ will be referred to as a $\delta$-near-support of the function $f$ of order $p$ in the direction $v$ (we write $\text{supp}_{\delta,v}^p(f)$).

Remark 2.2. We do not assert that the $\delta$-near-support is unique (cf. [8], §1.1.2 in the Russian translation). In the case of nonuniqueness, $\text{supp}_{\delta,v}^p(f)$ means an arbitrarily chosen $\delta$-near-support (which is fixed throughout a concrete argument).

Definition 2.4. A function $f$ is said to be $\delta$-near-centred of order $p$ in a direction $v$ if $\text{supp}_{\delta,v}^p(f) = [a, b]$, $a < 0 < b$, and

$$\int_{a<(x,v)<0} |f(x)|^p \, dx = \int_{0<(x,v)<b} |f(x)|^p \, dx = \frac{\|f\|_p^p - \delta}{2}.$$ 

A function $f$ is $\delta$-near-centred of order $p$ if $f$ is $\delta$-near-centred of order $p$ in the direction $e_j$ for all $j = 1, \ldots, d$.

Clearly, any function can be made $\delta$-near-centred by an appropriate shift. However, different shifts may be required for different $\delta$.

The operator of shift by a vector $a$ will be denoted by $T_a$; that is, $T_a f(x) = f(x-a)$.

Definition 2.5. A sequence of functions $f_n \in L_p$ will be called relatively tight if, for any $\delta > 0$,

$$\sup_n \sup_{\|v\|=1} D_{\delta,v}^p(f_n) < \infty.$$
A sequence of functions \( f_n \in L_p \) will be called \textit{tight} if, for any \( \delta > 0 \), it is \( \delta \)-\textit{near-compactly supported} (of order \( p \)); this means that there exists a cube \( Q \) in \( \mathbb{R}^d \) with edges parallel to the coordinate axes such that

\[
\int_Q |f_n(x)|^p \, dx > \| f_n \|^p_p - \delta.
\]

\textbf{Remark 2.3.} A sequence \( f_n \in L_p \) is \textit{tight} in the sense of Definition 2.5 if and only if the sequence of measures \( |f_n(x)|^p \, dx \) is \textit{uniformly dense} in accordance with the terminology adopted in measure theory (for example, see [9], vol. 2, §8.6).

It turns out (see Lemma 4.7) that a relatively tight sequence is tight if the \( \delta \)-near-compactly-supported condition holds for some single sufficiently small \( \delta \).

We are ready to present a high-level structure of the proof of Proposition 2.1. We now introduce the following classes of sequences of \( L_p \)-functions defined in terms of the imposed constraints.

1. The class \textit{Max} comprises all maximizing sequences (for the convolution operator \( K: L_p \to L_{r'} \)).
2. The class \textit{RTgt} comprises all relatively tight sequences.
3. The class \textit{Tgt} consists of all tight sequences.
4. The class \textit{WCvg} consists of all weakly convergent sequences.
5. The class \textit{LCvg} comprises all locally convergent sequences; that is, the sequences that converge with respect to the \( L_p \)-norm on any cube in \( \mathbb{R}^d \).
6. The class \textit{Cvg} comprises all \( L_p \)-convergent sequences.

Note that a subsequence of a sequence that belongs to any of these classes also belongs to that class.

\textbf{Scheme of the proof of Proposition 2.1.}

1. We start with an arbitrary ‘seed’ sequence \( (f_n) \in \text{Max} \).
2. We take some \( \delta_0 < 1/3 \) and find shift vectors \( a_n \) such that the functions \( g_n = T_{a_n}(f_n) \) are \( \delta_0 \)-near-centred.
3. Note that \textit{Max} \( \subset \text{RTgt} \) (Corollary 3.2). Clearly, \( (g_n) \in \text{Max} \). Therefore, \( (g_n) \in \text{RTgt} \).
4. By Lemma 4.6 we have \( (g_n) \in \text{Tgt} \).
5. The sequence \( (g_n) \) is \( L_p \)-bounded, and hence, since \( L_p \) is reflexive, there exists a weakly convergent subsequence (which we again denote by \( (\tilde{g}_m) \)).
6. We set \( \tilde{h}_m = K\tilde{g}_m \). Since \( (\tilde{g}_m) \in \text{WCvg} \), by Corollary 4.1 we conclude that \( (\tilde{h}_m) \in \text{LCvg} \).
7. We also have \( (Kg_n) \in \text{Tgt} \) by Lemma 4.3. As a result, \( (\tilde{h}_m) \in \text{LCvg} \cap \text{Tgt} \).
8. An appeal to Lemma 4.8 shows that \( (\tilde{h}_m) \in \text{Cvg} \). This concludes the proof of Proposition 2.1.

In the next two sections we prove the results underlying the steps of our construction.

The most lengthy is the proof of the result used in Step 3 (Lemma 3.7 and its Corollary 3.2). This proof, which is subdivided into short steps, is given separately in §3. The remaining lemmas required for the construction of a maximizer are given in §4.
§ 3. An estimate for the \( \delta \)-diameters of near-maximizers

The main technical result in this section is Lemma 3.7; the conceptual conclusion is Corollary 3.2. We approach the proof of Lemma 3.7 through a chain of preparatory results, in which Lemma 3.5 is a key step. The indicator (characteristic) function of a set \( \Omega \) will be denoted by \( I_\Omega \); if \( \Omega \) is a set defined by a property (predicate) \( P \), then the indicator function is written \( I_P \).

**Lemma 3.1.** If \( \gamma > 1 \), \( \lambda \in (0, 1/2) \), \( u \in [\lambda, 1 - \lambda] \), then
\[
 u^\gamma + (1 - u)^\gamma \leq 1 - \kappa,
\]
where \( \kappa = \kappa(\lambda, \gamma) = 2\lambda(1 - 2^{1 - \gamma}) > 0 \).

**Proof.** For \( \gamma > 1 \) the function \( h(u) = u^\gamma + (1 - u)^\gamma \) is convex and symmetric about \( u = 1/2 \). It can be assumed without loss of generality that \( u \in [\lambda, 1/2] \). Convexity implies that
\[
 h(u) = h\left( (1 - 2u) \cdot 0 + 2u \cdot \frac{1}{2} \right) \leq (1 - 2u)h(0) + 2uh\left( \frac{1}{2} \right)
= (1 - 2u) + 2u \cdot 2\left( \frac{1}{2} \right)^\gamma = 1 - 2u(1 - 2^{1 - \gamma}) \leq 1 - 2\lambda(1 - 2^{1 - \gamma}).
\]
The lemma is proved.

**Lemma 3.2.** Let \( \Omega \) be a measure space, \( \gamma > 1 \) and \( 0 < \lambda < 1/2 \). Suppose that \( g \in L_1(\Omega) \), \( \|g\|_1 = 1 \), is expressed as \( g = g_1 + g_2 \), where \( \|g_i\|_1 \geq \lambda \), \( i = 1, 2 \), and \( g_1g_2 = 0 \). Then
\[
 \|g_1\|^\gamma_1 + \|g_2\|^\gamma_1 \leq 1 - \kappa,
\]
where \( \kappa = \kappa(\lambda, \gamma) \) is the same as in Lemma 3.1.

**Proof.** Since \( g_1g_2 = 0 \), we have \( \|g_1\|_1 + \|g_2\|_1 = 1 \) and it remains to apply Lemma 3.1 with \( u = \|g_1\|_1 \). The lemma is proved.

**Lemma 3.3.** Let \( R > 0 \) and \( f \in L_1([-R, R]) \). Then for any \( a < R \) there exists \( t_0 \), \( |t_0| \leq R - a \), such that
\[
 \frac{1}{2a} \int_{|t-t_0|\leq a} |f(t)| \, dt \leq \frac{1}{R} \|f\|_1.
\]

**Remark 3.1.** Replacing the numerator 1 on the right by an appropriate constant, we can swap the quantifiers (\( \exists t_0 \forall a \)) in the conclusion of the lemma. This stronger result, which follows from the Hardy-Littlewood maximal inequality, is not required in our paper.

**Proof of Lemma 3.3.** We assume that \( a \leq R/2 \) (otherwise the inequality is a tautology). The function \( h(t) = \int_{|x-t|\leq a} |f(x)| \, dx \), defined for \( |t| \leq R - a \), is continuous and
\[
 2(R - a) \min h(t) \leq \int_{|t|\leq R-a} h(t) \, dt \leq 2a \|f\|_1.
\]
It suffices to take \( t_0 \) for which \( h(t_0) = \min h(t) \) and note that \( 2(R - a) \geq R \). The lemma is proved.
Lemma 3.4. Let \( f \in L_p(\mathbb{R}^d) \). Given \( R \) and \( a \), where \( R > a > 0 \), a unit vector \( v \in \mathbb{R}^d \), and \( c \in \mathbb{R} \), there exists \( t_0 \in [c - (R - a), c + (R - a)] \) such that
\[
\frac{1}{2a} \int_{|v(x) - t_0| \leq a} |f(x)|^p \, dx \leq \frac{1}{R} \int_{|v(x) - c| \leq R} |f(x)|^p \, dx.
\]

Proof. We can assume that \( v = (1, 0, \ldots, 0) \) and \( c = 0 \). The required result follows by applying Lemma 3.3 to the univariate function \( x_1 \mapsto \int |f(x)|^p \, dx_2 \ldots dx_d \) on the interval \( x_1 \in [-R, R] \). The lemma is proved.

Definition 3.1. Let \( A \) be a mapping from \( L \) into \( \tilde{L} \), where \( L \) and \( \tilde{L} \) are some spaces of measurable functions on \( \mathbb{R}^d \). Suppose that \( a > 0 \) and a unit vector \( v \in \mathbb{R}^d \) are given. The mapping \( A \) is an \( a \)-expander of the support in the direction \( v \) if the condition \( f(x) = 0 \) almost everywhere for \( t_1 < (x, v) < t_2 \), where \( -\infty \leq t_1 < t_2 \leq +\infty \), implies that \( Af(x) = 0 \) almost everywhere for \( t_1 + a < (x, v) < t_2 - a \).

In the next result we use the notation and concepts from Definition 2.2.

Lemma 3.5. Let \( A \) be a bounded linear operator from \( L_p(\mathbb{R}^d) \) into \( L_s(\mathbb{R}^d) \), where \( 1 \leq p < s < \infty \). Suppose that \( A \) is an \( a \)-expander of the support in the direction \( v \). Let \( \|f\|_p = 1 \) and \( D = D^p_{\delta, v}(f) \). Then for any \( \beta > 0 \), either (i) \( D \leq 8\beta a \) or (ii) \( D > 8\beta a \) and
\[
\|Af\|_s^\delta < \|A\|_s^\delta (1 - \kappa + \beta^{-\gamma}),
\]
where \( \gamma = s/p \) and \( \kappa = \kappa(\delta/2, \gamma) = \delta(1 - 2^{1-\gamma}) \) is defined as in Lemma 3.1.

Proof. The cases \( \delta \geq 1 \) (when \( D = 0 \)), \( \beta \leq 1 \) (when \( 1 - \kappa + \beta^{-\gamma} \geq 1 \)) and \( D \leq 8\beta a \) are trivial. So, assume that \( \delta < 1 \), \( \beta > 1 \) and \( D > 8\beta a \). Let \( c \) and \( R \) be such that
\[
\|fI_{(v,x)>c+R}\|_p^p = \|fI_{(v,x)<c-R}\|_p^p = \frac{\delta}{2}.
\]
It is clear that \( D \leq 2R \), and hence, \( R > 4a > 2a \). By Lemma 3.4 there exists \( t_0 \in [c - (R - 2a), c + (R - 2a)] \) such that
\[
\int_{|v(x) - t_0| \leq 2a} |f(x)|^p \, dx \leq \frac{4a}{R} \int_{|v(x) - c| \leq R} |f(x)|^p \, dx \leq \frac{8a}{D} < \beta^{-1}.
\]
Setting (Figure 1)
\[
f_l(x) = f(x)I_{(v,x)<t_0} \quad \text{and} \quad f_r(x) = f(x)I_{(v,x)\geq t_0},
\]
we have \( f(x) = f_l(x) + f_r(x) \), and moreover, \( f_l(x)f_r(x) = 0 \), \( \|f_l\|_p^p \geq \delta/2 \), and \( \|f_r\|_p^p \geq \delta/2 \). By Lemma 3.2, applied to the functions \( g_1 = |f_l|^p \) and \( g_2 = |f_r|^p \), we have\(^3\) \( \|f_l\|_p^s + \|f_r\|_p^s \leq 1 - \kappa(\delta/2, \gamma) \).

Consider the function
\[
f_m(x) = f(x)I_{(v,x)-t_0<2a}.
\]
Using (3.2) we have \( \|f_m\|_p^s < \beta^{-\gamma} \).

\(^3\)This place in our proof is close to Lions’ approach ‘to prevent the possible splitting of minimizing sequences by keeping them concentrated’ (see §1.2 of [5]) and reveals the role of the subadditivity property, as expressed by Lemma 3.2.
Consider the following subsets of $\mathbb{R}^d$:

$$
\Omega_l = \{x \mid (v, x) < t_0 - a\},
\Omega_m = \{x \mid |(v, x) - t_0| \leq a\},
\Omega_r = \{x \mid (v, x) > t_0 + a\}.
$$

These subsets are pairwise disjoint, and $\Omega_l \cup \Omega_m \cup \Omega_r = \mathbb{R}^d$. We have

$$
Af_r = 0 \quad \text{for} \quad x \in \Omega_l,
Af_l = 0 \quad \text{for} \quad x \in \Omega_r,
A(f - f_m) = 0 \quad \text{for} \quad x \in \Omega_m.
$$

Consequently,

$$
\|Af\|_s^s = \int_{\Omega_l} |Af_l(x)|^s \, dx + \int_{\Omega_r} |Af_r(x)|^s \, dx + \int_{\Omega_m} |Af_m(x)|^s \, dx
\leq \|A\|_s^s (\|f_l\|_p^s + \|f_r\|_p^s + \|f_m\|_p^s) < \|A\|_s^s (1 - \kappa + \beta^{-\gamma}),
$$

proving Lemma 3.5.

**Lemma 3.6.** Suppose that the operator $A$ satisfies the hypotheses of Lemma 3.5. Suppose further that $\|f\|_p = 1$ and $\|Af\|_s^s \geq \|A\|_s^s (1 - \tau)$. Then, for any $\delta > \tau(1 - 2^{1-\gamma})$, the $\delta$-diameter $D = D_{\delta, v}(f)$ satisfies the inequality

$$
D \leq 8a(\kappa - \tau)^{-1/\gamma}, \quad \text{where} \quad \kappa = \delta(1 - 2^{1-\gamma}).
$$

**Proof.** We set $\beta = (\kappa - \tau)^{-1/\gamma}$ and apply Lemma 3.5. Assuming that case (ii) takes place, we have

$$
1 - \tau \leq \frac{\|Af\|_s^s}{\|A\|_s^s} < 1 - \kappa + \beta^{-\gamma} = 1 - \tau,
$$

which is a contradiction. Therefore, case (i) holds and the lemma is proved.

**Lemma 3.7.** For $q > 1$ let $k \in L_q(\mathbb{R}^d)$, and let $K: L_p \to L_{r'}$ be the convolution operator with kernel $k$. Set $\gamma = r'/p > 1$. Also let $\varepsilon > 0$ and $\delta > \varepsilon r'(1 - 2^{1-\gamma})^{-1}$. Suppose that $\rho > 0$ is sufficiently small so that

$$
\varepsilon + \frac{2\rho^{1/q}}{\|K\|_{p, r}} \leq \delta \frac{1 - 2^{1-\gamma}}{r'}.
$$
Then, for any unit vector \( v \in \mathbb{R}^d \) and any \( \varepsilon \)-maximizer \( f \) of the operator \( K \),

\[
D_{\delta,v}(f) \leq c D_{\rho,v}(k), \tag{3.3}
\]

where

\[
c = 4 \left( \delta (1 - 2^{1-\gamma}) - r' \left( \varepsilon + \frac{2\rho^{1/q}}{\|K\|_{p,r}} \right) \right)^{-1/\gamma}.
\]

Proof. We set \( M = \|K\|_{p,r} \) and \( a = D_{\rho,v}(k)/2 \). We can assume without loss of generality that \( \text{supp}_{\rho,v} k = [-a,a] \). Let \( k_{\rho} = kI_{(v,x)|x|<a} \) and let \( K_{\rho} \) be the convolution operator with kernel \( k_{\rho} \). We have \( \|k_{\rho} - k\|_{q} = \rho \), and by Young’s inequality, \( \|K_{\rho} - K\|_{p,r} \leq \rho^{1/q} \). In particular, \( \|K_{\rho}\|_{p,r} \leq M + \rho^{1/q} \).

We fix an \( \varepsilon \)-maximizer \( f \in L_p \) for the operator \( K \). We have \( \|K_{\rho}f\|_{r'} \geq \|Kf\|_{r'} - \|K_{\rho} - K\|_{p,r} \|f\|_p \geq M(1 - \varepsilon) - \rho^{1/q} \), and so,

\[
\frac{\|K_{\rho}f\|_{r'}}{K_{\rho}\|_{p,r}} \geq \frac{M(1 - \varepsilon) - \rho^{1/q}}{M + \rho^{1/q}} > 1 - \varepsilon - \frac{2\rho^{1/q}}{M}.
\]

The operator \( K_{\rho} \) is an \( a \)-expander of the support in the direction \( v \). Applying Lemma 3.6 with \( A = K_{\rho} \) and \( s = r' \), we have \( \|Af\| = \|A\|^s(1 - \tau) \), where

\[
1 - \tau = \left( 1 - \varepsilon - \frac{2\rho^{1/q}}{M} \right)^{r'} > 1 - r' \left( \varepsilon + \frac{2\rho^{1/q}}{M} \right) > 1 - \delta (1 - 2^{1-\gamma})
\]

(here we have used Bernoulli’s inequality and the inequality relating \( \varepsilon, \delta \) and \( \rho \)). The estimate for \( D \) in Lemma 3.6 gives (3.3). The lemma is proved.

Corollary 3.1. Let \( q, k \) and \( K \) be as in Lemma 3.7. Suppose that \( \varepsilon, \delta, \rho \) and \( c \) are related as follows:

\[
\delta = \frac{4r'}{1 - 2^{1-\gamma} \varepsilon}, \quad \rho = (\|K\|_{p,r} \varepsilon)^q, \quad c = 4(\varepsilon r')^{-1/\gamma}. \tag{3.4}
\]

Then (3.3) holds for any unit vector \( v \in \mathbb{R}^d \) and any \( \varepsilon \)-maximizer \( f \) of the operator \( K \).

Corollary 3.2. Any maximizing sequence \( (f_n) \) of \( L_p \)-functions for the convolution operator \( K \) is relatively tight.

Indeed, let \( f_n \) be an \( \varepsilon_n \)-maximizer and let \( \varepsilon_n \to 0 \). Given \( \delta > 0 \), we define \( \varepsilon \) by (3.4) and choose \( n_0 \) in Definition 2.5 from the condition \( \varepsilon_n \leq \varepsilon \) for \( n \geq n_0 \).

§ 4. Lemmas for the construction of a convergent maximizing sequence

As before, we assume that \( 1/q + 1/p + 1/r = 2 \).

4.1. The compactness lemma.

Lemma 4.1. Let \( k \in L_q(\mathbb{R}^d) \) and let \( \chi \in L_{r'} \cap L_\infty(\mathbb{R}^d) \). Then the integral operator

\[
\chi K : f(x) \mapsto \chi(x) (k * f)(x)
\]

with kernel \( \chi(x)k(x-y) \) sends any weakly convergent sequence \( f_n \in L_p \) to a sequence converging in the norm of \( L_{r'} \).
Remark 4.1. Compositions of convolution operators and multiplication operators in Lebesgue spaces have been treated in great generality in [10]. Lemma 4.1 is a particular case of Theorem 6.4 in [10], but it seems no easier to establish this relationship than to give an independent proof, which is what we do here.

Proof of Lemma 4.1. We can assume without loss of generality that $\|f_n\|_p \leq 1$ for all $n$. Let $f_n \to f$ in $L_p$. Then $\|f\|_p \leq 1$.

We consider the case $k \in L_q \cap L_\infty$ first. Since $q < p' < \infty$, we have $k \in L_{p'}$, and hence the sequence $k \ast f_n$ is pointwise convergent. Moreover, $\|k \ast f_n\|_\infty \leq \|k\|_{p'} \|f_n\|_p \leq \|k\|_{p'}$, and so,

$$|\chi(x) \cdot (k \ast f_n)(x)| \leq \|k\|_{p'} |\chi(x)|.$$ 

The majorant on the right lies in $L_{p'}$, and hence, using the dominated convergence theorem, we conclude that $\|f_n - f\|_{p'} \to 0$.

We now remove the assumption that $k \in L_\infty$. Let $K_\lambda$ be the convolution operator with the truncated function $k_\lambda(x) = k(x) I_{|k(x)| \leq \lambda}(x) \in L_q \cap L_\infty$. By the above, $\|\chi K_\lambda(f_n - f)\|_{p'} \to 0$. We complete the proof as follows. Given $\varepsilon > 0$, we find $\lambda$ such that $\|\chi\|_\infty \|k - k_\lambda\|_q < \varepsilon/3$. Now let $n_0$ be such that $\|\chi K_\lambda(f_n - f)\|_{p'} < \varepsilon/3$ for $n \geq n_0$. Then, for $n \geq n_0$,

$$\|\chi K(f_n - f)\|_{p'} \leq \|\chi K_\lambda(f_n - f)\|_{p'} + \|k(\lambda) - k\|_{p'} + \|\chi(k_\lambda - k) \ast f\|_{p'} < \varepsilon.$$ 

(The second and third terms in the middle part are estimated using Young’s inequality.) The lemma is proved.

Corollary 4.1. Let $k \in L_q(\mathbb{R}^d)$ and let $\Omega \subset \mathbb{R}^d$ be of finite measure. Suppose that the sequence $(f_n)$ converges weakly in $L_p(\mathbb{R}^d)$. Then the sequence of restrictions of the convolution $(k \ast f_n)|_\Omega$ converges in the norm of $L_{p'}(\Omega)$.

4.2. Convolution preserves tightness.

Lemma 4.2. Let $f \in L_p$, $k \in L_q$, let $v$ be a unit vector in $\mathbb{R}^d$, and let $[a_1, b_1] = \text{supp}^p_{\delta_1,v}(f)$ and $[a_2, b_2] = \text{supp}^q_{\delta_2,v}(k)$. Set $a = a_1 + a_2$, $b = b_1 + b_2$ and $h = k \ast f$. Then, for sufficiently small $\delta_1$ and $\delta_2$,

$$\int_{a < (x,v) < b} |h(x)|^{r'} \, dx \geq \|h\|_{r'}^{r'} - c(\delta_1^{1/p} + \delta_2^{1/q}),$$

where the constant $c$ depends on $f(\cdot)$, $k(\cdot)$, $p$, $q$ and $r$, but is independent of $a_i$, $b_i$, $v$ and $\delta_i$, $i = 1, 2$.

Proof. Let $\chi_i(x)$ be the indicator function of the set $\{x \mid a_i < (x,v) < b_i\}$, $i = 1, 2$. We set $\widehat{f} = f \chi_1$, $\widehat{k} = k \chi_2$, $\widehat{h} = \widehat{k} \ast \widehat{f}$. By the definition of a $\delta$-near-support, we have $\|f - \widehat{f}\|_p = \delta_1^{1/p}$ and $\|k - \widehat{k}\|_q = \delta_2^{1/q}$. By Young’s inequality,

$$\|h - \widehat{h}\|_{p'} \leq \|k - \widehat{k}\|_q \|f\|_p + \|f - \widehat{f}\|_p \|k\|_q + \|k - \widehat{k}\|_q \|f - \widehat{f}\|_p \leq \|f\|_p \delta_2^{1/q} + \|k\|_q \delta_1^{1/p} + \delta_2^{1/q} \delta_1^{1/p}.$$
We denote the right-hand side of the last inequality by $A$. Since $\hat{h}(x) = 0$ for $x \notin [a,b]$, we have
\[
\int_{a<(x,v)<b} |h(x)|^r'\,dx \geq \|h\|_{r'}^r - \int_{a<(x,v)<b} |h(x) - \hat{h}(x)|^r'\,dx \geq (\|h\|_{r'} - \|\hat{h}\|_{r'})^r' - \|h - \hat{h}\|_{r'}^r \geq \|h\|_{r'}^r - r'\|\hat{h}\|_{r'}^{-1} A - A'
\]
(here we have used Bernoulli’s inequality in the second line). The conclusion of the lemma follows immediately from this estimate. The proof is complete.

**Lemma 4.3.** Let $(f_n)$ be a tight sequence in $L_p$ and let $k \in L_q$. Then $(k * f_n)$ is a tight sequence in $L_{r'}$.

**Proof.** We put $g_n = k * f_n$. Let $\delta > 0$ be given. We need to show that the sequence $(g_n)$ is $\delta$-near-compactly-supported of order $r'$.

We take $\delta_1$ and $\delta_2$ so as to have $c\delta_1^{1/p} = c\delta_2^{1/q} = \delta/(2d)$. Let $Q = \prod_{j=1}^d [a_{1j}, b_{1j}]$ be a cube for which
\[
\int_Q |f_n(x)|^p\,dx \geq \|f_n\|_p^p - \delta_1
\]
(see Definition 2.5). We set $[a_{2j}, b_{2j}] = \text{supp}_\delta^q (k)$, $j = 1, \ldots, d$. By Lemma 4.2 we have
\[
\int_{(x,e_j) \notin [a_j,b_j]} |g_n(x)|^r'\,dx \leq \frac{\delta}{d}, \quad j = 1, \ldots, d,
\]
where $a_j = a_{1j} + a_{2j}$ and $b_j = b_{1j} + b_{2j}$. Therefore,
\[
\int_{\prod_{j=1}^d [a_j,b_j]} |g_n(x)|^r'\,dx \geq \|g_n\|_{r'} - \delta.
\]

It remains to extend the parallelepiped $\prod_{j=1}^d [a_j, b_j]$ to a cube. Thus, the $\delta$-near-compactly-supported condition for the sequence $(g_n)$ is fulfilled. The lemma is proved.

### 4.3. Shifts, centring and tightness.

The lemmas in this subsection are just various technical expressions of the simple idea: if a mass is concentrated in a bounded neighbourhood of the origin, then a long distance shift is incompatible with mass centring.

We first consider the concepts in Definition 2.4 and prove that the set of shift vectors that provide $\delta$-near-centring of a given function with various small $\delta$ is bounded.

**Lemma 4.4.** Let $1 \leq p < \infty$. Fix $f \in L_p$, $\delta_0 < \|f\|_p^p/3$ and a unit vector $v \in \mathbb{R}^d$, and set $D = D^p_{\delta_0,v}(f)$. Let $a_0 \in \mathbb{R}^d$ be a vector such that the function $T_{a_0}f$ is $\delta_0$-near-centred in the direction $v$. Suppose that the function $T_{a}f$ is $\delta$-near-centred in the direction $v$ for some $\delta \leq \delta_0$ and $a \in \mathbb{R}^d$. Then $|(a - a_0, v)| \leq D$.

**Proof.** Assuming without loss of generality that $v = (1,0,\ldots,0)$ and considering the univariate function
\[
f_1(x_1) = \int_{\mathbb{R}^{d-1}} |f(x)|^p\,dx_2 \ldots dx_n,
\]
the general case can be reduced to the case \( d = 1, \, p = 1, \, f \geq 0 \) (where \( f \) now denotes the function \( f_1 \) from the above displayed formula).

Now \( a_0 \) and \( a \) are scalars. Let \( \|f\|_1 = m \). By the centring assumption we have, first,
\[
\int_{a_0 - D}^{a_0} f(x) \, dx \geq \frac{m - \delta_0}{2} \quad \text{and} \quad \int_{a_0}^{a_0 + D} f(x) \, dx \geq \frac{m - \delta_0}{2},
\]
and second,
\[
\int_{-\infty}^{a} f(x) \, dx \geq \frac{m - \delta}{2} \quad \text{and} \quad \int_{a}^{\infty} f(x) \, dx \geq \frac{m - \delta}{2}.
\]
Suppose that \( a > a_0 + D \). Then
\[
\frac{m}{3} < \frac{m - \delta}{2} \leq \int_{a}^{\infty} f(x) \, dx \leq \int_{a_0 + D}^{\infty} f(x) \, dx \leq m - \int_{a_0 - D}^{a_0} f(x) \, dx \leq \delta_0 < \frac{m}{3},
\]
a contradiction. The same analysis gives a contradiction under the assumption that
\( a < a_0 - D \). Therefore, \( |a - a_0| \leq D \). The lemma is proved.

The remaining lemmas in this subsection pertain to the concepts introduced in Definition 2.5.

**Lemma 4.5.** Let \( a_n \in \mathbb{R}^d \) be a bounded sequence of vectors. Assume that sequences \((f_n)\) and \((\widehat{f}_n)\) in \( L_p \) are related by shifts \( \widehat{f}_n = T_{a_n} f_n \) and that one of them is tight. Then the other sequence is too.

**Proof.** Let \( \|a_n\| \leq R \) for all \( n \). For any coordinate cube \( Q \), the shifted cube \( T_{a_n} Q \) is contained in the cube \( QR \), which is independent of \( n \), concentric with \( Q \) and whose side is greater by \( 2R \) than that of \( Q \). Therefore, for any \( \delta > 0 \) the sequence \((f_n)\) is \( \delta \)-near-compactly-supported if and only if the sequence \((\widehat{f}_n)\) is too. The lemma is proved.

**Lemma 4.6.** Let \((f_n) \subset L_p, \, 1 \leq p < \infty,\) be a relatively tight sequence and let \( \|f_n\|_p = 1 \) for all \( n \). Suppose that all functions \( f_n \) are \( \delta_0 \)-near-centred (of order \( p \)) for some \( \delta_0 < 1/3 \). Then the sequence \((f_n)\) is tight.

**Proof.** We can assume without loss of generality that \( \delta = \delta_0 \) and that the condition in Definition 2.5 is satisfied for all vectors \( e_j \) in the fixed orthonormal basis in \( \mathbb{R}^d \).
This means that there exists \( D_0 > 0 \) such that, for any \( n \geq 1 \),
\[
D^p_{\delta_0, e_j} (f_n) < D_0.
\]

Given \( \delta \in (0, 1/3) \), we verify the \( \delta \)-near-compactly-supported condition. We choose vectors \( a_n \) so that the shifted functions \( \widehat{f}_n = T_{a_n} f_n \) are \( \delta \)-near-centred. By Lemma 4.4, we have \( |(a_n, e_j)| \leq D_0 \) for all \( n \geq 1 \) and \( j = 1, \ldots, d \).

By the definition of a relatively tight sequence, \( D^p_{\delta/d, e_j} (f_n) \leq D \) for some \( D \), all \( n \) and \( j = 1, \ldots, d \). Therefore,
\[
\int_{|(x, e_j)| > D + D_0} \left| f_n(x) \right|^p \, dx \leq \int_{|x, e_j| > D} \left| \widehat{f}_n(x) \right|^p \, dx \leq \frac{\delta}{d}, \quad j = 1, \ldots, d.
\]
We set \( R = D + D_0 \) and consider the cube \( Q = [-R, R]^d \). The complement of \( Q \) is the union of the sets \( \{ x \mid |(x, e_j)| > R \} \), \( j = 1, \ldots, d \). Therefore, \( \int_Q |f_n(x)|^p \, dx \geq \| f_n \|^p_p - d \cdot (\delta/d) = 1 - \delta \).

This verifies the \( \delta \)-near-compactly-supported condition, proving the lemma.

The next lemma, though not used in the proof of Theorem 1.1, expresses most clearly the relationship between relative tightness and tightness.

**Lemma 4.7.** Let \( (f_n) \) be a relatively tight sequence in \( L_p, 1 \leq p < \infty \). Assume that \( (f_n) \) is normalized (\( \| f_n \|_p = 1 \) for all \( n \)) and is \( \delta_0 \)-near-compactly-supported for some \( \delta_0 < 1/3 \). Then \( (f_n) \) is tight.

**Proof.** Consider a \( \delta_0 \)-near-centred sequence \( (\hat{f}_n) \) obtained from \( (f_n) \) by suitable shifts, \( \hat{f}_n = T_{a_n} f_n \). We prove that the sequence \( (a_n) \) is bounded in \( \mathbb{R}^d \).

We can assume that the coordinate cube \( Q \) from the \( \delta_0 \)-near-compactly-supported condition has centre at the origin and is given by the inequalities \( |(x, e_j)| \leq R, j = 1, \ldots, d \). Assuming that \( (a_n, e_j) > R \), we have

\[
\int_{(x,e_j) > 0} |\hat{f}_n(x)|^p \, dx < \int_{(x,e_j) > R} |f(x)|^p \, dx \leq \delta_0,
\]

which contradicts the fact that the function \( \hat{f}_n \) is \( \delta_0 \)-near-centred. Hence, \( (a_n, e_j) \leq R \). A similar analysis shows that \( (a_n, e_j) \geq -R \). So \( \sup_n \| a_n \| \leq R\sqrt{d} \).

Now the required result follows by applying Lemma 4.6 to the sequence \( (\hat{f}_n) \) and then applying Lemma 4.5 to the sequences \( (f_n) \) and \( (\hat{f}_n) \). The lemma is proved.

### 4.4. The final lemma.

**Lemma 4.8.** Suppose that a sequence of functions \( (f_n) \) in \( L_p(\mathbb{R}^d) \) has the following properties:

(i) normalization: \( \| f_n \|_p = 1 \) for all \( n \);
(ii) tightness (see Definition 2.5);
(iii) local convergence: there exists a function \( f \in L_p \) to which \( f_n \) converges on bounded subsets \( \| I_\Omega (f_n - f) \|_p \to 0 \) for any bounded set \( \Omega \subset \mathbb{R}^d \).

Then \( f_n \to f \) in \( L_p \). In particular, \( \| f \|_p = 1 \).

**Proof.** Let \( \varepsilon > 0 \) be given. We take a bounded set \( U \) such that \( \| I_{\mathbb{R}^d \setminus U} f \|_p < \varepsilon/3 \).

By conditions (i) and (ii), there exist a number \( n_1 \) and a cube \( Q \) in \( \mathbb{R}^d \) such that

\[
\| I_{\mathbb{R}^d \setminus Q} f_n \|_p < \varepsilon/3 \quad \text{for} \quad n \geq n_1.
\]

We set \( \Omega = U \cup Q \). By condition (iii) there exists \( n_2 \) such that \( \| I_\Omega (f_n - f) \|_p < \varepsilon/3 \) for \( n \geq n_2 \). Clearly, for \( n \geq \max(n_1, n_2) \) we have

\[
\| f_n - f \|_p \leq \| I_\Omega (f_n - f) \|_p + \| I_{\mathbb{R}^d \setminus \Omega} f_n \|_p + \| I_{\mathbb{R}^d \setminus \Omega} f \|_p < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,
\]

proving the lemma.

Thus Proposition 2.1, and therefore also Theorem 1.1 are proved.
§ 5. Limit cases

We examine the question of whether a maximizer exists in the limit cases excluded from Theorem 1.1. For $1 \leq p, q \leq \infty$, relation (1.1) defines the exponent $r \in [1, \infty]$ if and only if $1/p + 1/q \geq 1$; that is, $q \leq p'$. For $q = p'$ we have $r' = \infty$. On the plane with coordinates $u = 1/q$, $v = 1/p'$ the domain corresponding to the admissible pairs $(q, p)$ in Young’s inequality is the triangle formed by the lines I: $v = 0$ (that is, $p = 1$), II: $u = v$ (that is, $r' = \infty$), and III: $u = 1$ (that is, $q = 1$). So, there are three limit cases, with subcases, corresponding to the vertices of the triangle (see Figure 2).

![Figure 2](image-url)

Figure 2. Existence/nonexistence of a maximizer: (a) a maximizer exists; (b) a maximizer may fail to exist; existence is unknown; (c) both possibilities; (d) a maximizer is given by a delta function.

5.1. Case I: $p = 1$. Among the maximizing sequences we have, in particular, delta sequences, and hence we always have $\|K_k\|_{1,r} = \|k\|_r$. Moreover, if we consider the operator $K$ in the extended space $\mathcal{M} = (C_0)^*$ of finite Borel measures (the dual space of the space of continuous functions that vanish at infinity), we see that the delta function is a maximizer: $k \ast \delta = k$. We shall enlarge on this later, once we have analyzed Case II. Now we consider the more delicate question of the existence of a maximizer in $L_1$.

Subcase I(A): $p = q = r' = 1$.

The existence of a maximizer is equivalent to the existence of a function $0 \neq f \in L_1$ for which the Minkowski inequality

$$\int \left| \int k(y-x)f(x)\,dy \right| \,dx \leq \|k\|_1 \|f\|_1 = \int \int |f(x)| |k(y-x)| \,dy \,dx$$

becomes an equality.

**Proposition 5.1.** (a) Let $k(x) = |k(x)|e^{i\alpha(x)} \in L_1$. The function $\alpha(\cdot)$ with values in $(-\pi, \pi]$ is defined on the set $A = \{x \mid k(x) \neq 0\}$. A necessary and sufficient condition for the convolution operator $K: L_1 \to L_1$ with kernel $k$ to have a maximizer is that there exist a set $B$ of positive measure, a measurable real function $\beta(\cdot)$ on $B$ and a function $\psi(\cdot)$ on $A + B$ such that, for almost all $(x, y) \in A \times B$,

$$\psi(x+y) = \alpha(x) + \beta(y) \mod 2\pi.$$  

(5.1)

(This is known as the Pexider equation; see [12], § 4.3.)
(b) The above condition is satisfied, in particular, if the kernel $k(x)$ reads as

$$k(x) = e^{itx} \sum_{m=1}^{M} g_m(x)e^{i\phi_m},$$  

(5.2)

where $M \in \mathbb{N} \cup \{\infty\}$, $t$ and the $\phi_m$ are real numbers and the functions $g_m \in L_1$ are nonnegative and such that there exists $\delta > 0$ such that, for $m \neq m'$, the distance between the (essential) supports of the functions $g_m$ and $g_{m'}$ is at least $\delta$.

Remark 5.1. In assertion (a) of Proposition 5.1 we face a technical problem, in that a sum of measurable sets may fail to be a measurable set. This can easily be circumvented by changing the function $k(\cdot)$ on a null set, if necessary, so that the set $A = \{x \mid k(x) \neq 0\}$ becomes a Borel set (or even an $F_\sigma$- or a $G_\delta$-set). In the same way, we can assume that $B$ is a Borel set. Then $A + B$ is measurable (see §1.10.9 in [9]). Moreover, the sets $A$ and $B$ can be corrected so that $A + B$ is open (see [11]).

Proof of Proposition 5.1. (a) The integral $L_1$-Minkowski inequality becomes an equality

$$\int \left| \int F(x, y) \, dy \right| \, dx = \int \int |F(x, y)| \, dx \, dy$$

if and only if $F(x, y) = |F(x, y)|e^{i\psi(x)}$ almost everywhere, where $\psi(x)$ is a real measurable function. In our case $F(x, y) = k(x - y)f(y)$ and the above condition assumes the form (5.1), where $B = \{y \mid f(y) \neq 0\}$, $f(y) = |f(y)|e^{i\beta(y)}$ on $B$. In the proof of necessity a maximizer $f$ exists by assumption, and in the proof of sufficiency we can take as $|f|$ any norm-1 function that vanishes outside $B$.

(b) Under condition (5.2), as a maximizer we can take an arbitrary function $f$ such that $f(x)e^{-itx} \geq 0$, $\|f\|_1 = 1$, and the support of $f$ lies in the $(\delta/2)$-neighbourhood of the origin. Indeed, in this case $k(y)f(y - x) = e^{-itx}g(y)|f(y - x)|$, and $\|k * f\|_1 = \sum_m \int g_m(x) * f(x) \, dx = \|k\|_1\|f\|_1$. (The convolutions of $|f(x)|$ with the ‘pieces’ $g_m(x)e^{i\phi_m}$ of the function $k(x)e^{-itx}$, where the argument is constant, do not overlap.) This proves Proposition 5.1.

Subcase I(B): $p = 1$ and $1 < q = r' < \infty$. In this case no maximizer exists. Indeed, any maximizer $0 \neq f \in L_1$ must satisfy the equality

$$\left\| \int T_y k(\cdot)f(y) \, dy \right\|_q = \|f\|_1 \|k\|_q = \int \|T_y k(\cdot)\|_q |f(y)| \, dy,$$

that is, the Minkowski inequality must be an equality for it. Since $L_q$ is uniformly convex, this would imply that there is a function $\lambda(y)$ such that $T_y k(x) = \lambda(y)k(x)$ for almost all $x$ and $(|f(y)| \, dy)$-almost all $y$. But, clearly, for $0 \neq k \in L_q$ this is impossible.

Subcase I(C): $p = 1$ and $q = r' = \infty$. Setting $M = \|k\|_\infty$, we have $\|K_k\|_{1,1} = M$. If $|k|^{-1}(M)$ is a set of positive measure, then a maximizer exists. (The proof is similar to the proof of sufficiency in assertion (b) of Proposition 5.2, which follows, but the argument is simpler.) We now give a criterion for the existence of a maximizer in the case when $|k|^{-1}(M)$ is a null set. Below $\mu(\Omega)$ denotes the Lebesgue measure of a set $\Omega$. 

On maximizers of a convolution operator 1143
Proposition 5.2. Assume that \( k \in L_\infty \) and \( \mu(|k|^{-1}(M)) = 0 \). Set \( U_a = \{ x \mid |k(x)| \geq M - a \} \).

a) A necessary condition for the convolution operator \( K_k : L_1 \to L_\infty \) to have a maximizer is that \( \mu(U_a) = \infty \) for any \( a > 0 \).

b) A necessary and sufficient condition for a maximizer to exist is that there exist a sequence of vectors \( v_n \in \mathbb{R}^d \) and a set \( \Omega \) of positive measure on which the sequence of functions \( T_{v_n}k \) converges in measure to a function \( m(x) \) with \( |m(x)| = M \).

Proof. a) Assume that \( \inf_a \mu(U_a) < \infty \). By the continuity of the measure, we have \( \lim_{a \to 0} \mu(U_a) = \mu(|k|^{-1}(M)) = 0 \). Let \( \int |f(x)| \, dx = 1 \). We prove that \( f \) is not a maximizer. Let \( \varepsilon \in (0,1) \). Since the Lebesgue integral is absolutely continuous (for example, see [9], vol. 1, Theorem 2.5.7), there exists \( \delta > 0 \) such that \( \int_\Omega |f(x)| \, dx < \varepsilon \) if \( \mu(\Omega) < \delta \). Let \( a \) be such that \( \mu(U_a) < \delta \). Then \( A(x) = \int_{U_a} |f(x-y)| \, dy < \varepsilon \) for any \( x \). Therefore,

\[
|k * f(x)| \leq \int_{U_a} |k(y)f(x-y)| \, dy + \int_{\mathbb{R}^d \setminus U_a} |k(y)f(x-y)| \, dy \\
\leq MA(x) + (M-a)(1-A(x)) \leq M-a(1-\varepsilon),
\]

implying that \( \|k * f\|_\infty < M \).

b) Sufficiency. We can assume that in the hypothesis of the proposition the set \( \Omega \) has finite measure. We claim that \( f(x) = (M\mu(\Omega))^{-1}I_\Omega(-x)m(-x) \) is a maximizer. By assumption, for any \( a > 0 \) there exists \( n \) such that the measure of \( D = \{ x \in \Omega : |T_{v_n}k(x) - M(x)| > a \} \) is less than \( a \). Hence

\[
|k * f(-v_n) - M| = \frac{1}{M\mu(\Omega)} \left| \int_\Omega k(x-v_n)m(x) \, dx - \int_\Omega |m(x)|^2 \, dx \right| \\
\leq \frac{1}{\mu(\Omega)} \left( \int_D + \int_D \right) |T_{v_n}k(x) - m(x)| \, dx \leq a + 2Ma.
\]

Now \( \|k * f\|_\infty \geq M \), because \( a \) is arbitrary and the convolution is continuous.

Necessity. Let \( f \) be some maximizer of the operator \( K \). Then there exists a sequence \( (x_n) \) such that \( |k * f(x_n)| \to M \). Passing to a subsequence if necessary, we assume that \( \arg(k * f(x_n)) \) converges; multiplying \( f \) by an appropriate \( e^{i\theta} \) we can assume that \( k * f(x_n) \to M \). We set \( v_n = -x_n \) and \( f(x) = |f(x)|e^{i\phi(x)} \).

Let \( c > 0 \) be a constant for which \( \Omega = \{ x : |f(-x)| \geq c \} \) is a set of positive measure. Setting

\[
m(x) = Me^{-i\phi(-x)} \quad \text{and} \quad g_n(x) = (k(x-v_n) - m(x))e^{i\phi(-x)},
\]

we will prove that the sequence \( (g_n) \) converges to zero in measure in \( \Omega \). Given \( \varepsilon > 0 \), we set

\[
V_n = \{ x \in \Omega : |g_n(x)| > \varepsilon \}.
\]

We use the following elementary geometric inequality: \( |z| \leq M, |z-M| \geq \varepsilon \) \( \Rightarrow \) \( \Re z \leq M - \varepsilon' \), where \( \varepsilon' = \varepsilon^2/(2M) \). Given \( x \in V_n \), if \( z = k(x-v_n)e^{i\phi(-x)} \), then

\[
\Re(k(x-v_n)f(-x)) \leq (M-\varepsilon')|f(-x)|.
\]
Therefore,
\[ \text{Re}(k \ast f(x_n)) = \left( \int_{\mathbb{R}^d \setminus V_n} + \int_{V_n} \right) \text{Re}(k(x - v_n)f(-x)) \, dx \]
\[ \leq \int_{\mathbb{R}^d} M|f(-x)| \, dx - \int_{V_n} \varepsilon'|f(-x)| \, dx \leq M - \varepsilon'\mu(V_n). \]

We have \( k \ast f(x_n) \to M \) as \( n \to \infty \), and hence \( \mu(V_n) \to 0 \), as required. Proposition 5.2 is proved.

**Remark 5.2.** If \( k(x) \geq 0 \), then the criterion for the existence of a maximizer in assertion (b) of Proposition 5.2 can be stated differently:

\[ \exists \Omega \subset \mathbb{R}^d, \quad \mu(\Omega) > 0, \quad \forall a > 0 \quad \inf_{v \in \mathbb{R}^d} \mu(\Omega \setminus T_v U_a) = 0. \]

For example (in the one-dimensional case), for \( k(x) = e^{-|x|} \) no maximizer exists, for \( k(x) = 1 + \tanh x \) a maximizer does exist, and for \( k(x) = \sin |x|^\alpha \) a maximizer exists if and only if \( 0 \leq \alpha < 1 \).

### 5.2. Case II: \( r' = \infty \) and \( q = p' \in [1, \infty) \).

This case is clear: a maximizer exists. In fact, for \( f(x) = |k(-x)|^{q/q'}k(-x) \) (the subcase \( q = 1, \quad q/q' = 0 \) is also included) we have \( \|f\|_p = \|k\|_q^{-1} \) and \( \|k\|_q = (k \ast f)(0) \). Therefore, Young’s inequality \( \|k\|_q\|f\|_p \geq \|k \ast f\|_\infty \) becomes an equality and \( f \) is a maximizer.

Note that Cases I(A), I(B) and II describe operators which are the transposes one of another. (Case I(C) is self-dual.) The corresponding bilinear form is formally the same, however the conclusions about the existence of maximizers are different. This is because the spaces \( L_1 \) and \( L_\infty \) are nonreflexive. To make this more clear, let \( k \in L_q, \quad 1 \leq q < \infty \). We denote the operator of convolution with \( k \) by \( K \); this operator acts from \( L_1 \) into \( L_q \) (this is Case I) and the dual operator, which acts from \( L_{q'} \) into \( L_\infty \) (this is Case II), is denoted by \( K' \). For \( K' \) there exists a maximizer \( g_0 \in L_{q'} \). We have

\[ \|K\| = \|K'\| = \max_{g: \|g\|_{q'} = 1} \|K'g\|_\infty = \|K'g_0\|_\infty = \sup_{f: \|f\|_1 = 1} |(f, K'g_0)|. \]

The supremum over \( f \in L_1 \) need not be attained in general; however, it is attained if we assume that \( f \in L_\infty^* \). More precisely, by a corollary to the Hahn-Banach theorem, there exists \( \phi \in L_\infty^*, \quad \|\phi\|_{L_\infty^*} = 1 \), such that \( |(\phi, K'g_0)| = \|K'g_0\|_\infty = \|K\| \). To be even more explicit, note that the range of the operator \( K' \) lies in the closed subspace \( C_0 \subset L_\infty \cap C \) of continuous functions that vanish at infinity; hence we can assume that \( \phi \in C_0^* = \mathcal{M} \). The concrete choice is \( \phi = \delta_{x_0} \), where \( x_0 \) is a point of maximum of the function \( |K'g_0(x)| \).

### 5.3. Case III: \( q = 1 \) and \( p = r' \in (1, \infty) \).

#### 5.3.1. The case of a positive kernel.

We claim that if \( k \geq 0 \), then no maximizer exists.

Here \( \|K_k\|_{p,p'} = \|k\|_1 \). As a maximizing sequence of pairs \( \{f_n, g_n\} \) for the bilinear form \( (k \ast f, g) \) we can take, for example,

\[ f_n(x) = n^{-1/p} I_{[0,n]}(x), \quad g_n(x) = n^{-1/p'} I_{[0,n]}(x). \]
Then \( \|f_n\|_p = \|g_n\|_{p'} = 1 \) and
\[
(k \ast f_n, g_n) = \frac{1}{n} \int_0^n \int_{x-n}^x k(t) dt \, dx = \int_{-n}^n k(t) \left(1 - \frac{|t|}{n}\right) dt \to \int k(t) dt.
\]

The situation is similar to that for Case I(B), swapping the functions \( k \) and \( f \).

For a tentative maximizer \( f \), the Minkowski inequality
\[
\left\| \int T_y f(\cdot) k(y) dy \right\|_p = \left\| k \right\|_1 \left\| f \right\|_p = \int \left\| T_y f(\cdot) \right\|_p k(y) dy
\]
should become an equality, which is impossible. (As distinct from Case I(B), even a maximizer which is a distribution does not exist here.)

5.3.2. The case of a general kernel. We show that in the general case (of functions with variable sign) a maximizer can exist. Let \( p = r' = 2 \). The operator \( K_k \) acts in the Hilbert space \( L_2 \) and is unitarily equivalent to the operator of multiplication by the continuous function \( w = \mathcal{F} k \), where \( \mathcal{F} \) is the Fourier transform with unitary normalization. Let \( m = \|w\|_{\infty} \). A maximizer exists if and only if the set \( \{\xi : |w(\xi)| = m\} \) has positive measure. This is possible. For example, let \( w \) be the standard 'hat' function \( w \in C_0^\infty \), \( 0 \leq w(\xi) \leq 1 \) everywhere, and let \( w(\xi) = 1 \) in some neighbourhood \( U \) of the origin. Then \( k = \mathcal{F}^{-1} w \in L_1 \) and
\[
\|k \ast \mathcal{F}^{-1} I_U\|_2 = \|w \cdot I_U\|_2 = \|I_U\|_2 \|K_k\|_{2,2}.
\]

Whether a maximizer can exist in the case \( p \neq 2 \) remains an open question.

Acknowledgement. We thank Prof. R. Frank (Munich) for a suggestion that led the proof of Theorem 1.1 being simplified.

Bibliography

[1] M. Pearson, “Extremals for a class of convolution operators”, *Houston J. Math.* 25:1 (1999), 43–54.
[2] G. H. Hardy, “The constants of certain inequalities”, *J. London Math. Soc.* 8:2 (1933), 114–119.
[3] K. I. Babenko, “An inequality in the theory of Fourier integrals”, *Izv. Akad. Nauk SSSR. Ser. Mat.* 25:4 (1961), 531–542; English transl. in Amer. Math. Soc. Transl. Ser. 2, vol. 44, Amer. Math. Soc., Providence, RI 1965, pp. 115–128.
[4] W. Beckner, “Inequalities in Fourier analysis”, *Ann. of Math.* (2) 102:1 (1975), 159–182.
[5] P. L. Lions, “The concentration-compactness principle in the calculus of variations. The locally compact case. I”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 1:2 (1984), 109–145.
[6] K. Tintarev and K.-H. Fieseler, *Concentration compactness. Functional-analytic grounds and applications*, Imperial College Press, London 2007, xii+264 pp.
[7] M. O. Korpusov and A. G. Sveshnikov, *Nonlinear functional analysis and mathematical modelling in physics. Methods for analysis of nonlinear operators*, KRASAND, Moscow 2011, 480 pp. (Russian)
[8] W. Hengartner and R. Theodorescu, *Concentration functions*, Probab. Math. Statist., vol. 20, Academic Press, New York–London 1973, xii+139 pp.; Russian transl., Nauka, Moscow 1980, 174 pp.
[9] V. I. Bogachev, *Measure theory*, vols. 1, 2, Regulyarnaya i Khaoticheskaya Dinamika, Moscow–Izhevsk 2003, 544 pp., 576 pp.; English transl., vols. I, II, Springer-Verlag, Berlin 2007, xviii+500 pp., xiv+575 pp.

[10] R. C. Busby and H. A. Smith, “Product-convolution operators and mixed-norm spaces”, *Trans. Amer. Math. Soc.* **263**:2 (1981), 309–341.

[11] J. H. B. Kemperman, “A general functional equation”, *Trans. Amer. Math. Soc.* **86**:1 (1957), 28–56.

[12] J. Aczél and J. Dhombres, *Functional equations in several variables*, Encyclopedia Math. Appl., vol. 31, Cambridge Univ. Press, Cambridge 1989, xiv+462 pp.

Gleb V. Kalachev
Faculty of Mechanics and Mathematics,
Lomonosov Moscow State University,
Moscow, Russia
E-mail: gleb.kalachev@yandex.ru

Sergey Yu. Sadov
Moscow, Russia
E-mail: serge.sadov@gmail.com