STUDIA MATHEMATICA 113 (2) (1995)

Moser's Inequality for a class of integral operators

by

FINBARR HOLLAND (Cork) and DAVID WALSH (Maynooth)

Abstract. Let $1 < p < \infty$, $q = p/(p-1)$ and for $f \in L^p(0,\infty)$ define $F(x) = (1/x)\int_0^x f(t)\,dt$, $x > 0$. Moser's Inequality states that there is a constant C_p such that

$$\sup_{a\le 1}\sup_{f\in B_p}\int_0^\infty \exp[ax^q|F(x)|^q - x]\,dx = C_p$$

where B_p is the unit ball of L^p. Moreover, the value $a = 1$ is sharp. We observe that $F = K_1 f$ where the integral operator K_1 has a simple kernel K. We consider the question of for what kernels $K(t,x)$, $0 \le t$, $x < \infty$, this result can be extended, and proceed to discuss this when K is non-negative and homogeneous of degree -1. A sufficient condition on K is found for the analogue of Moser's Inequality to hold. An internal constant ψ, the counterpart of the constant a, arises naturally. We give a condition on K that ψ be sharp. Some applications are discussed.

Introduction. Let $1 < p < \infty$. Let q denote the exponent conjugate to p: $q = p/(p-1)$. Let L^p denote the class of Lebesgue measurable functions f on $(0,\infty)$ such that

$$\|f\|_p = \Big(\int_0^\infty |f(t)|^p\,dt\Big)^{1/p} < \infty,$$

and let B_p denote the unit ball of the resulting Banach space under the norm $f \to \|f\|_p$.

If f is Lebesgue measurable on $(0,\infty)$, we define

$$F(x) = \frac{1}{x}\int_0^x f(t)\,dt, \qquad x > 0.$$

The inequality of Moser mentioned in the title may be formulated as follows.

1991 *Mathematics Subject Classification*: 26D10, 47G10.
Key words and phrases: Moser's Inequality, integral operator, distribution function.

THEOREM 1. *There is a constant C_p, depending on p only, such that*

$$\sup_{a\le 1}\sup_{f\in B_p}\int_0^\infty \exp[ax^q|F(x)|^q - x]\,dx = C_p \tag{1}$$

while

$$\sup_{f\in B_p}\int_0^\infty \exp[ax^q|F(x)|^q - x]\,dx = \infty$$

for every $a > 1$.

The purpose of this paper is to extend this result to the case where F is given in terms of f by any one of a class of integral operators.

Moser [11] proved a slightly different, but equivalent, formulation of (1) for non-negative functions in B_p on the assumption that $2 \le p$, a result that was subsequently extended to the full range of p by Jodeit [8], Adams [1], and Marshall [9] amongst others. Apparently the argument given by Adams in particular is based on an idea of Garsia which was never published. The above statement follows easily from their work since B_p is invariant under the mapping $f \to |f|$ and

$$|F(x)| \le \frac{1}{x}\int_0^x |f|(t)\,dt, \qquad x > 0.$$

There are now several proofs of this result in the literature, but the best constant C_p remains to be determined, although Chang and Carleson [3] have shown that the supremum is attained when p is an integer ≥ 2. However, their work does not yield a description of the extremum function nor the value of the supremum.

Jodeit and Adams proved the above result by deriving estimates for the distribution function

$$\lambda \to d_f(\lambda) = |\{x > 0 : x - x^q F(x)^q \le \lambda\}|$$

of $0 \le f \in B_p$, where $|E|$ stands for the Lebesgue measure of a subset E of $(-\infty,\infty)$.

Recently, McCarthy [10] obtained the following sharp estimate for $\lambda \to d_f(\lambda)$:

$$\sup_{0\le f\in B_p} d_f(\lambda) = \mu(p)\lambda,$$

where

$$\mu(p) = (p-1)e^{q^p - q} + 1;$$

equality is attained for every $\lambda > 0$ by the function $t \to \lambda^{1/q}\phi(t/\lambda)$, where

$$\phi(t) = \begin{cases} \zeta(t) & \text{if } 0 \le t \le \mu(p),\\ \zeta(\mu(p)) & \text{if } t \ge \mu(p), \end{cases}$$

and

$$\zeta(t) = \begin{cases} (p-1)^{1/q} t/p & \text{if } 0 \le t \le p,\\ (t-1)^{1/q} & \text{if } t \ge p. \end{cases}$$

To motivate the results of this paper we note first of all that the relationship between f and F is linear, $F = K_1 f$, say, and that K_1 is a bounded operator from L^p to L^p with norm $\|K_1\|_p = q$. This is a classical result, which had its beginnings in early work by Hardy, and which was later refined by him and others. A companion result [7] was shown by Hardy and his co-workers to be true for the adjoint operator K_1^* defined by $G = K_1^* g$, where

$$G(y) = \int_y^\infty \frac{g(s)}{s}\,ds, \qquad y > 0.$$

Indeed, a duality argument establishes that $K_1^* : L^p \to L^p$ and $\|K_1^*\|_p = p$.

This circle of ideas suggests the following questions: Are there positive constants b_p, C_p^* such that

$$\sup_{b\le b_p}\sup_{0\le g\in B_p}\int_0^\infty \exp\left[by^q G(y)^q - y\right]\,dy = C_p^*?$$

If so, what are the best such constants?

Before concluding this section we show that this question has a positive answer when $p = 2$. The result that follows motivates the rest of the paper, which is devoted to a more complete analysis of a family of operators acting on L^p of which K_1 is one.

THEOREM 2. *In the notation above, $b_2 = 1$ and $C_2^* = C_2$. In other words,*

$$\sup_{b\le 1}\sup_{0\le g\in B_2}\int_0^\infty \exp\left[by^2 G(y)^2 - y\right]dy = C_2 \tag{†}$$

and

$$\sup_{0\le g\in B_2}\int_0^\infty \exp\left[by^2 G(y)^2 - y\right]dy = \infty$$

for every $b > 1$.

Proof. Let $0 \le g \in B_2$ and define $f = K_1^* g - g$. Then $f \in L^2$ and

$$\|f\|_2^2 = \|K_1^* g - g\|_2^2 = \|G\|_2^2 - 2\int_0^\infty G(y)g(y)\,dy + \|g\|_2^2$$

$$= \|g\|_2^2 + \int_0^\infty G(y)^2\,dy + \int_0^\infty y\frac{d}{dy}G(y)^2\,dy = \|g\|_2^2$$

after an integration by parts, the integrated terms vanishing because $y(G(y))^2 \to 0$ both as $y \to 0$ and as $y \to \infty$. Thus $f \in B_2$. By Moser's theorem,

$$\int_0^\infty \exp[bx^2|F(x)|^2 - x]\,dx \le C_2$$

for every $b \le 1$. But it readily follows from the definition of f and an integration by parts that

$$F(x) = K_1 f(x) = \frac{1}{x}\int_0^x f(t)\,dt = G(x), \quad x > 0.$$

Hence the left hand side of (†) is less than or equal to C_2; thus $b_2 \ge 1$ and if $b_2 = 1$ then $C_2^* \le C_2$.

We have shown above that, given $g \in L_2$, the solution of $f = K_1^* g - g$ is $g = K_1 f - f$. This reasoning can be reversed. Given $f \in B_2$, let $g = K_1 f - f = F - f$. We observe that $F'(x) = -F(x)/x + f(x)/x$. In the same way as we did above, by applying an integration by parts we can show that $\|g\|_2 = \|f\|_2$. Indeed, it can be shown that $f = K_1^* g - g$ and therefore $G = F$ as before.

It is now clear that $b_2 \le 1$ and further that if $b_2 = 1$ then $C_2 \le C_2^*$. The result is immediate.

Preliminaries. We return to Moser's Inequality as given by (1). In fact, Moser observed that the integrals are finite for all real a but are unbounded for $a > 1$. As noted above, $F = K_1 f$, where the integral operator K_1 has a simple kernel K. We pose the following question: For what kernels $K = K(t,x)$ can the result (1) be extended? More precisely, suppose $K = K(t,x)$ is a measurable function on $(0,\infty)\times(0,\infty)$ and is non-negative and homogeneous of degree -1 in t and x. We assume that

$$\int_0^\infty K(t,x)^q\,dt = \frac{1}{x^{q-1}}\int_0^\infty K(u,1)^q\,du = \frac{\psi}{x^{q-1}}$$

exists for each $x > 0$. For each $f \in L^p$ we define

$$H(x) = H(f,x) = x\int_0^\infty K(t,x)f(t)\,dt. \tag{2}$$

It follows from Hölder's Inequality that

$$|H(x)| \le x\Big(\int_0^\infty K(t,x)^q\,dt\Big)^{1/q}\|f\|_p,$$

from which we obtain

$$\psi^{-1}|H(x)|^q \le x \tag{3}$$

for every $x > 0$ and for every $f \in B_p$. From (3) it is immediate that

$$\sup_{f\in B_p}\int_0^\infty \exp[a|H(x)|^q - x]\,dx \le \frac{1}{1-a\psi}$$

for every $a < 1/\psi$.

Our aim is to find further conditions on K which ensure the following: Let $p > 1$ and let H be defined as in (2). Then there exists a constant C_p such that

$$\sup_{f\in B_p}\int_0^\infty \exp[\psi^{-1}|H(x)|^q - x]\,dx = C_p. \tag{4}$$

Before proceeding further there are a number of remarks to be made about (4).

Remarks. 1. In the first place the integral exists for all $f \in L^p$; this is a consequence of the fact that for fixed $f \in L^p$,

$$\lim_{x\to\infty}|H(x)|^q/x = 0. \tag{5}$$

In order to see this let $\varepsilon > 0$ be given. Choose y_1 such that $\int_{y_1}^\infty |f(t)|^p\,dt < \varepsilon^p$ and choose $\delta > 0$ such that $\int_0^\delta K(u,1)^q\,du < \varepsilon^q$. Writing

$$H(x) = x\int_0^\infty K(t,x)f(t)\,dt = \int_0^\infty K(t/x,1)f(t)\,dt$$
$$= \Big(\int_0^{\delta x} + \int_{\delta x}^\infty\Big)K(t/x,1)f(t)\,dt = I_1 + I_2,$$

we have

$$|I_1| \le \Big(\int_0^{\delta x}|f|^p\,dt\Big)^{1/p}\Big(\int_0^\delta xK(u,1)^q\,du\Big)^{1/q}$$
$$\le \|f\|_p\Big(x\int_0^\delta K(u,1)^q\,du\Big)^{1/q} \le \varepsilon\|f\|_p x^{1/q},$$

while for the second term we have

$$|I_2| \le \Big(\int_{\delta x}^\infty |f|^p\,dt\Big)^{1/p}\Big(x\int_\delta^\infty K(u,1)^q\,du\Big)^{1/q} \le \varepsilon(\psi x)^{1/q}$$

if $\delta x \ge y_1$, i.e. if $x \ge y_1/\delta$. Therefore, for such x, $|H(x)| \le \varepsilon(\|f\|_p + \psi^{1/q})x^{1/q}$ and (5) follows.

2. We note that Theorem 1 is the special case of (4) in which

$$K(t,x) = \begin{cases} 1/x, & 0 \le t \le x,\\ 0, & t > x.\end{cases} \tag{6}$$

3. We observe that the special kernel (6) is discontinuous at $t = x$. This simple fact is no accident; it points the way towards the extra condition required in order to prove the result indicated in (4). Indeed, we shall show later that some such condition must be imposed on the kernel, for if $K(t,x) = 1/(t+x)$ $(t,x > 0)$, then K is smooth and satisfies the previous requirements but (4) does not hold for K.

The main result. We can now formulate our main result. We continue to assume the kernel K satisfies the earlier conditions.

THEOREM 3. *Suppose the kernel $K(t,x)$ lacks smoothness at $t = x$ in the manner prescribed by the following property: For $0 < u < 1$ let*

$$G_1(u)^q = \int_{1-u}^{1} K(t,1)^q\,dt, \qquad G_2(u)^q = \int_{1}^{1+u} K(t,1)^q\,dt;$$

there exist numbers α, D, $0 < \alpha \le 1, D > 0$, such that

$$G_1(u)^q = O(u^\alpha), \quad G_2(u)^q = O(u^\alpha) \quad (u \to 0)$$

and

(7) $$Du^\alpha \le |G_1(u)^q - G_2(u)^q| \quad (u \to 0).$$

Then there exists a constant C_p, depending on p only, such that

(8) $$\sup_{f \in B_p} \int_0^\infty \exp[\psi^{-1}|H(x)^q| - x]\,dx = C_p.$$

Before giving the proof it is appropriate to make a number of qualifying remarks about the statement above.

Remarks. 1. For an arbitrary $f \in L^p$ the integral in (8) is finite. This follows from (5).

2. The condition on K is satisfied if for instance $K(u,1)$ has a simple discontinuity at $u = 1$.

3. The fact that the kernel is assumed to lack smoothness on the diagonal is not significant. The lack of smoothness may occur on any line $t = mx$ for any $m > 0$.

4. The constant $1/\psi$ in the exponent may be shown to be best possible if we impose a certain integral-mean condition on K.

5. This theorem should be compared and contrasted with Lemma 1 of Adams [1]. Moser's original inequality is a particular case of both results. However, neither result is contained in the other since the kernels used are entirely different.

Proof of Theorem 3. In the course of the argument we shall invoke a number of lemmas whose proof will be given in a later section. Without loss

of generality we shall assume henceforth that $\psi = 1$ and f is non-negative, with $\|f\|_p = 1$. We shall write

$$\beta_1 = \int_0^1 K(t,1)^q\,dt, \quad \beta_2 = \int_1^\infty K(t,1)^q\,dt, \quad \text{so that } \beta_1 + \beta_2 = 1.$$

We proceed by adopting the method given in [1] of analysing the distribution function of the exponent in (8). The basis of this approach is a subtle use of Hölder's Inequality and a close examination of the cases where equality almost holds there. To continue, let λ be real, $0 < f \in B_p$ and set

$$E_\lambda(f) = \{x \in (0,\infty) : x - H(x)^q \le \lambda\}, \qquad d_f(\lambda) = |E_\lambda(f)|.$$

It is clear from (3) that we need only consider $\lambda > 0$. It is a consequence of (5) that for a given f, $d_f(\lambda) \le 2\lambda$ for large λ. For, let us take $\varepsilon = 1/2^{1+1/q}$, $\lambda \ge y_1/\delta$, in the notation following (5). Then if $x \in E_\lambda$ and $H(x)^q < x/2$, we have $x \le \lambda + H(x)^q \le \lambda + x/2$, which implies $x \le 2\lambda$. On the other hand,

$$E_\lambda \cap \{x : H(x)^q \ge x/2\} \subset \{x : H(x)^q \ge x/2\} \subset [0, y_1/\delta] \subset [0,\lambda],$$

and therefore $d_f(\lambda) \le 2\lambda$. As a result the integral which appears on the left hand side of (8) may be written

$$\int_0^\infty e^{-\lambda}\,dd_f(\lambda) = [e^{-\lambda}d_f(\lambda)]_0^\infty + \int_0^\infty e^{-\lambda}d_f(\lambda)\,d\lambda = \int_0^\infty e^{-\lambda}d_f(\lambda)\,d\lambda.$$

If we can show that a number N exists independent of λ, f such that $d_f(\lambda) \le 2N\lambda$ then Theorem 3 will follow with $C = 2N$.

In order to simplify the exposition we introduce the group of isometries $\{\Gamma_\lambda : \lambda > 0\}$ on L_p defined by $\Gamma_\lambda f(x) = \lambda^{1/p} f(\lambda x)$, $f \in L_p$. It is clear that Γ_λ maps L_p onto itself and that $(\Gamma_\lambda)^{-1} = \Gamma_{1/\lambda}$. For any kernel K satisfying the conditions above, the integral operator K_1, defined on L_p by $K_1 f(x) = \int_0^\infty K(t,x)f(t)\,dt$, has the property that it commutes with Γ_λ:

$$\Gamma_\lambda K_1 = K_1 \Gamma_\lambda.$$

To see this, let $f \in L_p$. Then

$$\begin{aligned}\Gamma_\lambda K_1 f(x) = \lambda^{1/p} K_1 f(\lambda x) &= \lambda^{1/p} \int_0^\infty K(t,\lambda x) f(t)\,dt \\ &= \lambda^{1/p} \int_0^\infty (1/\lambda) K(t/\lambda, x) f(t)\,dt \\ &= \lambda^{1/p} \int_0^\infty K(u,x) f(\lambda u)\,du = K_1 \Gamma_\lambda f(x).\end{aligned}$$

Next we shall show that $E_\lambda(f) = \lambda E_1(\Gamma_\lambda f)$ for $\lambda > 0$. We recall that

$$E_\lambda(f) = \{x \in (0,\infty) : x - x^q(K_1 f(x))^q \le \lambda\}.$$

From the commutativity property,

$$\begin{aligned} E_1(\Gamma_\lambda f) &= \{y : y - y^q[K_1\Gamma_\lambda f(y)]^q \le 1\} = \{y : y - y^q[\Gamma_\lambda K_1 f(y)]^q \le 1\} \\ &= \{y : y - y^q\lambda^{q-1}[K_1 f(\lambda y)]^q \le 1\} \\ &= \{y : \lambda y - (\lambda y)^q[K_1 f(\lambda y)]^q \le \lambda\} \\ &= \{y : \lambda y \in E_\lambda(f)\} = \frac{1}{\lambda}E_\lambda(f), \end{aligned}$$

which is the desired result. It follows that $d_f(\lambda) = \lambda d_{\Gamma_\lambda f}(1)$, which in turn gives

$$\sup_{f\in B_p} d_f(\lambda) = \lambda \sup_{f\in B_p} d_f(1).$$

It suffices therefore to take $\lambda = 1$ in the sequel and we proceed to obtain a bound on $d_f(1)$. Let N be a positive number to be fixed later and let

$$E = E_1(f) = (E \cap (0,N]) \cup (E \cap (N,\infty)) \equiv E' \cup E''.$$

Suppose $E'' \ne \emptyset$ and let $x \in E''$. For any such x we have by Hölder's Inequality

$$\text{(9)} \qquad \begin{aligned} x - 1 \le H(x)^q &= x^q\Big\{\int_0^x f(t)K(t,x)\,dt + \int_x^\infty f(t)K(t,x)\,dt\Big\}^q \\ &\le x\Big\{\Big(\int_0^x f^p\,dt\Big)^{1/p}\Big(\int_0^1 K(t,1)^q\,dt\Big)^{1/q} \\ &\quad + \Big(\int_x^\infty f^p\,dt\Big)^{1/p}\Big(\int_1^\infty K(t,1)^q\,dt\Big)^{1/q}\Big\}^q \\ &\le x\Big\{\beta_1^{1/q}\Big(\int_0^x f^p\,dt\Big)^{1/p} + \beta_2^{1/q}\Big(\int_x^\infty f^p\,dt\Big)^{1/p}\Big\}^q. \end{aligned}$$

Equality holds in Hölder's Inequality if and only if

$$\text{(10)} \qquad f(t) = x^{1/(p(p-1))}K(t,x)^{q-1} = x^{-1/p}K(t/x,1)^{q-1} = v_x(t),$$

the extremal function at x. The term in curly brackets on the right hand side of (9) is less than or equal to 1 by another application of Hölder's Inequality. It is clear from (10) that x can be arbitrarily large depending on f. Consequently, if $1/x$ is small, it follows from (9), on division by x, that we would have what we may call approximate equality in Hölder's Inequality and we would expect that f is close to the extremal function v_x. This is indeed so and it is of key importance in the proof to be able to measure the difference. We can do this by means of Clarkson's Inequalities, and the result given in Lemma 1 says: *there exists a constant A such that*

$$\text{(11)} \qquad \begin{aligned} \|f - v_x\|_p^p &\le A/x, \quad 2 \le p < \infty, \\ \|f - v_x\|_p^q &\le A/x, \quad 1 < p < 2, \end{aligned}$$

provided $1/x < 1$.

Already from these inequalities we can deduce that E'' is contained in a finite interval $[0, Q(f)]$. However, we can deduce more than this on exploiting the homogeneity of the kernel. Let $x, y \in E''$ with $x < y$. It is not hard to see that if $t \in E_1(v_s)$ then $at \in E_a(v_{as})$ for all $a > 0$. This suggests that we consider the ratios y/x, $1/x$, so we introduce the new variables

$$\text{(12)} \qquad u = \frac{y-x}{x}, \quad \mu = \frac{1}{x}.$$

We use Hölder's Inequality once more as in (9), this time writing the integral for H as a sum of three parts and applying the inequality to each part separately to get

$$\begin{aligned} x - 1 \le x\Big\{&\Big(\int_0^x f^p\,dt\Big)^{1/p}\Big(\int_0^1 K(t,1)^q\,dt\Big)^{1/q} \\ &+ \Big(\int_x^y f^p\,dt\Big)^{1/p}\Big(\int_1^{y/x} K(t,1)^q\,dt\Big)^{1/q} \\ &+ \Big(\int_y^\infty f^p\,dt\Big)^{1/p}\Big(\int_{y/x}^\infty K(t,1)^q\,dt\Big)^{1/q}\Big\}^q. \end{aligned}$$

It is convenient at this point to introduce the function

$$Y(u)^q = \int_{1/(1+u)}^1 K(t,1)^q\,dt,$$

and we note that $Y(u)^q \sim G_1(u)^q$ as $u \to 0$. The last inequality now becomes

$$\text{(13)} \qquad \begin{aligned} 1 - 1/x \le \Big\{&\beta_1^{1/q}\Big(\int_0^x f^p\,dt\Big)^{1/p} + \Big(\int_x^y f^p\,dt\Big)^{1/p} G_2(u) \\ &+ (\beta_2 - G_2(u)^q)^{1/q}\Big(\int_y^\infty f^p\,dt\Big)^{1/p}\Big\}^q. \end{aligned}$$

It is appropriate also to include here the companion inequality of (13); it is a consequence of the fact that $y \in E''$ and is derived in the same way (see the case $1 < p < 2$ below for the details):

(13′) $$1-1/y \le \Big\{(\beta_1 - Y(u)^q)^{1/q}\Big(\int\limits_0^x f^p\,dt\Big)^{1/p} + \Big(\int\limits_x^y f^p\,dt\Big)^{1/p} Y(u) + \beta_2^{1/q}\Big(\int\limits_y^\infty f^p\,dt\Big)^{1/p}\Big\}^q.$$

We note that the right hand side in both cases is bounded by 1. Invoking Lemma 2 allows us to assume that there is a number M, depending on p, K only, such that $u \le M$ for all u satisfying (13). Indeed, it is further true that if $\mu \to 0$ then so does u, independently of f. The proof of this statement is contained in Lemma 3.

The final step in the proof is to show that there exists N such that $u \le N\mu$ for all u, μ, f for which (13) holds. For, once this has been established, then

$$y - x \le N$$

and hence $|E''| \le N$ giving in turn $d_f(1) \le 2N$. As has already been said, this would complete the proof of the theorem with $C = 2N$.

Suppose that the claim above is not true. Then, for arbitrarily large N we would be able to find functions f and numbers u, μ satisfying (13) such that $\mu < u/N$. Since we know $u \le M$ it follows that $\mu \to 0$ as $N \to \infty$ and consequently that also $u \to 0$.

To finish the argument we shall show that these statements are in conflict with (7) and this will establish the claim.

Proceeding to the proof of the final step we suppose for definiteness that $G_1(u) > G_2(u)$ for small u. We shall split the argument into two cases and shall assume first that $2 \le p < \infty$.

The case $2 \le p < \infty$. We recall (11) says that

(14) $$\|f - v_y\|_p^p < A/y \le A\mu$$

when $y \in E''$, $y > x$, $\mu = 1/x$ is small. In other words, f is then close to v_y, and it is convenient here to introduce numbers η_i, $1 \le i \le 3$, which describe the difference more explicitly. Accordingly we set

(15) $$\Big(\int\limits_0^x f^p\,dt\Big)^{1/p} = (\beta_1 - Y(u)^q)^{1/p} + \eta_1,$$
$$\Big(\int\limits_y^\infty f^p\,dt\Big)^{1/p} = \beta_2^{1/p} + \eta_2, \qquad \Big(\int\limits_x^y f^p\,dt\Big)^{1/p} = Y(u)^{q-1} + \eta_3.$$

These definitions make sense if we recall what v_y, β_1, β_2, $Y(u)$ are. It is readily seen that $\int_0^x v_y^p\,dt = \beta_1 - Y(u)^q$, $\int_x^y v_y^p\,dt = Y(u)^q$ and $\int_y^\infty v_y^p\,dt = \beta_2$.

Clearly then,

$$|\eta_1| = \Big|\Big(\int\limits_0^x f^p\,dt\Big)^{1/p} - \Big(\int\limits_0^x v_y^p\,dt\Big)^{1/p}\Big| \le \Big(\int\limits_0^x |f - v_y|^p\,dt\Big)^{1/p} \le \|f - v_y\|_p \le c\mu^{1/p},$$

with similar estimates holding for η_2 and η_3. Here c is a positive constant, though not always the same one, and this convention applies hereafter. Henceforth also, we shall often write Y for $Y(u)$, and G_2 for $G_2(u)$. The numbers η_i may be positive or negative and they satisfy the consistency equation

(16) $$[(\beta_1 - Y(u)^q)^{1/p} + \eta_1]^p + (\beta_2^{1/p} + \eta_2)^p + (Y(u)^{q-1} + \eta_3)^p = 1.$$

From this equation we deduce the second important piece of information we need about the η_i. It is contained in the following relation, whose proof is deferred to Lemma 4:

(17) $$\eta_1\beta_1^{1/q} + \eta_2\beta_2^{1/q} = Y^q/p - (Y^{q-1} + \eta_3)^p/p - \tfrac{1}{2}(p-1)(\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p}) + \eta_1 Y^q/(q\beta_1^{1/p}) + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q).$$

Returning now to (13) we have, on substituting from (15) and dropping the power q,

$$1 - 1/x \le \beta_1^{1/q}((\beta_1 - Y(u)^q)^{1/p} + \eta_1) + G_2(u)[Y(u)^{q-1} + \eta_3] + (\beta_2^{1/p} + \eta_2)(\beta_2 - G_2(u)^q)^{1/q},$$

which yields, on expanding by Taylor's Theorem,

$$1 - 1/x \le \beta_1^{1/q}\{\beta_1^{1/p}(1 - Y^q/(p\beta_1) + O(Y^{2q}) + \eta_1\} + (\beta_2^{1/p} + \eta_2)\beta_2^{1/q}\{1 - G_2^q/(q\beta_2) + O(G_2^{2q})\} + Y^{q-1}G_2 + G_2\eta_3.$$

On simplifying we get

$$Y^q/p + G_2^q/q - Y^{q-1}G_2 \le 1/x + \eta_1\beta_1^{1/q} + \eta_2\beta_2^{1/q} + G_2\eta_3 - G_2^q\eta_2/(q\beta_2^{1/p}) + O(Y^{2q}).$$

We notice that we can substitute for two of the terms on the right hand side using (17), whose significance will now become apparent. On doing this and recalling that $\mu \le u/N$ we obtain

(18) $$Y^q/p + G_2^q/q - Y^{q-1}G_2 \le u/N + Y^q/p - (Y^{q-1} + \eta_3)^p/p + \{\eta_3 G_2 - G_2^q\eta_2/(q\beta_2^{1/p}) + \eta_1 Y^q/(q\beta_1^{1/p})\} - \tfrac{1}{2}(p-1)[\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p}] + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q).$$

Our purpose is to derive a contradiction from (18) after some initial simplification. While the right hand side may seem inordinately unwieldy, in fact many of the terms can be incorporated in a single term of the form $cu^\alpha/N^{1/p}$ because they are of the same order of magnitude. The estimates for the η_i, obtained by means of Clarkson's Inequality, now allow us to do this for the three terms in curly brackets. Thus

$$|\eta_3 G_2(u)| \le cu^{1/p}u^{\alpha/q}/N^{1/p} \le cu^\alpha/N^{1/p}, \quad |\eta_1 Y(u)^q| \le cu^{\alpha+1/p}/N^{1/p},$$

with a similar inequality for the remaining term.

Equally important is the fact that all the terms involving η_1, η_2 may now be discarded from the right hand side. This fortuitous circumstance owes its validity to the simple observation that the two terms involving η_1^2, η_2^2 have coefficients which are constants with negative sign. Since the term in the square brackets dominates the higher order terms in the η_i, $i = 1, 2$, we may safely discard all these terms while still preserving the inequality. The remaining "O" term is of the form $O(u^{2\alpha})$, which, according to Lemma 3, is of the order $cu^\alpha/N^{1/p}$. Combining all these remarks we can say that the right hand side of (18) is less than

$$cu^\alpha/N^{1/p} + Y^q/p - (Y^{q-1} + \eta_3)^p/p.$$

Since $|\eta_3/Y^{q-1}| \le c_1(u^{\alpha/p}/N^{1/p})(c_2/u^{\alpha/p}) \le c/N^{1/p}$ the second and third terms may be combined to give $Y^q\{-p\eta_3/Y^{q-1} - \binom{p}{2}(\eta_3/Y^{q-1})^2 + \ldots\}/p$, which is of the order $u^\alpha/N^{1/p}$. Finally, therefore, we have in place of (18),

$$Y^q/p + G_2^q/q - Y^{q-1}G_2 \le cu^\alpha/N^{1/p}. \tag{19}$$

However, this is impossible by virtue of our initial assumption (7). The Arithmetic-Geometric Mean Inequality tells us that the left hand side of (19) is greater than or equal to 0 with equality if and only if $Y = G_2$. What we have in (19) is approximate equality in the Arithmetic-Geometric Mean Inequality, as we previously had in the Hölder Inequality, and an estimate for the size of the difference between Y and G_2 may be obtained from this. We carry out this analysis in Lemma 5 and the result is

$$\begin{aligned} Y(u)^q - G_2(u)^q &\le G_2^q\sqrt{\frac{3pq}{Y^q}}\sqrt{cu^\alpha/N^{1/p}} \\ &\le cY^{q/2}u^{\alpha/2}/N^{1/(2p)} \le cu^\alpha/N^{1/(2p)}. \end{aligned}$$

But this is in direct conflict with our hypothesis (7), which says that $Y(u)^q - G_2(u)^q \ge Du^\alpha$ as $u \to 0$, bearing in mind that N may be arbitrarily large. It is seen here that the assumption (7) was designed to prevent Y and G_2 from being too close. Therefore (19) cannot hold for arbitrarily large N and so there exists a constant N such that $u \le N\mu$ for all u, μ, f for which (13) holds.

We assumed above that both $\beta_1 > 0$ and $\beta_2 > 0$; the argument is easier if one of these is zero and this applies also to what follows. The final step has therefore been demonstrated and the proof of the Theorem is complete for the case $2 \le p < \infty$.

The case $1 < p < 2$. In this case the estimate which replaces (14) is, by Clarkson's Inequality,

$$\|f - v_y\|_p \le c(1/y)^{1/q}$$

with an exponent $1/q$ on the right hand side, whereas in the previous case the exponent was $1/p$. This difference marks a change in the analysis and unfortunately the straightforward estimates that we used in the last case no longer suffice entirely. This will become obvious in a moment.

From the inequality above we have the following estimates for the η_i: $|\eta_i| \le c\|f - v_y\|_p \le c(1/y)^{1/q}$, $1 \le i \le 3$. The inequality (18) follows as before but we shall not be able to discard terms on the right hand side as easily as before. The difficulty arises with the term η_3. We cannot a priori assume that

$$|\eta_3 G_2(u)| \le cu^\alpha/N^{1/p}, \tag{20}$$

as we had in the previous case. Indeed, on the basis of this estimate for η_3 we can only say that $|\eta_3 G_2(u)| \le cu^{1/q}u^{\alpha/q}/N^{1/q}$, and the exponent of u may well be less than α. In order to surmount this difficulty we shall need to derive another estimate for η_3 similar to that in the previous case, which will in turn give an analogue of (20). To achieve this, we note that $y \in E''$ implies that $y - 1 \le H(y)^q$ and we simply write down the counterpart of (13) for this case.

This yields, on first using the homogeneity and then dropping the power q,

$$\begin{aligned} y - 1 &\le y^q\Big\{\Big(\int_0^x f^p\,dt\Big)^{1/p}\Big(\int_0^x K(t,y)^q\,dt\Big)^{1/q} \\ &\quad + \Big(\int_x^y f^p\,dt\Big)^{1/p}\Big(\int_x^y K(t,y)^q\,dt\Big)^{1/q} \\ &\quad + \Big(\int_y^\infty f^p\,dt\Big)^{1/p}\Big(\int_y^\infty K(t,y)^q\,dt\Big)^{1/q}\Big\}^q \\ &\le y\{[(\beta_1 - Y^q)^{1/p} + \eta_1](\beta_1 - Y^q)^{1/q} \\ &\quad + [Y^{q-1} + \eta_3]Y + (\beta_2^{1/p} + \eta_2)\beta_2^{1/q}\}. \end{aligned}$$

Consequently,

$$1 - 1/y \le \beta_1 - Y^q + \eta_1(\beta_1 - Y^q)^{1/q} + \eta_3 Y + Y^q + \beta_2 + \eta_2\beta_2^{1/q},$$

which reduces to

$$0 \le 1/y + \eta_3 Y + \eta_1 \beta_1^{1/q} + \eta_2 \beta_2^{1/q} - \eta_1 Y^q/(q\beta_1^{1/p}) + O(Y^{2q}).$$

On substituting from (17) we get

$$\begin{aligned}0 \le\ & 1/y + \eta_3 Y + Y^q/p - (Y^{q-1}+\eta_3)^p/p\\ & - \tfrac{1}{2}(p-1)(\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p})\\ & + \eta_1 Y^q/(q\beta_1^{1/p}) - \eta_1 Y^q/(q\beta_1^{1/p}) + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q).\end{aligned}$$

Cancelling and rearranging gives

$$\begin{aligned}(21)\qquad (Y^{q-1}+\eta_3)^p/p \le\ & \mu + \eta_3 Y + Y^q/p\\ & - \tfrac{1}{2}(p-1)(\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p})\\ & + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q)\\ \le\ & (Y^{q-1}+\eta_3)Y - Y^q/q + \mu + O(Y^{2q}).\end{aligned}$$

Put $z = \eta_3 + Y^{q-1}$ so that $z \ge 0$. Then z satisfies

$$(22)\qquad z^p \le pYz - (p-1)Y^q + p\mu + O(Y^{2q}).$$

This imposes a bound on z which in turn imposes a bound on η_3 as we shall see. We define $F(z) = z^p - pYz + (p-1)Y^q$, so that (22) becomes

$$(23)\qquad F(z) \le p\mu + O(Y^{2q}).$$

Appealing to Lemma 6 we get the required bound on η_3, namely $|\eta_3| \le cu^{\alpha/p}/N^{1/(2q)}$. The proof now proceeds as in the case $2 \le p < \infty$. We look at (13) again and derive the inequality (18) as before:

$$\begin{aligned}(18)\quad Y^q/p + G_2^q/q - Y^{q-1}G_2 &\\ \le\ & u/N + Y^q/p - (Y^{q-1}+\eta_3)^p/p\\ & + \{\eta_3 G_2 - G_2^q\eta_2/(q\beta_2^{1/p}) + \eta_1 Y^q/(q\beta_1^{1/p})\}\\ & - \tfrac{1}{2}(p-1)[\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p}]\\ & + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q).\end{aligned}$$

We now have the right estimates for the term in curly brackets:

$$|\eta_3 G_2| \le cu^\alpha/N^{1/(2q)}, \qquad |\eta_1 Y^q| \le cu^{\alpha+1/q}/N^{1/q},$$

with a similar estimate for its third part, so that this term may certainly be replaced by $cu^\alpha/N^{1/(2q)}$. We estimate the sum of the second and third terms just as before and find that this is of the order $u^\alpha/N^{1/(2q)}$. Similar considerations as before now lead finally to

$$Y^q/q + G_2^q/p - YG_2^{q-1} \le cu^\alpha/N^{1/(2q)}.$$

But we know that this gives a contradiction.

The proof of the Theorem is therefore complete.

Auxiliary lemmas. We give here the proofs of the lemmas invoked in the course of the proof of the Theorem.

Suppose (Ω, μ) is a measure space and p, q are conjugate exponents with $1 < p < \infty$. If f, g are measurable functions on Ω with values in $[0, \infty]$, and $\|f\|_p = \|g\|_q = 1$ then, by Hölder's Inequality,

$$0 \le 1 - \int_\Omega fg\, d\mu$$

with equality holding if and only if $f = g^{q-1}$. The question arises: how small is $f - g^{q-1}$ if $0 < 1 - \int_\Omega fg\, d\mu$? An answer to this is given by Clarkson's Inequalities [1]:

THEOREM (Clarkson). *Suppose $u, v \in L^p$.*

(i) *If $2 \le p < \infty$, then*

$$(24)\qquad \left\|\frac{u+v}{2}\right\|_p^p + \left\|\frac{u-v}{2}\right\|_p^p \le \frac{\|u\|_p^p + \|v\|_p^p}{2}.$$

(ii) *If $1 < p \le 2$, then*

$$(25)\qquad \left\|\frac{u+v}{2}\right\|_p^q + \left\|\frac{u-v}{2}\right\|_p^q \le \left(\frac{\|u\|_p^p + \|v\|_p^p}{2}\right)^{q-1}.$$

With the aid of these inequalities we provide an answer to the question posed.

LEMMA 1. *Let f, g be non-negative measurable functions on Ω with $\|f\|_p = \|g\|_q = 1$. Suppose that*

$$\varepsilon = 1 - \int_\Omega fg\, d\mu > 0.$$

(i) *If $2 \le p < \infty$, then*

$$\|f - g^{q-1}\|_p^p \le p2^{p-1}\varepsilon.$$

(ii) *If $1 < p < 2$, then*

$$\|f - g^{q-1}\|_p^q \le q2^{q-1}\varepsilon.$$

Proof. Let $v = g^{q-1}$. Then $v \in L^p(d\mu)$ and $\|v\|_p^p = \|g\|_q^q = 1$. By the Hahn–Banach Theorem we have

$$\begin{aligned}\left\|\frac{f+v}{2}\right\|_p &= \sup\left\{\frac{1}{2}\left|\int_\Omega (f+v)h\, d\mu\right| : h \in L^q(d\mu),\ \|h\|_q = 1\right\}\\ &\ge \frac{1}{2}\left|\int_\Omega (f+v)g\, d\mu\right| = \frac{1}{2}\left\{\int_\Omega fg\, d\mu + \int_\Omega g^q\, d\mu\right\}\\ &\ge \frac{1}{2}(2-\varepsilon) = 1 - \frac{\varepsilon}{2}.\end{aligned}$$

Now let $2 \le p < \infty$. We deduce from the last displayed inequality and (24) that

$$\left\|\frac{f-v}{2}\right\|_p^p \le 1 - \left(1-\frac{\varepsilon}{2}\right)^p \le \frac{p\varepsilon}{2},$$

from which (i) follows. The proof of (ii) is similar. This completes the proof of Lemma 1.

COROLLARY. *Let a, b be non-negative sequences such that $\sum a_i^p = \sum b_i^q = 1$ and*

$$1-\varepsilon \le \sum a_i b_i.$$

Then, if $2 \le p < \infty$, we have

$$\sum |a_i - b_i^{q-1}|^p \le p2^{p-1}\varepsilon,$$

while if $1 < p < 2$, we have

$$\Big(\sum |a_i - b_i^{q-1}|^p\Big)^{q/p} \le q2^{q-1}\varepsilon.$$

LEMMA 2. *There is a constant M depending on p, K only, independent of the values of x, y, f, such that $u \le M$ for all u satisfying* (13) *or* (13′).

Proof. Suppose otherwise and choose functions f_n and numbers $x_n, y_n \in E''$, $x_n < y_n$, such that the corresponding sequence $\{u_n\}$ tends to infinity. Then $x_n/y_n \to 0$ and $1/y_n = (1/x_n)(x_n/y_n) \le x_n/(Ny_n) \to 0$ as $n \to \infty$; furthermore,

$$\lim_n Y(u_n)^q = \beta_1.$$

Recall the inequality (13′) related to (13):

$$1 - 1/y_n \le \Big\{\Big(\int_0^{x_n} f_n^p\,dt\Big)^{1/p}(\beta_1 - Y(u_n)^q)^{1/q} + \Big(\int_{x_n}^{y_n} f_n^p\,dt\Big)^{1/p} Y(u_n) + \beta_2^{1/q}\Big(\int_{y_n}^{\infty} f_n^p\,dt\Big)^{1/p}\Big\}^q.$$

Thus the right hand side can be made arbitrarily close to one. We define finite sequences $(a_i(n))$, $(b_i(n))$, $1 \le i \le 3$, as follows:

$$a_1(n)^p = \int_0^{x_n} f_n^p\,dt,\quad a_2(n)^p = \int_{x_n}^{y_n} f_n^p\,dt,\quad a_3(n)^p = \int_{y_n}^{\infty} f_n^p\,dt,$$

$$b_1(n)^q = \beta_1 - Y(u_n)^q,\quad b_2(n) = Y(u_n),\quad b_3(n)^q = \beta_2,$$

and note that $\sum_{i=1}^3 a_i(n)^p = 1 = \sum_{i=1}^3 b_i(n)^q$, and $\sum_{i=1}^3 a_i(n)b_i(n) \ge 1 - 1/y_n$. Assuming first that $2 \le p < \infty$ and applying the Corollary we deduce that

$$\sum |a_i - b_i^{q-1}|^p \le A/y_n, \quad \text{and therefore} \quad \lim_n a_1(n) = 0.$$

But we already know that $\|f_n - v_{x_n}\|_p^p \le A/x_n \le A/N \le \beta_1/3$ if N is large enough. Since $\int_0^{x_n} v_{x_n}^p\,dt = \beta_1$, this gives a contradiction and establishes our claim in case $\beta_1 > 0$.

If $\beta_1 = 0$ then $\beta_2 = 1$ and $Y(u_n) = 0$ for all $n \ge 1$. The argument above yields $\int_{x_n}^{y_n} f_n^p\,dt \le A/y_n$. But the fact that $\|f_n - v_{x_n}\|_p^p \le A/x_n$ implies that $(\int_{x_n}^{y_n} v_{x_n}^p\,dt)^{1/p} \le (A/x_n)^{1/p} + (A/y_n)^{1/p}$. However,

$$\int_{x_n}^{y_n} v_{x_n}^p\,dt = \int_{x_n}^{y_n} \frac{1}{x_n} K(t/x_n, 1)^q\,dt = \int_1^{1+u_n} K(t,1)^q\,dt \to 1 \quad \text{as } n \to \infty.$$

These two estimates give a contradiction if we assume, as we may, that N is sufficiently large. The case $1 < p < 2$ follows similarly.

Next we show that if u, μ, f satisfy (13′) and $\mu \to 0$ then also $u \to 0$, independently of f.

LEMMA 3. *Suppose u, μ, f satisfy* (13′). *Then*

$$u^\alpha \le \begin{cases} c\mu^{1/p} & \text{if } 2 \le p < \infty, \\ c\mu^{1/q} & \text{if } 1 < p < 2. \end{cases}$$

Proof. If μ is bounded away from 0 then since u is bounded by Lemma 2, the inequality above is clearly satisfied. It suffices therefore to consider values of μ which are close to 0. Accordingly, we shall take a sequence $\{f_n\}$ of functions and real sequences $\{x_n\}, \{y_n\}$ with $x_n, y_n \in E_1''(f_n)$, $x_n < y_n$, such that (13′) is satisfied and $\lim_n x_n = \infty$. We assume $1 < p < 2$ and apply the Corollary above to (9) on writing

$$a_1(n)^p = \int_0^{x_n} f_n^p\,dt, \quad a_2(n)^p = \int_{x_n}^{\infty} f_n^p\,dt,$$

$$b_1^q = \beta_1, \quad b_2^q = \beta_2,$$

and we obtain $|(\int_0^{x_n} f_n^p\,dt)^{1/p} - \beta_1^{1/p}| \le c/x_n^{1/q}$, with a similar inequality with x_n replaced by y_n. From these we deduce that

$$\Big(\int_{x_n}^{y_n} f_n^p\,dt\Big)^{1/p} < [(\beta_1^{1/p} + c/y_n^{1/q})^p - (\beta_1^{1/p} - c/x_n^{1/q})^p]^{1/p} < c/x_n^{1/(pq)} = c\mu_n^{1/(pq)}.$$

It follows that $\lim_n \int_{x_n}^{y_n} f_n^p\,dt = 0$ and we note that the convergence to 0 depends only on μ_n. We now apply the Corollary once more, this time to

(13′) with the sequences

$$a_1(n)^p = \int_0^{x_n} f_n^p\,dt,\quad a_2(n)^p = \int_{x_n}^{y_n} f_n^p\,dt,\quad a_3(n)^p = \int_{y_n}^{\infty} f_n^p\,dt,$$

$$b_1(n)^q = \beta_1 - Y(u_n)^q,\quad b_2(n) = Y(u_n),\quad b_3(n)^q = \beta_2,$$

and we conclude that $|a_2(n) - b_2(n)^{q-1}| < c/y_n^{1/q}$. Consequently,

$$Y(u_n)^{q-1} \le c\mu_n^{1/(pq)} + c\mu_n^{1/q} \le c\mu_n^{1/(pq)}$$

or $Y(u_n)^q \le c\mu_n^{1/q}$. Since $Y(u_n)^q \ge Du_n^\alpha$ the result follows for this case. The case $2 \le p < \infty$ is similar. This completes the proof.

Remark. We note that the hypothesis (7) has been used above.

Our next step is to prove a relation between the η_i.

LEMMA 4. *The quantities η_i introduced in* (15) *satisfy*

$$(17)\qquad \begin{aligned}\eta_1\beta_1^{1/q} + \eta_2\beta_2^{1/q} &= Y^q/p - \tfrac{1}{2}(p-1)(\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p})\\ &\quad - (Y^{q-1}+\eta_3)^p/p + \eta_1 Y^q/(q\beta_1^{1/p})\\ &\quad + O(\eta_1^3, \eta_2^3, Y^{2q}, \eta_1^2 Y^q).\end{aligned}$$

Proof. We begin by setting down the consistency equation for the η_i:

$$(16)\qquad (\beta_2^{1/p} + \eta_2)^p + [(\beta_1 - Y(u)^q)^{1/p} + \eta_1]^p + (Y(u)^{q-1} + \eta_3)^p = 1.$$

The usual Taylor expansion gives

$$\begin{aligned}\beta_2\left[1 + p\eta_2/\beta_2^{1/p} + \binom{p}{2}(\eta_2/\beta_2^{1/p})^2 + O(\eta_2^3)\right] + (Y^{q-1}+\eta_3)^p&\\ + (\beta_1 - Y^q)[1 + \eta_1(\beta_1 - Y^q)^{-1/p}]^p &= 1.\end{aligned}$$

Also,

$$\begin{aligned}[1 + \eta_1(\beta_1 - Y^q)^{-1/p}]^p &= 1 + p\eta_1\beta_1^{-1/p}(1 - Y^q/\beta_1)^{-1/p}\\ &\quad + \binom{p}{2}\eta_1^2\beta_1^{-2/p}(1 - Y^q/\beta_1)^{-2/p} + O(\eta_1^3)\\ &= 1 + p\eta_1\beta_1^{-1/p}\{1 + Y^q/(p\beta_1) + O(Y^{2q})\}\\ &\quad + \binom{p}{2}\eta_1^2\beta_1^{-2/p}\left\{1 + \frac{2Y^q}{p\beta_1} + O(Y^{2q})\right\} + O(\eta_1^3)\\ &= 1 + p\eta_1\beta_1^{-1/p} + \eta_1 Y^q/\beta_1^{1+1/p}\\ &\quad + \binom{p}{2}\eta_1^2\beta_1^{-2/p} + \binom{p}{2}\frac{2\eta_1^2 Y^q}{p\beta_1^{1+2/p}} + O(Y^{2q}, \eta_1^3).\end{aligned}$$

Therefore, on multiplying this by $\beta_1 - Y^q$, we get on the right hand side,

$$\begin{aligned}\beta_1 + p\eta_1\beta_1^{1/q} - Y^q - \frac{p\eta_1 Y^q}{\beta_1^{1/p}} + \binom{p}{2}\eta_1^2\beta_1^{1-2/p}&\\ + \eta_1 Y^q/\beta_1^{1/p} + O(Y^{2q}, \eta_1^3, \eta_1^2 Y^q)&\\ = \beta_1 + p\eta_1\beta_1^{1/q} - Y^q - (p-1)\eta_1 Y^q/\beta_1^{1/p}&\\ + \binom{p}{2}\eta_1^2\beta_1^{1-2/p} + O(Y^{2q}, \eta_1^3, \eta_1^2 Y^q)&.\end{aligned}$$

Combining our results we get, from (16),

$$\begin{aligned}1 &= (\beta_1 + \beta_2) + p(\eta_1\beta_1^{1/q} + \eta_2\beta_2^{1/q}) - Y^q\\ &\quad + (Y^{q-1} + \eta_3)^p + \binom{p}{2}[\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p}]\\ &\quad - (p-1)\eta_1 Y^q/\beta_1^{1/p} + O(Y^{2q}, \eta_1^3, \eta_2^3, \eta_1^2 Y^q),\end{aligned}$$

whence

$$\begin{aligned}\eta_1\beta_1^{1/q} + \eta_2\beta_2^{1/q} &= Y^q/p - \tfrac{1}{2}(p-1)(\eta_1^2\beta_1^{1-2/p} + \eta_2^2\beta_2^{1-2/p})\\ &\quad - (Y^{q-1} + \eta_3)^p/p + \eta_1 Y^q/(q\beta_1^{1/p})\\ &\quad + O(Y^{2q}, \eta_2^3, \eta_1^3, \eta_1^2 Y^q),\end{aligned}$$

which is (17), the required result.

We turn next to the situation of approximate equality in the Arithmetic-Geometric Mean Inequality and an estimation of the closeness of the terms. We formulate the result in general terms.

LEMMA 5. *Let $1 < p < \infty$ and let q be the conjugate index to p. Suppose $a > b > 0$ and $0 < a/p + b/q - a^{1/p}b^{1/q} < \delta$. Then $a - b < \sqrt{3pqb\delta}$ when δ is small.*

Proof. Put $a = e^s$, $b = e^t$ where $s > t$. Then

$$\begin{aligned}e^s/p + e^t/q - e^{s/p+t/q} &= e^s\{1/p + e^{t-s}/q - e^{(t-s)/q}\}\\ &= e^s\left\{\frac{1}{p} + \frac{1 + (t-s) + (t-s)^2/2 + \ldots}{q}\right.\\ &\qquad \left.- \left(1 + (t-s)/q + ((t-s)/q)^2\frac{1}{2!} + \ldots\right)\right\}\\ &= e^s\left\{\frac{(t-s)^2}{2qp} + \frac{(t-s)^3}{3!q}(1 - 1/q^2) + O((t-s)^4)\right\}\\ &= a\left\{\frac{1}{2pq}\left(\log\frac{a}{b}\right)^2 - \frac{(1-1/q^2)(\log a/b)^3}{3!q} + O\left(\left(\log\frac{a}{b}\right)^4\right)\right\}\end{aligned}$$

$$= \left(\frac{a-b}{b}\right)^2 \frac{a}{2pq} - \left(\frac{a-b}{b}\right)^3 \frac{a[1+(1+1/q)/3]}{2pq} + O\left(\left(\frac{a-b}{b}\right)^4\right) > \frac{a((a-b)/b)^2}{3pq}$$

when $(a-b)/b$ is small enough. Since the left hand side is less than δ we get $((a-b)/b)^2 < 3pq\delta/a$ or $(a-b)/b < \sqrt{3pq\delta/a}$. The required result now follows.

COROLLARY. *If* (19) *holds then*

$$Y(u)^q - G_2(u)^q \le cu^\alpha/N^{1/(2p)}.$$

Proof. Let $a = Y^q$, $b = G_2^q$ in Lemma 5.

For the final lemma we take $1 < p < 2$ and we recall that $z = \eta_3 + Y^{q-1} \ge 0$, $F(z) = z^p - pYz + (p-1)Y^q$, and the inequality (23) holds, namely $F(z) \le p\mu + O(Y^{2q})$. Since $Y(u)^{2q} \le cu^\alpha u^\alpha \le cu^\alpha \mu^{1/q}$ by Lemma 3, and since $\mu \le u/N$, we may rewrite (23) as follows:

$$F(z) \le cu^\alpha/N^{1/q}. \tag{23}$$

LEMMA 6. *Under the above conditions there is a constant* c *depending only on* p, *such that* $|\eta_3| \le cu^{\alpha/p}/N^{1/(2q)}$.

Proof. It is readily checked that $F(z)$ has a local minimum at $z = Y^{q-1}$, that is, when $\eta_3 = 0$, and that $F(z) = 0$ at this point. Since, according to (23), F is small, we expect that z is close to the minimum value and we shall therefore write $z = (1+\delta)Y^{q-1}$. Consequently,

$$F(z) = Y^q\{(1+\delta)^p - p(1+\delta) + p - 1\} = Y^q L(\delta),$$

say. From (23), $Y^q L(\delta) \le cu^\alpha/N^{1/q}$. Since, according to our assumption, $Y^q \ge Du^\alpha$, it follows that $L(\delta)$ must be small. Elementary calculus shows that $L(\delta)$ has a local, indeed an absolute minimum, at $\delta = 0$, with value 0, and is strictly decreasing to the left and strictly increasing to the right of 0. This implies that δ can be made arbitrarily small depending on N. We may therefore expand $L(\delta)$ to obtain

$$Y^q\left\{\binom{p}{2}\delta^2 + O(\delta^3)\right\} \le cu^\alpha/N^{1/q}.$$

There is therefore a constant c such that $\delta^2 < c/N^{1/q}$, i.e. $|\delta| \le c/N^{1/(2q)}$, and so $|\eta_3| = |\delta Y^{q-1}| \le cu^{\alpha/p}/N^{1/(2q)}$. The proof is complete.

This completes the proofs of the auxiliary lemmas.

Some applications. We present here a few applications to demonstrate the power of Theorem 3.

I. The kernel for fractional integration of order r with $1/p < r \le 1$ is

$$K(t,x) = \begin{cases} (x-t)^{r-1}x^{-r} & \text{if } t < x, \\ 0 & \text{if } t > x. \end{cases}$$

We require $r > 1/p$ to ensure that $H(x) = H(f,x)$ is defined for all $f \in L^p$. We then have, with our earlier notation,

$$H(x) = x \int_0^x x^{-r}(x-t)^{r-1} f(t)\,dt,$$

while

$$\psi = \int_0^\infty K(u,1)^q\,du = \int_0^1 (1-u)^{q(r-1)}\,du = \frac{1}{1-q(1-r)} = \frac{1}{\alpha},$$

where $0 < \alpha \le 1$. In this case

$$G_1(u)^q = \int_{1-u}^1 K(t,1)^q\,dt = \int_{1-u}^1 (1-t)^{q(r-1)}\,dt = u^\alpha/\alpha,$$

while $G_2 = 0$. The condition (7) is satisfied with $\alpha = 1 - q(1-r)$. While Theorem 3 of course applies and (8) follows, it is worth pointing out that the analysis simplifies enormously and a much simpler proof can be given. We note that $\beta_2 = 0$ while (9) gives

$$x - 1 \le \alpha x\left\{(1/\alpha)^{1/q}\left(\int_0^x f^p\,dt\right)^{1/p}\right\}^q = x\left(\int_0^x f^p\,dt\right)^{q-1},$$

from which we deduce that $\int_x^\infty f^p\,dt \le c/x$. This inequality can be used instead of (11) and Clarkson's Inequality is not needed. The rest follows routinely. The case $r = 1$ gives the original inequality of Moser.

II. We consider the operator adjoint to the previous one. With the same kernel as in I, the integral operator in this case is, for $f \in L^p$,

$$H(x) = x\int_0^\infty K(x,t) f(t)\,dt = x\int_x^\infty t^{-r}(t-x)^{r-1} f(t)\,dt;$$

also,

$$\psi = \int_0^\infty K(1,u)^q\,du = \int_1^\infty (u-1)^{q(r-1)} u^{-rq}\,du = B(\alpha, q-1).$$

Again (7) is satisfied with $G_1(u) = 0$ and

$$G_2(u)^q = \int_1^{1+u} (t-1)^{q(r-1)} t^{-rq}\, dt \sim u^\alpha/\alpha \quad (u \to 0).$$

The result therefore holds for this kernel.

We remark that if $r = 1$ then $\alpha = 1$ and $\psi^{-1} = (B(1, q-1))^{-1} = q - 1$, which equals 1 if and only if $q = 2$. In other words, the following dual to the original Moser Inequality holds:

Let $H(x) = x \int_x^\infty (f(t)/t)\, dt$. There is a constant C_p such that

$$\int_0^\infty \exp\{(q-1)H(x)^q - x\}\, dx \le C_p$$

for all $f \in B_p$.

One might wonder as to what can be said if the kernels in I and II were combined. Thus define

$$K(t,x) = \begin{cases} (x-t)^{r-1} x^{-r} & \text{if } 0 < t < x, \\ (t-x)^{r-1} t^{-r} & \text{if } t > x. \end{cases}$$

We readily infer from what has been said above that $G_1(u)^q \sim G_2(u)^q$, as $u \to 0$, and therefore (7) does not hold. However, a slight modification of the kernel does work.

III. Define

$$K(t,x) = \begin{cases} d(x-t)^{r-1} x^{-r} & \text{if } 0 < t < x, \\ (t-x)^{r-1} t^{-r} & \text{if } t > x, \end{cases}$$

where d is positive and unequal to 1. It follows that $\psi = \int_0^\infty K(t,1)^q\, dt = d^q/\alpha + B(\alpha, q-1)$ and $|G_1(u)^q - G_2(u)^q| \sim |d^q - 1| u^\alpha/\alpha$ as $u \to 0$, so that (7) holds. The result (8) is true in consequence.

A counter-example. We show next that some lack of smoothness on the kernel K is necessary in order that (8) holds. To see this we shall take a smooth kernel $K(t,y) = 1/(t+y)$, $y, t > 0$, which does not satisfy (7), but does have all the other properties required of it, and shall show that (8) does not hold for K.

Suppose $q = 2$ so that $\psi = \int_0^\infty (1+u)^{-2}\, du = 1$. We fix $x > 0$ for the moment and take a particular function f, namely, the normalized extremal at x, $f(t) = \sqrt{x}/(t+x)$, $\|f\|_2 = 1$. Then

$$H(y) = y \int_0^\infty \frac{f(t)}{t+y}\, dt = y\sqrt{x} \int_0^\infty \frac{1}{(t+y)(t+x)}\, dt = \frac{y\sqrt{x}}{x-y} \log \frac{x}{y} \tag{26}$$

on using partial fractions and evaluating. We wish to show that

$$\int_0^\infty \exp\{H(y)^2 - y\}\, dy = \int_0^\infty \exp\left\{ \frac{y^2 x}{(x-y)^2} \left(\log \frac{x}{y} \right)^2 - y \right\} dy \tag{27}$$

is unbounded as $x \to \infty$. Under the change of variable $y = xs$, the right hand side of (27) becomes

$$x \int_0^\infty \exp\left\{ x \left[\frac{s^2}{(1-s)^2} \left(\log \frac{1}{s} \right)^2 - s \right] \right\} ds.$$

For our purpose it is enough to consider values of s in the range $0 < s < 1$ for which $(1-s)/s \le \varepsilon$ where ε is fixed and small, say $\varepsilon < 1/10$. In other words, s lies in the interval $[1/(1+\varepsilon), 1)$. The Taylor expansion for the logarithm yields

$$\log \frac{1}{s} = \log \left(1 + \frac{1-s}{s} \right) = \frac{1-s}{s} - \frac{1}{2} \left(\frac{1-s}{s} \right)^2 + \frac{1}{3} \left(\frac{1-s}{s} \right)^3 + O\left(\left(\frac{1-s}{s} \right)^4 \right),$$

and so

$$\frac{\log(1/s)}{1-s} = \frac{1}{s} \left\{ 1 - \frac{1}{2} \left(\frac{1-s}{s} \right) + \frac{1}{3} \left(\frac{1-s}{s} \right)^2 + O\left(\left(\frac{1-s}{s} \right)^3 \right) \right\}.$$

Hence, on squaring,

$$\left(\frac{\log(1/s)}{1-s} \right)^2 = \frac{1}{s^2} \left\{ 1 - \frac{1-s}{s} + \left(\frac{1-s}{s} \right)^2 \left(\frac{2}{3} + \frac{1}{4} \right) + O\left(\left(\frac{1-s}{s} \right)^3 \right) \right\} > \frac{1}{s^2} \left\{ 1 - \frac{1-s}{s} + \frac{1}{2} \left(\frac{1-s}{s} \right)^2 \right\}$$

for the values of ε considered. From (27), therefore, we have

$$\int_0^\infty \exp\{H(y)^2 - y\}\, dy > x \int_{1/(1+\varepsilon)}^1 \exp\left\{ x \left(1 - \frac{1-s}{s} + \frac{1}{2} \left(\frac{1-s}{s} \right)^2 - s \right) \right\} ds > x \int_{1/(1+\varepsilon)}^1 \exp\{x(1-s)(s-1)/s\}\, ds = I,$$

say. Make a change of variable, $u=(1-s)\sqrt{x}$, so that

$$I=\sqrt{x}\int_0^{\varepsilon\sqrt{x}/(1+\varepsilon)}\exp(-u^2/s)\,du.$$

Now $u^2/s\le u^2(1+\varepsilon)$, and consequently,

$$I>\sqrt{x}\int_0^{\varepsilon\sqrt{x}/(1+\varepsilon)}\exp(-u^2(1+\varepsilon))\,du.$$

Letting $v=u\sqrt{(1+\varepsilon)}$, whence $0\le v\le\varepsilon\sqrt{x/(1+\varepsilon)}$, we have

$$I>\Big(\int_0^{\varepsilon\sqrt{x/(1+\varepsilon)}}\exp(-v^2)\,dv\Big)\sqrt{x/(1+\varepsilon)}\to\infty\quad\text{as } x\to\infty,$$

because the integral on the right hand side tends to $\sqrt{\pi}$ since ε is fixed. It follows that the left hand side of (27) tends to ∞ at least as fast as $c\sqrt{x}$ and the assertion is proved.

A similar argument works for other values of q. We take $f(t)=A^{1/p}/(t+x)^{q-1}$ where $A=(q-1)x^{q-1}$, so that $\|f\|_p=1$. Then,

$$\int_0^\infty\frac{f(t)}{t+y}\,dt=A^{1/p}\int_0^\infty\frac{dt}{(t+x)^{q-1}(t+y)}$$
$$=\frac{A^{1/p}}{yx^{q-2}(q-1)}F(1,1,q,1-x/y),$$

where we have recourse to a hypergeometric function for the evaluation of the integral [5]. We also note that $\psi=1/(q-1)$. We choose values of y for which y/x is close to 1, expand $F(1,1,q,1-x/y)$ by the hypergeometric series, proceed as above, and the required result follows.

The sharpness of the constant. At this point we come back to the question whether the constant ψ^{-1} in (8) is sharp. In considering this question we looked at the case of the kernel of fractional integration of order r. When $r=1$, the original case of Moser, we know it is sharp. Our investigation showed that it is sharp also for general r, $1/p<r<1$, and that this depends on a special property of the fractional integral operator. Specifically, if $f\in L^p$ and if

$$f_r(x)=\int_0^x(x-t)^{r-1}f(t)\,dt,$$

then Hardy and Littlewood [6] showed that f_r belongs to the little Lipschitz class λ_δ with $\delta=r-1/p$. Using this property it is not hard to prove our assertion. Moving now to the general case, we would like to state the property we require explicitly in terms of the kernel K. Accordingly, we impose the following integral-mean condition on K: There exist positive constants M,λ_0,δ with $\lambda_0<1$ such that

$$\Big\{\int_0^\infty|K(u,1)-K(\lambda u,1)|^q\,du\Big\}^{1/q}\le M(1-\lambda)^\delta,\quad\lambda_0\le\lambda\le1.\tag{28}$$

THEOREM 4. *If, in addition, the kernel K satisfies the condition* (28) *then the constant $1/\psi$ in* (8) *is best possible.*

Proof. Suppose that $x>0$ is fixed for the moment and that f is the normalized extremal at x: $f(t)=AK(t,x)^{q-1}$ where $A^p=x^{q-1}/\psi$. With $H(y)=Ay\int_0^\infty K(t,y)K(t,x)^{q-1}\,dt$, we have $H(x)=Ax\psi/x^{q-1}=(x^{q-1}/\psi)^{1/p}\psi x/x^{q-1}=(x\psi)^{1/q}$. It can now be shown that there exists $M_1>0$ such that

$$|H(y)-H(x)|\le M_1x^{1/q}\Big|\frac{y-x}{x}\Big|^\delta,\quad x/2<y<x.\tag{29}$$

Assuming for the moment that (29) holds, and, without loss of generality, that $\psi=1$, we proceed to prove our assertion. By (29) we have, for $x/2<y<x$,

$$H(y)\ge H(x)-M_1x^{1/q}\Big|\frac{y-x}{x}\Big|^\delta.$$

Suppose now that b is any number greater than one. We then have

$$bH(y)^q-y\ge bx\Big(1-M_1\Big|\frac{y-x}{x}\Big|^\delta\Big)^q-y\ge0,$$

provided that $1-M_1|(y-x)/x|^\delta\ge b^{-1/q}(y/x)^{1/q}$. That this is possible for y close to x is a consequence of the following argument: Let

$$F(u)=1-M_1(1-u)^\delta-b^{-1/q}u^{1/q},\quad 0<u\le1.$$

Since F is continuous and $F(1)=1-b^{-1/q}>0$, it is clear that there exists $u_0<1$, depending only on F, such that

$$F(u)\ge F(1)/2>0\quad\text{for } u_0<u\le1.$$

Consequently, for $u_0x<y<x$, $bH(y)^q-y>0$, and

$$\int_{u_0x}^x\exp[bH(y)^q-y]\,dy\ge x(1-u_0).$$

Since u_0 is independent of x the integral tends to infinity as $x\to\infty$. We have shown that (8) is not valid for arbitrary $b>1$ and therefore that the constant is best possible.

It remains to prove (29). We therefore take $x/2 \le y \le x$ and once more, without loss of generality, $\psi = 1$. We have, on substituting $u = t/y$,

$$\begin{aligned} H(y) - H(x) &= \frac{y}{x^{1/p}} \int_0^\infty K(u,1)K(uy/x,1)^{q-1}\,du - \frac{x}{x^{1/p}} \int_0^\infty K(u,1)^q\,du \\ &= \frac{y}{x^{1/p}} \int_0^\infty \{K(u,1)K(uy/x,1)^{q-1} - K(uy/x,1)^q\}\,du \\ &\quad + \frac{y}{x^{1/p}} \int_0^\infty \{K(uy/x,1)^q - K(u,1)^q\}\,du \\ &\quad + \frac{y-x}{x^{1/p}} \int_0^\infty K(u,1)^q\,du \\ &= I_1 + I_2 + I_3. \end{aligned}$$

By Hölder's Inequality,

$$\begin{aligned} |I_1| &\le \frac{y}{x^{1/p}} \int_0^\infty K(uy/x,1)^{q-1}|K(u,1) - K(uy/x,1)|\,du \\ &\le \frac{y}{x^{1/p}} \Big\{ \int_0^\infty K(uy/x,1)^q\,du \Big\}^{1/p} \Big\{ \int_0^\infty |K(u,1) - K(uy/x,1)|^q\,du \Big\}^{1/q} \\ &\le \frac{y}{x^{1/p}} \Big\{ \frac{x}{y} \int_0^\infty K(v,1)^q\,dv \Big\}^{1/p} M \Big|\frac{x-y}{x}\Big|^\delta \le M y^{1/q} \Big|\frac{x-y}{x}\Big|^\delta. \end{aligned}$$

Next,

$$|I_3| \le \Big|\frac{x-y}{x^{1/p}}\Big| = x^{1/q}\Big|\frac{x-y}{x}\Big|.$$

To handle I_2, we first observe that the inequality $a^q - b^q \le q(a-b)a^{q-1}$ holds for $0 < b < a$. Now,

$$\begin{aligned} \int_0^\infty |K(u,1)^q - K(uy/x,1)^q|\,du &\le q \int_0^\infty |K(u,1) - K(uy/x,1)| \\ &\qquad \times \max\{K(u,1)^{q-1}, K(uy/x,1)^{q-1}\}\,du \\ &\le q\Big\{ \int_0^\infty |K(u,1) - K(uy/x,1)|^q\,du \Big\}^{1/q} \\ &\qquad \times \Big\{ \int_0^\infty [K(u,1)^q + K(uy/x,1)^q]\,du \Big\}^{1/p} \\ &\le 3^{1/p} q M \Big|\frac{y-x}{x}\Big|^\delta, \end{aligned}$$

since $x/y \le 2$. In other words,

$$|I_2| \le 3^{1/p} q M x^{1/q} \Big|\frac{y-x}{x}\Big|^\delta.$$

Combining these estimates and assuming $0 < \delta \le 1$, we get

$$|H(y) - H(x)| \le (M + 3^{1/p} q M + 1) x^{1/q} \Big|\frac{y-x}{x}\Big|^\delta,$$

which is (29). This completes the proof.

Remarks. 1. The condition (28) is satisfied for the case of the fractional integral operator. In fact, the argument of Hardy–Littlewood referred to above can be used virtually unchanged to show that (28) holds with $\delta = r - 1/p$.

2. The assumption (28) makes for some redundancy in the hypothesis of Theorem 3. For then we have $G_1(u)^q = O(u^\alpha)$ as $u \to 0$ with $\alpha = \delta$, and similarly for G_2. To see this, let $\lambda < 1, u = 1 - \lambda$. Then $\int_{1-u}^1 K(t,1)^q\,dt = \int_0^1 [K(t,1)^q - \lambda K(\lambda t,1)^q]\,dt$. Now the estimate for I_2 obtained above gives

$$\int_0^1 |K(t,1)^q - K(\lambda t,1)^q|\,dt \le M_2(1-\lambda)^\delta,$$

for some M_2. Consequently,

$$\begin{aligned} \int_0^1 [K(t,1)^q - \lambda K(\lambda t,1)^q]\,dt &= \int_0^1 [K(t,1)^q - K(\lambda t,1)^q]\,dt \\ &\quad + \int_0^1 [K(\lambda t,1)^q - \lambda K(\lambda t,1)^q]\,dt, \end{aligned}$$

and taking absolute values,

$$\int_{1-u}^1 K(t,1)^q\,dt \le M_2(1-\lambda)^\delta + (1-\lambda)2\beta_1 \le M_3(1-\lambda)^\delta = M_3 u^\delta,$$

on supposing, as we may, that $1/2 \le \lambda < 1$. This gives the desired estimate for G_1.

3. The kernel $K(t,y) = 1/(t+y)$, $y, t > 0$ satisfies (28) (with $\delta = 1$, $\lambda_0 = 1/2$, when $p = 2$) but (7) does not hold.

Acknowledgments. We would like to thank the referee for helpful comments.

References

[1] D. R. Adams, *A sharp inequality of J. Moser for higher order derivatives*, Ann. of Math. 128 (1988), 385–398.
[2] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
[3] L. Carleson and S. Y. A. Chang, *On the existence of an extremal function for an inequality of J. Moser*, Bull. Sci. Math. (2) 110 (1986), 113–127.
[4] S. Y. A. Chang, *Extremal functions in a sharp form of Sobolev inequality*, in: Proc. Internat. Congress of Mathematicians, Berkeley, Calif., 1986.
[5] I. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, translated from the fourth Russian ed., Scripta Technica, Academic Press, 1965.
[6] G. H. Hardy and J. E. Littlewood, *Some properties of fractional integrals I*, Math. Z. 27 (1928), 565–606.
[7] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge Univ. Press, Cambridge, 1967.
[8] M. Jodeit, *An inequality for the indefinite integral of a function in L^q*, Studia Math. 44 (1972), 545–554.
[9] D. E. Marshall, *A new proof of a sharp inequality concerning the Dirichlet integral*, Ark. Mat. 27 (1989), 131–137.
[10] P. McCarthy, *A sharp inequality related to Moser's inequality*, Mathematika 40 (1993), 357–366.
[11] J. Moser, *A sharp form of an inequality by N. Trudinger*, Indiana Univ. Math. J. 20 (1971), 1077–1092.
[12] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

DEPARTMENT OF MATHEMATICS
UNIVERSITY COLLEGE
CORK, IRELAND

DEPARTMENT OF MATHEMATICS
MAYNOOTH COLLEGE
CO. KILDARE, IRELAND

Received March 2, 1994
Revised version September 27, 1994
(3240)

STUDIA MATHEMATICA 113 (2) (1995)

The stability radius of an operator of Saphar type

by

CHRISTOPH SCHMOEGER (Karlsruhe)

Abstract. A bounded linear operator T on a complex Banach space X is called an operator of Saphar type if its kernel is contained in its generalized range $\bigcap_{n=1}^{\infty} T^n(X)$ and T is relatively regular. For T of Saphar type we determine the supremum of all positive numbers δ such that $T-\lambda I$ is of Saphar type for $|\lambda|<\delta$.

I. Terminology and introduction. Throughout this paper let X denote a Banach space over the complex field $\mathbb{C}$ and let $\mathcal{L}(X)$ denote the algebra of all bounded linear operators on X. If $T \in \mathcal{L}(X)$, we denote by $N(T)$ the kernel and by $T(X)$ the range of T. The *generalized range* of T is defined by

$$T^{\infty}(X) = \bigcap_{n=1}^{\infty} T^n(X).$$

We write $\sigma(T)$ for the spectrum of T and $\varrho(T)$ for the resolvent set $\mathbb{C}\setminus\sigma(T)$. The spectral radius of T is denoted by $r(T)$.

In [6, Theorem 3] T. Kato showed that for T in $\mathcal{L}(X)$ the set

$$\varrho_{\mathrm{K}}(T) = \{\lambda \in \mathbb{C} : (T-\lambda I)(X) \text{ is closed and } N(T-\lambda I) \subseteq (T-\lambda I)^{\infty}(X)\}$$

is an open subset of $\mathbb{C}$. Since $\varrho(T) \subseteq \varrho_{\mathrm{K}}(T)$, the complement $\sigma_{\mathrm{K}}(T) = \mathbb{C}\setminus \varrho_{\mathrm{K}}(T)$ is a compact subset of $\sigma(T)$. We showed in [10, Satz 2] that $\partial\sigma(T) \subseteq \sigma_{\mathrm{K}}(T)$, thus $\sigma_{\mathrm{K}}(T) \neq \emptyset$.

We call $T \in \mathcal{L}(X)$ *relatively regular* if $TST = T$ for some $S \in \mathcal{L}(X)$. In this case TS is a projection on $T(X)$ (hence $T(X)$ is closed), $I - ST$ is a projection on $N(T)$, and we say that S is a *pseudo-inverse* of T.

T is called an *operator of Saphar type* if T is relatively regular and $N(T) \subseteq T^{\infty}(X)$. This class of operators has been studied by P. Saphar [9] (see also [2] and [12]). Operators in this class have an important property:

THEOREM 1. *$T \in \mathcal{L}(X)$ is of Saphar type if and only if there is a neighbourhood $U \subseteq \mathbb{C}$ of 0 and a holomorphic function $F : U \to \mathcal{L}(X)$ such*

1991 *Mathematics Subject Classification*: Primary 47A10.