SMOOTHNESS OF EXTREMIZERS OF A CONVOLUTION INEQUALITY

MICHAEL CHRIST AND QINGYING XUE

Abstract. Let \( d \geq 2 \) and \( T \) be the convolution operator
\[
T f(x) = \int_{\mathbb{R}^{d-1}} f(x' - t, x_d - |t|^2) \, dt,
\]
which is bounded from \( L^{(d+1)/d}(\mathbb{R}^d) \) to \( L^{d+1}(\mathbb{R}^d) \). We show that any critical point \( f \in L^{(d+1)/d}(\mathbb{R}^d) \) of the functional \( \|T f\|_{d+1}/\|f\|_{(d+1)/d} \) is infinitely differentiable, and that \( |x|^{\delta}f \in L^{(d+1)/d} \) for some \( \delta > 0 \). In particular, this holds for all extremizers of the associated inequality. This is done by exploiting a generalized Euler-Lagrange equation, and certain weighted norm inequalities for \( T \).

1. Introduction

Optimal constants and extremizers have been determined for some of the most fundamental \( L^p \) inequalities of Fourier and real analysis. Among such achievements are the celebrated works of Beckner [2], Burkholder [5], Lieb [16],[17], and Pichorides [18]. Certain multilinear inequalities, governed by linear geometric structure, have more recently been treated in [4]. Still more recently, optimal constants and extremizers have been determined for Fourier restriction/extension inequalities for paraboloids, in the lowest dimensions, in works of Foschi [13], Hundertmark and Zharnitsky [14], and Bennett, Bez, Carbery, and Hundertmark [3]. The geometry which underlies restriction inequalities features curvature.

The present paper is one of a series [10],[11],[9],[8],[7],[6],[20] which treat questions concerning extremals for certain \( L^p \) norm inequalities, whose form is determined by the influence of curvature and singularities. These works focus on less fine questions such as the existence of extremizers, precompactness of extremizing sequences, and qualitative and quantitative properties of extremizers. The present paper is concerned with such properties of extremizers, for one particular inequality.

Let \( d \geq 2 \). Points of \( \mathbb{R}^d \) will be represented as \( x = (x', x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}^1 \). Our object of investigation is the convolution operator
\[
T f(x) = \int_{\mathbb{R}^{d-1}} f(x' - t, x_d - |t|^2) \, dt.
\]
This operator is bounded from $L^{(d+1)/d}(\mathbb{R}^d)$ to $L^{d+1}(\mathbb{R}^d)$, and satisfies no other $L^p \to L^q$ inequalities. The curvature of the parabola $x_d = |x'|^2$ and scaling symmetry of the measure $dx = |x'|^2$ are the crucial ingredients in this theory. This operator $T$ enjoys a rich symmetry structure discussed in [6], and is perhaps the most prototypical representative of the class of operators $f \mapsto f \ast \mu$, where $\mu$ is a measure supported on a nonflat submanifold of $\mathbb{R}^d$.

Let $p_0 = p_0(d) = (d + 1)/d$ and $q_0 = q_0(d) = d + 1$. Denote by $A_d$ the optimal constant in the inequality
\begin{equation}
(1.1) \quad \|Tf\|_{q_0} \leq A_d\|f\|_{p_0}.
\end{equation}

An $\varepsilon$-quasiextremal for inequality (1.1) is a function satisfying $\|Tf\|_{q_0} \geq \varepsilon\|f\|_{p_0}$. A characterization of quasiextremals is established in [6], which includes some quantitative though non-optimal control as $\varepsilon \to 0$. It is shown in [7] that extremizers for the inequality (1.1) exist, and that any nonnegative extremizing sequence of functions is precompact modulo action of the group of all geometric symmetries of the inequality.

In the present paper we take a third step by establishing two properties of extremizers: smoothness, and some improved decay. These are established for all critical points of the functional $\|Tf\|_{q_0}/\|f\|_{p_0}$. We formulate a conjecture concerning the precise decay rate of nonnegative extremizers. The extremizers and optimal constant $A_d$ remain unknown, and it remains unknown whether extremizers are unique modulo natural symmetries.

A technical device which underlies the analysis, and which may be of some independent interest, is a family of weighted norm inequalities for $T$. Lemma 2.2 formulates a one parameter family of rather sharp weighted inequalities. These involve pairs of exponents different from $(p_0, q_0)$, are not consequences of (1.1), and are suited to our purpose.

The transpose $T^*$ of $T$ takes the form
\begin{align*}
T^* f(x) &= \int_{\mathbb{R}^{d-1}} f(x' + t, x_d + |t|^2) \, dt = \int_{\mathbb{R}^{d-1}} f(x' - t, x_d + |t|^2) \, dt.
\end{align*}

$T^*$ is equal to $T$ conjugated with the norm-preserving operator associated to the transformation $(y', y_d) \mapsto (y', -y_d)$ of $\mathbb{R}^d$. In particular, $T^*$ is likewise bounded from $L^{(d+1)/d}(\mathbb{R}^d)$ to $L^{d+1}(\mathbb{R}^d)$.

Real-valued critical points of the functional $\|Tf\|_{q_0}/\|f\|_{p_0}$ are characterized by the generalized Euler-Lagrange equation
\begin{equation}
(1.2) \quad f = \lambda \left( T^* [(Tf)^d] \right)^{d/d_0}
\end{equation}

where
\begin{equation}
(1.3) \quad \lambda = A_d^{-d_0} \|f\|_{p_0}^{d_0 - d_0}.
\end{equation}

Complex-valued critical points are characterized by this same equation (1.2) with $\lambda = \|Tf\|_{q_0}^{-d_0}/\|f\|_{p_0}^{d_0}$, provided that powers of complex numbers on the right-hand side of (1.2) are interpreted as follows: If $z \in \mathbb{C}$ and $0 \neq s \in \mathbb{R}$, then $z^s$ should be interpreted as $|z|^s$. When $s = d$ is an even integer, this is not a product of positive integer powers of $z$ and $\bar{z}$.
The main result of this paper is:

**Theorem 1.1.** Let $d \geq 2$, and let $\lambda \in \mathbb{R}$. Let $f \in L^{p_0(d)}(\mathbb{R}^d)$ be any real-valued solution of the generalized Euler-Lagrange equation (1.2). Then $f \in C^\infty$. Moreover, all partial derivatives of $f$ are bounded functions, and there exists $\delta > 0$ such that $(1 + |x|)^{\delta} \nabla f(x) \in L^{p_0}(\mathbb{R}^d)$ for all $k \geq 0$.

The same conclusion holds for all complex-valued solutions if $d$ is even and $\lambda \in \mathbb{C}$.

This means, of course, that there exists a $C^\infty$ function which is equal almost everywhere to $f$.

Inequality (1.1) is invariant under parabolic scaling. There are no a priori inequalities which assert that if $f \in L^{p_0}$, then $S(f) = (T^*[(Tf)^d])^d$ has additional decay or smoothness properties, which would lead to a simple proof of the theorem via a bootstrapping argument. Instead, we will analyze the linearization (and all but the highest order terms in its finite Taylor series) of the multilinear operator $S$ about a dense class of functions, and will show that these operators do improve decay. The key in using this fact in conjunction with (1.2) and a fixed-point argument, is to find Banach spaces which encode more rapid decay than does $L^{(d+1)/d}$, and which are preserved by $S$. We do this by developing a limited theory of weighted inequalities for $T$.

We will demonstrate Theorem 1.1 only in the real case. The same reasoning applies to the complex case, with small and straightforward modifications in formulas to accommodate various complex conjugations.

Define

\begin{align*}
(1.4) & \quad v(x) = \min \left(1, |x'|^{-d}, |x_d - |x'|^2|^{-d}\right) \\
(1.5) & \quad v_*(x) = \min \left(1, |x'|^{-d}, |x_d + |x'|^2|^{-d}\right)
\end{align*}

The functions $v, v_*$ are $O(|x|^{-\delta})$ as $|x| \to \infty$ for some $\delta > 0$, and have another noteworthy aspect. Consider the parabolic dilations $\delta_r(x', x_d) = (rx', r^2x_d)$, for $r > 0$, and the associated operators $\delta_r(f)(x') = f(\delta_r(x'))$. With respect to these dilations, $T$ enjoys the symmetry $\delta_r(Tf) = r^{d-1}T(\delta_r(f))$. The weight $v$ equals $w^{-d}$ where $w(x)$ is the maximum of the three quantities $w_0(x) = 1$, $w_1(x) = |x'|$ and $w_2(x) = |x_d - |x'|^2|$. Each is homogeneous with respect to the dilations $\delta_r$, but $w_j$ is homogeneous of degree $j$ for $j = 0, 1, 2$. We believe that $v$ accurately expresses the behavior of extremals, in the following sense.

**Conjecture 1.2.** Let $\lambda \in \mathbb{C}$, and let $f \in L^{p_0(d)}(\mathbb{R}^d)$ be any solution of the generalized Euler-Lagrange equation (1.2).

(i) There exists $C = C(f, \lambda) < \infty$ such that for almost every $x \in \mathbb{R}^d$, $f(x) \leq C v(x)$.

(ii) If $\lambda > 0$, $f$ is nonnegative and $\|f\|_{p_0(d)} > 0$, then there exists $c = c(f, \lambda) > 0$ such that for almost every $x \in \mathbb{R}^d$, $f(x) \geq c v(x)$.

---

1 The three are related: $|w_j(x) - w_j(\tilde{x})| = O(w_{j-1}(x))$ if $|x - \tilde{x}| = O(1)$, for $j = 1, 2$. 

---

**SMOOTHNESS OF EXTREMIZERS 3**
2. Weighted inequalities

The following elementary inequalities provide a foundation for our analysis. The proof of Lemma 2.1, deferred to §8, is thoroughly elementary but is not short.

**Lemma 2.1.** For each \( d \geq 2 \) there exists \( C_d < \infty \) such that

\[
T(v^*) \leq C_d v^{1/d}
\]

\[
T^*(v) \leq C_d v_1^{1/d}.
\]

For \( \theta \in [0, 1] \) define exponents \( p_\theta, q_\theta \) to be

\[
p^{-1}_\theta = (1 - \theta)p_0(d)^{-1}
\]

\[
q^{-1}_\theta = (1 - \theta)q_0(d)^{-1}.
\]

Then \( q_\theta = p_\theta/d \).

**Lemma 2.2 (Weighted Inequalities).** There exists \( C < \infty \) such that for every \( t \in [0, 1] \), for every nonnegative function \( f \),

\[
\left( \int_{\mathbb{R}^d} (Tf)^q \cdot u^{-tq/d} \right)^{1/q} \leq C \left( \int_{\mathbb{R}^d} f^{p_0} \cdot u^{-tp_0} \right)^{1/p_0}.
\]

\[
\left( \int_{\mathbb{R}^d} (T^*f)^q \cdot u^{-tq/d} \right)^{1/q} \leq C \left( \int_{\mathbb{R}^d} f^{p_0} \cdot u^{-tp_0} \right)^{1/p_0}.
\]

**Proof.** Consider the analytic family of operators \( T_z f = v^{-z/d} T(v^*_z f) \) on the strip \( \{ z : 0 \leq \text{Re}(z) \leq 1 \} \). When \( \text{Re}(z) = 0 \), \( T_z \) is bounded from \( L^{p_0} \) to \( L^{q_0} \). When \( \text{Re}(z) = 1 \), \( T_z \) is bounded from \( L^\infty \) to \( L^\infty \), by Lemma 2.1. Both bounds are uniform in \( \text{Im}(z) \). The first conclusion of Lemma 2.2 follows by complex interpolation. The inequality for \( T^* \) is proved in the same way. \( \square \)

Lemma 2.2 has consequences which are conveniently expressed in terms of certain nonlinear operators and certain function spaces. Define nonlinear operators

\[
\mathcal{T}(f) = (Tf)^d,
\]

\[
\mathcal{T}_*(f) = (T^*f)^d,
\]

\[
S(f) = \mathcal{T}_*(\mathcal{T}(f)).
\]

In these terms, the Euler-Lagrange equation (1.2) becomes \( f = \lambda S(f) \).

Set

\[
w = v^{-1}_* \text{ and } w_* = v^{-1}.
\]

The following scales of Banach spaces \( X_\theta, X_\theta^*, Y_\theta, Y_\theta^* \) are adapted to \( T \) and \( T^* \). Define these spaces to be the sets of all equivalence class of measurable functions on \( \mathbb{R}^d \) for
Lemma 2.5. For any

\[ \text{Combining Lemmas 2.2 and 2.3 gives a result which will be useful in our proofs.} \]

\section*{Corollary 2.4. Let $\delta > 0$ and suppose that $f \in X_\delta$. Then the function $t \mapsto \|f\|_{X_t}$ is continuous on $[0, \delta]$.}

If $f \equiv 0$, this is trivial; otherwise it is an immediate consequence of (2.15) and the fact that any nonvanishing log-convex function is continuous.

\section*{Lemma 2.5. For any $\theta \in [0, 1]$, $T$ maps $X_\theta$ to $Y_{*,\theta}$, $T$ maps $X_\theta$ to $X_{*,\theta}$, and $T_*$ maps $X_{*,\theta}$ to $X_\theta$. Therefore exists $C < \infty$ such that for all $\theta \in [0, 1]$}

(2.16) \[ \|T(f)\|_{Y_{*,\theta}} \leq C\|f\|_{X_\theta} \]

(2.17) \[ \|T(f)\|_{X_{*,\theta}} \leq C\|f\|_{X_\theta}^d \]

(2.18) \[ \|T_*(f)\|_{X_\theta} \leq C\|f\|_{X_{*,\theta}}^d \]

Likewise $S$ maps $X_\theta$ to $X_\theta$ and

(2.19) \[ \|S(f)\|_{X_\theta} \leq C\|f\|_{X_\theta}^d \]

for all $f \in X_\theta$.\]
We will need to apply inequalities of Calderón-Zygmund/Littlewood-Paley type at certain points in the proof, with respect to weighted $L^p$ norms. There is a well-known condition on the weight which ensures that such operators are bounded. Denote by $A_p = A_p(\mathbb{R}^d)$ the usual Muckenhoupt classes of weights [19]. For $1 < p < \infty$, $A_p$ is the set of all locally integrable nonnegative functions $w$ for which the quantity

$$\left[ u \right]_{A_p} = \sup_B \left( |B|^{-1} \int_B u \right) \left( |B|^{-1} \int_B u^{-1/(p-1)} \right)^{p-1}$$

is finite. Operators of Calderón-Zygmund and Littlewood-Paley type are bounded on $L^p(w)$ for $u \in A_p$ [19].

**Lemma 2.6.** Let $P > 1$. There exists $\delta > 0$ such that

$$w^t \in A_p$$

for all $|t| \leq \delta$ and $p \in [P, \infty]$.  

**Lemma 2.7.** Let $u, v : \mathbb{R}^d \to [0, \infty)$ be measurable functions and $p \in (1, \infty)$. If $u, v \in A_p$ then $\max(u, v) \in A_p$.

**Proof.** Let $B \subset \mathbb{R}^d$ be any ball of finite radius. If $\int_B v \leq \int_B u$, then the following reasoning applies:

$$|B|^{-1} \int_B w = |B|^{-1} \int_B \max(u, v) \leq |B|^{-1} \int_B (u + v) \leq 2|B|^{-1} \int_B u.$$

Therefore

$$|B|^{-1} \int_B w \cdot \left( |B|^{-1} \int_B w^{-1/(p-1)} \right)^{p-1} \leq 2|B|^{-1} \int_B u \cdot \left( |B|^{-1} \int_B w^{-1/(p-1)} \right)^{p-1}$$

$$\leq 2 |B|^{-1} \int_B u \cdot \left( |B|^{-1} \int_B u^{-1/(p-1)} \right)^{p-1} \leq 2[u]_{A_p}$$

where $[u]_{A_p}$ is the $A_p$ constant of $u$.

If on the other hand $\int_B v \geq \int_B u$, then the same reasoning yields the bound $2[v]_{A_p}$.

Thus $[\max(u, v)]_{A_p} \leq 2 \max([u]_{A_p}, [v]_{A_p})$.

**Proof of Lemma 2.6.** $w(x', x_d) = \max(1, |x'|^d, |x_d + |x'|^2|^d)$. Therefore by repeated applications of Lemma 2.7 it suffices to prove that $|x'|^{d} \in A_p$ and $|x_d + |x'|^2|^d \in A_p$.

Let $s \geq 0$. It is well known that $u(x') = |x'|^s$ belongs to $A_p(\mathbb{R}^{d-1})$ if and only if $u^{-1/(p-1)} \in L_{\text{loc}}^1(\mathbb{R}^{d-1})$, thus if and only if $s < (p-1)(d-1)$. It follows easily that $u(x', x_d) = |x'|^s$ belongs to $A_p(\mathbb{R}^d)$ for the same range of $s$, that is, if and only if $s < (p-1)(d-1)$.

It is elementary that $|x_d + |x'|^2|^s \in A_p(\mathbb{R}^d)$ whenever $s/(p-1) < 1$ and $2s/(p-1) < d - 1$; details are left to the reader. Alternatively, a general result [19] p. 219 asserts that for any polynomial $P$ of degree $D$, $|P|^s \in A_p(\mathbb{R}^d)$ whenever $sD < p - 1$; in our case this implies that $|x_d + |x'|^2|^s \in A_p$ for all $s < (p-1)/2$.

**Remark 2.1.** The weight $w^{|p|}$ therefore belongs to $A_q$ whenever $dtp_t < q - 1$. Substituting $p_t = (1 - t)^{-1} p_0 = (1 - t)^{-1} (d + 1)/d$ gives the sufficient condition

$$q > 1 + dt(1 - t)^{-1} (d + 1)/d = 1 + t(1 - t)^{-1} (d + 1)$$
For $q = p_t$ this becomes
\[ p_t > 1 + dtp_t \Leftrightarrow 1 > p_t^{-1} + dt \Leftrightarrow 1 > (1 - t)d/(d + 1) + dt. \]
This is clearly not satisfied for $t = 1$, but is satisfied for $0 \leq t < 1/d^2$.

Denote by $|D|^r$ the differentiation operators $\hat{|D|^r}f(\xi) = |\xi|^r \hat{f}(\xi)$. We will use the notation $\langle z \rangle = (1 + |z|^2)^{1/2}$ for $z \in \mathbb{C}$. Let $\nabla^*$ denote the divergence of a vector field.

**Lemma 2.8.** There exists $C \in (0, \infty)$ such that for all sufficiently small $t \geq 0$,
\[ C^{-1}\|h\|_{X_t} \leq \|\nabla|D|^{-1}h\|_{X_t} \leq C\|h\|_{X_t} \]
for all $h \in X_t$.

*Proof.* The operators $\nabla \circ |D|^{-1}$ and $|D|^{-1} \circ \nabla^*$ are Calderón-Zygmund operators of classical type. Moreover, $|D|^{-1}\nabla^* \circ \nabla|D|^{-1}$ is the identity operator. The weight used to define $X_t$ belongs to $A_{p_t}$, provided that $t$ is sufficiently small. Therefore (2.20) follows from the theory of weighted Calderón-Zygmund inequalities. \qed

**Lemma 2.9.** For all sufficiently small $\varrho \geq 0$,
\[ \|\langle z \rangle \| \leq \|\nabla f\|_{X_{\varrho}} \leq \|f\|_{X_{\varrho}} \]
for all $f \in X_t$.

*Sketch of proof.* Consider the analytic family of operators $z \mapsto |D|^z$. For all sufficiently small $t \geq 0$, $|D|^{|\sigma|}$ is bounded on $X_t$ with a norm $\lesssim \langle \sigma \rangle^C$, uniformly for all $\sigma \in \mathbb{R}$. This inequality holds by Lemma 2.6 and the theory of $A_p$ weighted inequalities for Calderón-Zygmund operators. It follows that
\[ \|\langle z \rangle \| \leq \|\nabla f\|_{X_{\varrho}} \leq \|f\|_{X_{\varrho}}. \]

But $|D|f$ may be replaced by $\nabla f$, by Lemma 2.8. \qed

3. Multilinear Bounds

Define the multilinear operators
\[ \tilde{S}(f_i,j)^{d}_{i,j=1} = \prod_{i=1}^{d} T_i^* \left( \prod_{j=1}^{d} T(f_{i,j}) \right). \]

Thus $Sf = \tilde{S}(f, f, \ldots, f)$. We will sometimes write this more simply as $\tilde{S}(\vec{f})$ where $\vec{f} = (f_\alpha : \alpha \in A)$, with $A = \{1, 2, \ldots, d\}^2$. Repeated applications of Hölder’s inequality lead to the inequality
\[ |\tilde{S}(\vec{f})| \leq \left( \prod_{\alpha \in \{1, 2, \ldots, d\}^2} S(|f_\alpha|) \right)^{1/d^2}. \]
Indeed, for any nonnegative functions \( g_i \),

\[
T^* \prod_{i=1}^{d} g_i = \int_{\mathbb{R}^{d-1}} \prod_{i=1}^{d} g_i(x' - t, x_d + |t|^2) \, dt
\]

\[
\leq \prod_{i=1}^{d} \left( \int_{\mathbb{R}^{d-1}} g_i(x' - t, x_d + |t|^2) \, dt \right)^{1/d}
\]

\[
= \prod_{i=1}^{d} (T^*(g_i^d))^{1/d}.
\]

Applying this to \( g_i = |T(f_{i,j})| \leq T(|f_{i,j}|) \) gives (3.2).

**Lemma 3.1.** Let \( A \) be any finite index set. Let \( \theta_{\alpha}, t_{\alpha} \in [0, 1] \) for each \( \alpha \in A \). Suppose that \( \sum_{\alpha \in A} \theta_{\alpha} = 1 \) and

\[
1 - t = \sum_{\alpha \in A} \theta_{\alpha}(1 - t_{\alpha}).
\]

Then for any nonnegative functions \( f_{\alpha} \),

\[
\| \prod_{\alpha \in A} f_{\alpha}^{\theta_{\alpha}} \|_{X_t} \leq \prod_{\alpha \in A} \| f_{\alpha} \|_{X_{t_{\alpha}}}^{q_{\alpha}}.
\]

**Proof.** In this proof, products with respect to \( \alpha \) are always taken over all \( \alpha \in A \). Recall that \( p_t^{-1} = (1 - t)p_0^{-1} \). Introduce the exponents

\[
q_{\alpha}^{-1} = \frac{\theta_{\alpha}(1 - t_{\alpha})}{1 - t} = \frac{\theta_{\alpha}p_t}{p_{t_{\alpha}}}
\]

and

\[
r_{\alpha} = \frac{p_0 t_{\alpha} \theta_{\alpha}}{1 - t} = p_{t_{\alpha}} t_{\alpha} \theta_{\alpha}.
\]

Since

\[
\sum_{\alpha} r_{\alpha} = p_0 (1 - t)^{-1} \sum_{\alpha} t_{\alpha} \theta_{\alpha} = p_0 (1 - t)^{-1} \sum_{\alpha} (t_{\alpha} - 1) \theta_{\alpha} + p_0 (1 - t)^{-1} \sum_{\alpha} \theta_{\alpha}
\]

\[
= p_0 (1 - t)^{-1} \cdot (t - 1) + p_0 (1 - t)^{-1} \cdot 1 = p_0 t (1 - t)^{-1} = tp_t,
\]

one can write

\[
\| \prod_{\alpha} f_{\alpha}^{\theta_{\alpha}} \|_{X_t} = \int_{\mathbb{R}^d} \prod_{\alpha} f_{\alpha}^{p_t \theta_{\alpha} w^{p_t}}
\]

\[
= \prod_{\alpha} \left( f_{\alpha}^{p_t \theta_{\alpha} w^{r_{\alpha}}} \right).
\]

Since \( \sum_{\alpha} q_{\alpha}^{-1} = 1 \) by the hypothesis (3.3), Hölder’s inequality gives

\[
\int \prod_{\alpha} \left( f_{\alpha}^{p_t \theta_{\alpha} w^{r_{\alpha}}} \right) \leq \prod_{\alpha} \left( \int f_{\alpha}^{q_{\alpha} p_t \theta_{\alpha} w^{r_{\alpha} q_{\alpha}}} \right)^{1/q_{\alpha}}.
\]
The exponents in this last expression can be simplified: \( q_\alpha p_\beta \theta_\alpha = p_t \), while
\[
\tau_\alpha q_\alpha = p_0 (1 - t_\alpha)^{-1} t_\alpha = t_\alpha p_t.
\]
Thus the last expression is simply
\[
\prod_\alpha \left( \int f_\alpha^{p_t} L_\alpha t_\alpha p_\alpha^{p_\beta} \right)^{p_t \theta_\alpha / p_\beta} = \prod_\alpha \| f_\alpha \|_{X_t \alpha}^{\theta_\alpha p_t}.
\]
Raising everything to the power \( 1/p_t \) establishes the lemma. \( \square \)

Now let \( A = \{1, \ldots, d\}^2 \).

**Corollary 3.2.** Let \( t \in [0, 1/d^2] \). Let \( \beta \in A \). Let \( f_\alpha : \alpha \in A \) satisfy \( f_\beta \in X_{d^2t} \) and \( f_\alpha \in X_0 \) for all \( \alpha \neq \beta \). Then \( \vec{S}(\vec{f}) \in X_t \), and
\[
\| \vec{S}(\vec{f}) \|_{X_t} \leq C \| f_\beta \|_{X_{d^2t}} \prod_{\alpha \neq \beta} \| f_\alpha \|_{X_0}.
\]

**Proof.** \( \| \vec{S}(\vec{f}) \| \leq \prod_\alpha S(\| f_\alpha \|)^{1/d^2} \). Let \( \theta_\alpha = 1/d^2 \) for all \( \alpha \in A \). Let \( t_\alpha = 0 \) for all \( \alpha \neq \beta \), and \( t_\beta = d^2 t_\). These parameters satisfy the hypotheses (3.3) of Lemma 3.1.

Therefore
\[
\| \vec{S}(\vec{f}) \|_{X_t} \leq \prod_\alpha S(\| f_\alpha \|)^{1/d^2} \|_{X_t} \leq \prod_\alpha \| S(\| f_\alpha \| \|_{X_t \alpha}^{\theta_\alpha d_2} \leq C \prod_\alpha \| f_\alpha \|_{X_t \alpha}^{-\theta_\alpha d_2} = C \| f_\beta \|_{X_{d^2t}} \prod_{\alpha \neq \beta} \| f_\alpha \|_{X_0}.
\]

\( \square \)

## 4. Smoothing

Consider the operators \( T_\rho f(x) = \int_{|t| \leq \rho} f(x' - t, x_d - |t|^2) \, dt \).

**Lemma 4.1.** There exists \( \alpha_0 > 0 \) such that for any \( \alpha \in [0, \alpha_0] \), there exist \( C, A < \infty \) such that for all \( \rho \in [1, \infty) \),
\[
\| D^\alpha T_\rho f \|_{L^2(\mathbb{R}^d)} \leq C \rho^A \| f \|_{L^2(\mathbb{R}^d)}
\]
for all \( f \in L^2(\mathbb{R}^d) \).

For any \( p \in (1, \infty) \) there exist \( \eta > 0 \) and \( C, A < \infty \) such that for all \( \rho \in [1, \infty) \),
\[
\| D^\alpha T_\rho f \|_{L^p(\mathbb{R}^d)} \leq C \rho^A \| f \|_{L^p(\mathbb{R}^d)}
\]
for all \( f \in L^p(\mathbb{R}^d) \).

The proof of the first conclusion is a routine application of the method of stationary phase. See [19] for calculations of this type.

\( T_\rho \) is defined by convolution with a finite measure of total variation \( O(\rho^d) \), and consequently satisfies \( \| T_\rho f \|_{L^p} \leq C \rho^{d-1} \| f \|_{L^p} \) for all \( \rho \in [1, \infty] \). The second conclusion is by interpolating between these simple bounds and the first conclusion, using an analytic family of operators \( z \mapsto |D|^\eta \circ T_\rho \).

Let \( B_R = \{ x \in \mathbb{R}^d : |x| \leq R \} \).
Corollary 4.2. Let $t > 0$. There exists $\gamma = \gamma(t) > 0$ such that for any $f \in X_t$, $|D|^\gamma(Tf) \in L^q_{\text{loc}}$. More quantitatively, for any $R < \infty$ there exists $C < \infty$ such that for any $f \in X_t$, 
$$
\| |D|^\gamma(Tf)\|_{L^q_0(B_R)} \leq C \| f \|_{X_t}.
$$

Proof. There exists $\delta > 0$ such that whenever $|x| \geq 1$, $w(x) \geq c|x|^{d/2}$. Indeed, 
$$
\begin{align*}
    &w(x) = \max(1, |x'|^d, |x_d - |x'|^2|^d) \\
    &\quad \geq \max(\langle x' \rangle, \langle x_d - |x'|^2 \rangle^{1/2})^d \\
    &\quad \geq c \max(\langle x' \rangle, \langle x_d \rangle^{1/2})^d.
\end{align*}
$$

Let $1 \leq R, \rho < \infty$. Define $T^s f$ to be the restriction of $T f$ to $B(0, R)$. Let $t > 0$ and consider any function $g \in X_t$ supported in $B(0, \rho) \setminus B(0, \rho/2)$ satisfying $\| g \|_{X_t} \leq 1$. Then $\| g \|_{X_{t/2}} = O(\rho^{-\delta})$, where $\delta > 0$ depends only on $t$. Therefore $\| T g \|_{Y_{*, t/2}} \lesssim \rho^{-\delta}$. The space $Y_{*, t/2}$ embeds continuously into $L^q_t$ for some $q_1 > q_0$, yielding

\begin{equation}
\| T^s g \|_{L^q_t} \lesssim \rho^{-\delta}.
\end{equation}

On the other hand, because $g$ is supported in $B(0, \rho)$ and $T^s g$ is the restriction of $T g$ to $B(0, R)$, $T^s g$ equals the restriction of $T_s g$ to $B(0, R)$ where $s = C(R + \rho) \leq CR\rho$. Therefore by Lemma 4.1 there exists $\eta > 0$ such that

\begin{equation}
\| |D|^\eta T^s g\|_{L^q_0} \leq CR^A \| g \|_{L^q_0} \leq CR^A \| T g \|_{X_t} \leq CR^A \rho^\eta
\end{equation}

for a certain finite constant $A$, which depends only on the dimension $d$.

By interpolating between (4.1) and (4.2) using the natural analytic family of operators, we find that for any $\theta \in [0, 1]$,

\begin{equation}
\| |D|^\theta T^s g\|_{L^q_0} \leq CR^A \rho^A (1 - \theta)\delta
\end{equation}

where $Q(\theta)^{-1} = \frac{\theta}{\eta} + \frac{1}{\eta_0}(1 - \theta)$. Then $Q(0)^{-1} = q_1^{-1} < q_0^{-1}$. Therefore for all sufficiently small $\theta > 0$, $A\theta - (1 - \theta)\delta < 0$ and $Q(\theta)^{-1} < q_0^{-1}$. Fix one such parameter $\theta$. By Hölder’s inequality,

$$
\| |D|^\theta T^s g\|_{L^q_0} \leq CR^C \| D |^\theta T^s g\|_{L^q(\theta)}.
$$

Therefore in all,

$$
\| |D|^\theta T^s g\|_{L^q_0} \leq CR^C \rho^{-\varepsilon}
$$

for some $C < \infty$ and $\varepsilon, \theta, \eta > 0$.

We have proved that

\begin{equation}
\| |D|^\theta T g\|_{L^q_0(B(0, R))} \leq CR^C \rho^{-\varepsilon},
\end{equation}

provided that $g$ is supported on $B(0, \rho) \setminus B(0, \rho/2)$. Here $\varepsilon, \gamma > 0$. The same reasoning gives

\begin{equation}
\| |D|^\gamma T g\|_{L^q_0(B(0, R))} \leq CR^C \rho^C
\end{equation}

if $g$ is merely assumed to be supported on $B(0, \rho)$.

The proof of Corollary 4.2 is concluded by decomposing a general function $f$ as $\sum_{k=0}^\infty f_k$ where $f_0$ is supported on $B(0, R)$ and $f_k$ on $B(0, 2^k) \setminus B(0, 2^{k-1})$ for all $k \geq 1$. Apply (4.4) to the contribution of $f_k$ for all $k \geq 1$, and (4.5) for $k = 0$, and sum over $k$. \qed
5. Gaining some decay

Our goal here is to prove:

**Proposition 5.1.** Let $d \geq 2$ and $\lambda \in \mathbb{C}$. Let $f \in L^{p_0}(\mathbb{R}^d) = X_0(\mathbb{R}^d)$ be a solution of the generalized Euler-Lagrange equation $f = \lambda Sf$. Then there exists $t > 0$ such that $f \in X_t$.

To begin the proof, consider any decomposition $f = \varphi + g$ where $\varphi \in L^\infty$ has bounded support. Rewrite the equation $f = \lambda Sf$ as

$$
(5.1) 
\quad g = \lambda Sg + \mathcal{L}(\varphi, g)
$$

Then

$$
(5.2) 
\quad \mathcal{L}(\varphi, g) = \lambda S(\varphi + g) - \lambda Sg - \varphi.
$$

by the representation $\mathcal{L}(\varphi, g) = g - \lambda Sg$ and the basic $X_t$ bound for $S$. On the other hand, by expanding $S(\varphi + g)$ as a sum of $d^2$ terms $S(\cdot)$ and invoking (3.2) along with the bound $\|S(h)\|_{X_t} \leq C\|h\|_{X_t}^{d^2}$ gives an alternative majorization

$$
\|\mathcal{L}(\varphi, g)\|_{X_t} \leq C\varphi + C\varphi\|g\|_{X_t}^{d^2-1}.
$$

This bound can be improved; the operator $g \mapsto \mathcal{L}(\varphi, g)$ improves integrability in the following sense.

**Lemma 5.2.** For any bounded, compactly supported function $\varphi$ there exists $C_\varphi < \infty$ such that for all $g \in X_0$, the function $\mathcal{L}(\varphi, g)$ belongs to $X_{1/d^2}$, and

$$
(5.3) 
\quad \|\mathcal{L}(\varphi, g)\|_{X_{1/d^2}} \leq C_\varphi + C_\varphi\|g\|_{X_0}^{d^2-1}.
$$

**Proof.** By assumption, $\varphi \in X_1 \subset X_{1/d^2}$, so it suffices to show that $S(\varphi + g) - S(g)$ satisfies the required bound. Let $A = \{1, 2, \cdots, d\}^2$. $S(\varphi + g) - S(g)$ can be expanded as a sum of $d^2 - 1$ terms, each of which is of the general form $\bar{S}(\tilde{f})$ where $\tilde{f} = (f_\alpha : \alpha \in A)$, where each $f_\alpha$ equals either $\varphi$ or $g$, and where for each such term, there exists at least one index $\beta$ in $A$ for which $f_\beta = \varphi$. The required bound therefore follows directly from Corollary 3.2, again since $\varphi \in X_1$. \(\square\)

We continue with the proof of Proposition 4.1. Let $\varepsilon > 0$. Decompose $f = \varphi_\varepsilon + g_\varepsilon$ where $\|g_\varepsilon\|_{X_0} < \varepsilon$, and $\varphi_\varepsilon \in L^\infty$ has bounded support. Define

$$
A_\varepsilon(h) = \lambda Sh + \mathcal{L}(\varphi_\varepsilon, g_\varepsilon).
$$

This operator depends of course on $\varphi_\varepsilon, g_\varepsilon$, and is defined in such a way that $A_\varepsilon(g_\varepsilon) = g_\varepsilon$, that is, $g_\varepsilon$ is one solution of the fixed point equation $A_\varepsilon(h) = h$ in the space $X_0$.

**Lemma 5.3.** Let $\lambda \in \mathbb{C}$, and let $f \in L^{p_0(d)}(\mathbb{R}^d)$ be any solution of $f = \lambda Sf$. For each $\varepsilon > 0$, let $f = \varphi_\varepsilon + g_\varepsilon$ be any decomposition with $\varphi_\varepsilon$ bounded and having bounded support, and with $\|g_\varepsilon\|_{L^{p_0}} < \varepsilon$. Then there exists $\varepsilon_0 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0]$, there exists $t_\varepsilon > 0$ such that for all $t \in [0, t_\varepsilon]$, the fixed point equation

$$
A_\varepsilon(h) = h
$$

has a unique solution $h \in X_t$ satisfying $\|h\|_{X_t} \leq \varepsilon_0^{1/2}$. 
It bears emphasis that there are no a priori bounds for $\varepsilon_0$ or $t_\varepsilon$; these depend on $f$ in some uncontrolled manner.

**Proof.** We know that
\[
\| L(\varphi_\varepsilon, g_\varepsilon) \|_{X_0} \leq \varepsilon + C\varepsilon^{d^2},
\]
and that
\[
L(\varphi_\varepsilon, g_\varepsilon) \in X_{1/d^2}.
\]
By convexity of the $X_t$ norms, for each sufficiently small $\varepsilon > 0$ there exists $t_\varepsilon > 0$ such that
\[
\| L(\varphi_\varepsilon, g_\varepsilon) \|_{X_{t_\varepsilon}} \leq \varepsilon^{3/4}.
\]
Henceforth we consider only such small $\varepsilon$.

Let $B_\varepsilon$ be the ball of radius $\varepsilon^{1/2}$ in $X_{t_\varepsilon}$, centered at 0. If $h \in B_\varepsilon$ then
\[
\| A_\varepsilon(h) \|_{X_{t_\varepsilon}} \leq |\lambda| \cdot \| Sh \|_{X_{t_\varepsilon}} + \| L(\varphi_\varepsilon, g_\varepsilon) \|_{X_{t_\varepsilon}}
\leq C\| h \|_{X_{t_\varepsilon}}^{d^2} + \varepsilon^{3/4}
\leq C\varepsilon^{d^2/2} + \varepsilon^{3/4}
\leq \varepsilon^{1/2},
\]
so $A_\varepsilon(B_\varepsilon) \subset B_\varepsilon$. For any $h, \tilde{h} \in B_\varepsilon$,
\[
\| A_\varepsilon(h) - A_\varepsilon(\tilde{h}) \|_{X_{t_\varepsilon}} = |\lambda| \cdot \| Sh - Sh \|_{X_{t_\varepsilon}}
\leq C\| h - \tilde{h} \|_{X_{t_\varepsilon}} \cdot (\| h \|_{X_{t_\varepsilon}} + \| \tilde{h} \|_{X_{t_\varepsilon}})^{d^2-1}
\leq C\varepsilon^{1/2} \| h - \tilde{h} \|_{X_{t_\varepsilon}}.
\]
Therefore $A_\varepsilon : B_\varepsilon \rightarrow B_\varepsilon$ is a strict contraction, for each sufficiently small $\varepsilon$. Therefore there exists a unique $h_\varepsilon \in X_{t_\varepsilon}$ satisfying both $\| h_\varepsilon \|_{X_{t_\varepsilon}} \leq \varepsilon^{1/2}$ and $A_\varepsilon(h_\varepsilon) = h_\varepsilon$.

Exactly the same reasoning applies in $X_t$ for any $0 \leq t \leq t_\varepsilon$. □

**Proof of Proposition 5.7.** Suppose that $0 \leq s \leq t \leq t_\varepsilon$, and that both $h \in X_s$ and $\tilde{h} \in X_t$ are solutions of $A_\varepsilon(h) = h$, satisfying $\| h \|_{X_s} \leq \varepsilon^{1/2}$ and $\| \tilde{h} \|_{X_t} \leq \varepsilon^{1/2}$. Then
\[
\| \tilde{h} \|_{X_s} = \| A_\varepsilon \tilde{h} \|_{X_s} \leq C\| A_\varepsilon \tilde{h} \|_{X_t} \leq C\varepsilon^{3/4} < \varepsilon^{1/2},
\]
provided that $\varepsilon$ remains sufficiently small. Therefore $\tilde{h} = h$ by the uniqueness of solutions.

In particular, since $g_\varepsilon$ is a solution in $X_0$, this uniqueness of solutions implies that
\[
g_\varepsilon = h_\varepsilon \in X_{t_\varepsilon}
\]
for all sufficiently small $\varepsilon > 0$, as was to be proved. □
6. Smoothness

We have shown that any solution of the Euler-Lagrange equation enjoys some extra decay, beyond that encoded by the finiteness of its $L^p$ norm. We will next show how such extra decay can be used in conjunction with the Euler-Lagrange equation to demonstrate some smoothness. Our initial goal is to prove the following a priori inequality.

**Lemma 6.1.** Let $\rho > 0$ be sufficiently small. Then for any $0 \leq \varrho < \rho$ there exists $C < \infty$ such that for any solution $f$ of $f = \lambda Sf$, if $f \in X_\rho$ and $\nabla f \in X_\varrho$ then

$$
\|\nabla f\|_{X_\varrho} \leq C \|f\|_{X_\rho}^{d^2}.
$$

Here $C$ depends only on $\rho, \varrho, \lambda, d$.

Here $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_d}\right)$. It suffices to prove this under the assumption that $\|f\|_{X_\rho} = 1$, which will be assumed for the remainder of §6. Indeed, for general $f$, consider the function $F = f/\|f\|_{X_\rho}$. It satisfies the modified equation $F = \tilde{\lambda} SF$ where $\tilde{\lambda} = \lambda \|f\|_{X_\rho}^{\sigma}$ for a certain exponent $\sigma$. Thus we only have to replace $\lambda$ by $\tilde{\lambda}$ in order to assume $\|f\|_{X_\rho} = 1$.

**Lemma 6.2.** Let $\rho > 0$ be sufficiently small and $\lambda \in \mathbb{C}$. Let $0 < \varrho < \rho$. There exists $R < \infty$ such that for any function $f$ satisfying $f = \lambda Sf$ and $\|f\|_{X_\rho} = 1$, with $\nabla f \in X_\varrho$,

$$
\|\nabla f\|_{X_\varrho} \leq C \|T\nabla f\|_{L^{\infty}(B_R)}
$$

where $C, R, a$ depend only on $d, \rho, \varrho, \lambda, \|f\|_{X_\rho}$.

**Proof of Lemma 6.2.** Write

$$
\nabla f = \lambda \nabla (Sf) = d^2 \lambda \tilde{S}(f, f, \ldots, f, \nabla f)
$$

$$
= d^2 \lambda (T^*([Tf]^d))^{d-1} \cdot T^*([Tf]^d-1) \cdot \nabla T f.
$$

(6.2)

Here $\tilde{S}(f, f, \ldots, f, \nabla f)$ stands for the vector with $d$ components, whose $j$-th component equals $\tilde{S}(f, f, \ldots, f, \partial f/\partial x_j)$.

Therefore

$$
\|\nabla f\|_{X_\varrho} \leq C \|f\|_{X_\rho}^{d^2-1} \|\nabla T f\|_{Y_{\varrho', \rho}^d} = C \|\nabla T f\|_{Y_{\varrho', \rho}^d}
$$

(6.3)

for a certain $\varrho'< \varrho; \varrho'$ does not depend on $f$. Since

$$
\|\nabla T f\|_{Y_{\varrho', \rho}^d} \leq \|\nabla T f\|_{Y_{\varrho', \rho}^d}^{1-\theta} \|\nabla T f\|_{L^{\infty}}^{\theta}
$$

$$
= \|\nabla T f\|_{Y_{\varrho', \rho}^d}^{1-\theta} \|\nabla T f\|_{L^{\infty}}^{\theta}
$$

$$
\leq \|\nabla f\|_{X_\rho}^{1-\theta} \|\nabla f\|_{L^{\infty}}^{\theta}
$$

for some $\theta \in (0, 1)$, we deduce that

$$
\|\nabla f\|_{X_\varrho} \leq C \|\nabla T f\|_{L^{\infty}}
$$

(6.4)
where \( C \) depends only on \( \rho, \varrho, \lambda \). Now for any \( R < \infty \),

\[
\int_{|x| \geq R} |\nabla T f(x)|^{q_0} \, dx \leq R^{-q_0 \tau} \| \nabla T f \|_{Y_\varrho, q_0}^{q_0} = R^{-q_0 \tau} \| T \nabla f \|_{Y_\varrho, q_0}^{q_0} \leq CR^{-q_0 \tau} \| \nabla f \|_{Y_\varrho, q_0}^{q_0}
\]

for a certain exponent \( \tau > 0 \). Therefore

\[
\| \nabla f \|_{X_\varrho} \leq C \| \nabla T f \|_{L^{q_0}(B_R)} + CR^{-\tau} \| \nabla f \|_{X_\varrho}.
\]

Define \( R \) by the equation \( CR^{-\tau} = \frac{1}{2} \) to obtain

(6.5)

\[
\| \nabla f \|_{X_\varrho} \leq 2C \| \nabla T f \|_{L^{q_0}(B_R)}\]

where \( C, R \) depend only on \( \rho, \varrho, \lambda, d \). \( R \) will henceforth remain fixed. This same reasoning can be carried out for all dimensions \( d \) with very minor changes. \( \square \)

**Proof of Lemma 6.1.** We will use the representation (6.2) in order to obtain a bound for \( \| T \nabla f \|_{L^{q_0}(B_R)} \) in terms of \( \| f \|_{X_\rho} \), where \( R \) is as defined above.

Let \( \gamma \in (0, 1) \) be a small constant, to be chosen below. Writing \( T \nabla f = |D|^{\gamma} T(|D|^{-\gamma} f) \),

Corollary 4.2 gives

\[
\| T \nabla f \|_{L^{q_0}(B_R)} = \| |D|^{\gamma} T(|D|^{-\gamma} f) \|_{L^{q_0}(B_R)} \lesssim \| \nabla |D|^{-\gamma} f \|_{X_\varrho}
\]

provided that \( \gamma \) is a sufficiently small function of \( \varrho, d \) alone.

Therefore by Lemma 2.8, Lemma 2.9, and a second application of Lemma 2.8,

\[
\| \nabla |D|^{-\gamma} f \|_{X_\varrho} \leq C \| D \|^{1-\gamma} f \|_{X_\varrho} \leq C \| \nabla f \|_{X_\varrho, \rho}^{1-\gamma} \| f \|_{X_\varrho}^{\gamma}
\]

for some \( \gamma \in (0, 1) \).

Thus

\[
\| T \nabla f \|_{L^{q_0}(B_R)} \lesssim \| \nabla f \|_{X_\varrho, \rho}^{1-\gamma} \| f \|_{X_\varrho}^{\gamma} = \| \nabla f \|_{X_\varrho, \rho}^{1-\gamma}
\]

and therefore by Lemma 6.2,

\[
\| \nabla f \|_{X_\varrho} \lesssim \| \nabla f \|_{X_\varrho, \rho}^{1-\gamma}.
\]

Recall that \( \| \nabla f \|_{X_\varrho, \rho}^{1-\gamma} \) is assumed to be finite. It follows from this last inequality that \( \| \nabla f \|_{X_\varrho} \lesssim 1 \). This completes the proof of Lemma 6.1. \( \square \)

7. Mollified Derivatives and Conclusion of Proof

Lemma 6.1 presupposes that \( \nabla f \in X_\rho \), which we seek to prove. In order to remove the extraneous assumption, we approximate \( \nabla \) by a one-parameter family of operators which are individually bounded on the spaces \( X_t \).

For any \( s \geq 0 \) and \( \Lambda \geq 1 \) define

(7.1) \[\widehat{D}_s^\Lambda f(\xi) = \min \left( 1 + |\xi|^2, (1 + \Lambda^2)^{s/2} \right) \cdot \widehat{f}(\xi).\]

These operators are bounded on all \( L^p \) spaces, and likewise on all spaces \( X_t \) for \( t \in [0, 1] \). For \( s = 1 \) we write simply \( D_\Lambda \).
In order to prove that $\nabla f \in X_\rho$, it suffices to show that $\|D_\Lambda f\|_{X_\rho} \leq A$ for some finite constant $A$ which is independent of $\Lambda$. The proof of Lemma 6.1 relied on Leibniz’ rule for derivatives of products. There is no corresponding formula for $D_\Lambda(fg)$, but the following lemma provides an adequate substitute.

**Lemma 7.1.** Let $u \geq 0$ be a locally integrable function. Let $s \in (0, \infty)$. Suppose that $r^{-1} = p_j^{-1} + q_j^{-1}$ for $j = 1, 2$ and that all exponents $r, p_j, q_j$ belong to the open interval $(1, \infty)$. Suppose that the weight $u$ belongs to $A_r$ and that $u = u_1v_1 = u_2v_2$ where $u_j^{p_j/r} \in A_{p_j}$ and $v_j^{q_j/r} \in A_{q_j}$. Then there exists $C < \infty$ such that $D_\Lambda^s(fg) \in L^r$, and the following inequality holds, whenever the right-hand side is finite:

$$
\|D_\Lambda^s(fg)\|_{L^r(u)} \leq C\|D_\Lambda^s f\|_{L^{p_1}(u_1^{p_1/r})} \|g\|_{L^{q_1}(v_1^{q_1/r})} + C\|f\|_{L^{p_2}(u_2^{p_2/r})} \|D_\Lambda^s g\|_{L^{q_2}(v_2^{q_2/r})}.
$$

A proof will be given in §9.

**Corollary 7.2.** Let $s \in (0, \infty)$. Let $\rho > 0$ be sufficiently small, and let $0 < q < \rho$. There exist $q' \in (0, q)$ and $C < \infty$ such that for all $\Lambda \geq 1$ and all vector-valued functions $\tilde{f} \in X_\rho$,

$$
\|D_\Lambda^s \tilde{S}(\tilde{f})\|_{X_\rho} \leq C \sum_j \prod_{i \neq j} \|f_i\|_{X_\rho} \cdot \|D_\Lambda^s T f_j\|_{Y_{s, q'}}.
$$

The constant $C$ may be taken to be independent of $\Lambda$ while $s, \rho$ remain fixed.

Together, the proof of Lemma 6.1 and Corollary 7.2 establish:

**Lemma 7.3.** Let $s \in (0, \infty)$ and $\Lambda \in \mathbb{C}$. Let $\rho > 0$ be sufficiently small, and let $0 < q < \rho$. Let $f \in X_{\rho}$ be any solution of the Euler-Lagrange equation (1.2). There exists $C < \infty$ such that for all $\Lambda < \infty$,

$$
\|D_\Lambda^s f\|_{X_\rho} \leq C.
$$

Because this bound is uniform in $\Lambda$, combining this lemma with Proposition 5.1 yields:

**Corollary 7.4.** Let $\Lambda \in \mathbb{C}$. Let $f \in X_{\rho}$ be any solution of the Euler-Lagrange equation (1.2). Then there exists $t > 0$ such that for all $s \geq 0$, $|D|^s f \in X_t$.

It is now an easy consequence of Sobolev embedding that any solution of (1.2) is $C^\infty$, completing the proof of Theorem 1.1.

8. **Proof of Lemma 2.1**

Recall the definitions

$$
T^* f(x) = \int_{\mathbb{R}^{d-1}} f(x' + t, x_d + |t|^2) \, dt
$$

$$
v(x) = \min \{ 1, |x'|^{d-1}, |x_d - |x'|^2|^{d-1} \}
$$

$$
v_*(x) = \min \{ 1, |x'|^{d-1}, |x_d + |x'|^2|^{d-1} \}
$$

where $x \in \mathbb{R}^d$ as written as $x = (x', x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}^1$. Lemma 2.1 states that $T^* v \lesssim v_*^{1/d}$, with a corresponding inequality for $T$. 
Each of the weights \( v, v_* \) is equal to a minimum of three functions having three different degrees of homogeneity 0, 1, 2 with respect to the parabolic dilation group \( x \mapsto (rx', r^2 x_d) \), so there is no dilation invariance to simplify the analysis. Viewing \( T^*v(x) \) as an integral with respect to a second variable \( y \in \mathbb{R}^d \), and comparing the result to \( v_*(x)^{1/d} \), the estimation of \( T^*v(x) \) splits naturally into \( 3 \times 3 = 9 \) cases. This factor of 9 accounts largely for the length of the proof which we now present; in actuality some cases are combinable, but various subcases also arise.

**Proof of Lemma 2.1.** The two conclusions of Lemma 2.1 can be shown to be equivalent by the change of variables \( (x', x_d) \mapsto (x', -x_d) \), along with the substitution \( t \mapsto -t \) in the integrals defining \( T, T^* \). So we will prove only the inequality \( T^*(v) \lesssim v_*^{1/d} \).

Write

\[
T^*v(x) = \int_{\mathbb{R}^{d-1}} v(x' + t, x_d + |t|^2) dt = \int_{\mathbb{R}^{d-1}} v(s, x_d + |s-x'|^2) ds
\]

\[
= \int_{\mathbb{R}^{d-1}} \min \left( 1, |s|^{-d}, |x_d + |s-x'|^2 - |s|^2|^{-d} \right) ds.
\]

Observe that

\[
T^*v(x) \lesssim \int_{\mathbb{R}^{d-1}} \langle s \rangle^{-d} ds \lesssim 1 \text{ uniformly for all } x \in \mathbb{R}^{d-1}.
\]

This satisfies the required bound \( C v_*^{1/d}(x) \) provided that \( v_*(x) \) remains uniformly bounded below. Therefore we may assume throughout the rest of the analysis of the contribution of \( T^*v(x) \) that

\[
\max(|x'|, |x_d + x'|^2) \gg 1.
\]

In the same way, because the integrand is \( \leq |s|^{-d} \) and because \( \int_{|s|<\lambda} |s|^{-d} dx \lesssim \lambda^{-1} \), the contribution made to the integral by the set of all \( s \) satisfying \( |s| \geq \frac{1}{4} \max(|x'|, |x_d + |x'|^2|) \) is

\[
\lesssim \max(|x'|, |x_d + |x'|^2|)^{-1} = v_*(x)^{1/d}.
\]

It remains to discuss the contribution of those \( s \) which satisfy

\[
|s| < \frac{1}{4} \max(|x'|, |x_d + |x'|^2|).
\]

For each \( x \in \mathbb{R}^d \), partition the set of all such \( s \in \mathbb{R}^{d-1} \) into two regions

\[
\mathcal{R}_1(x) = \{ s \in \mathbb{R}^{d-1} : |s| \geq |x_d + |x'|^2 - 2x' \cdot s| \text{ and } |s| < \frac{1}{4} \max(|x'|, |x_d + |x'|^2|) \};
\]

\[
\mathcal{R}_2(x) = \{ s \in \mathbb{R}^{d-1} : |s| < |x_d + |x'|^2 - 2x' \cdot s| \text{ and } |s| < \frac{1}{4} \max(|x'|, |x_d + |x'|^2|) \}.
\]

Thus we have shown that

\[
T^*v(x) \leq C v_*(x)^{1/d} + J_1(x) + J_2(x)
\]

where

\[
J_i(x) = \int_{\mathcal{R}_i(x)} \min \left( 1, |s|^{-d}, |x_d + |s-x'|^2 - |s|^2|^{-d} \right) ds.
\]
More succinctly,
\[ J_1(x) \asymp \int_{\mathcal{R}_1(x)} \langle s \rangle^{-d} \, ds \]
\[ J_2(x) \asymp \int_{\mathcal{R}_2(x)} \langle x_d + |x'|^2 - 2x' \cdot s \rangle^{-d} \, ds. \]

We will often write \( J_i, \mathcal{R}_i \) as shorthand for \( J_i(x), \mathcal{R}_i(x) \).

- **Estimate for \( J_1(x) \) in the case** \( |x'| \geq |x_d + |x'|^2| \). In this case,
  \[ |x_d + |x'|^2 - 2x' \cdot s| \leq |s| \leq \frac{1}{4} \max(|x'|, |x_d + |x'|^2|) < |x'| \]
  by definition of \( \mathcal{R}_1(x) \), so one of the following two subcases occurs:
  \[(8.6) \quad |x_d + |x'|^2 - 2x' \cdot s| \leq 1 \text{ or } 1 < |x_d + |x'|^2 - 2x' \cdot s| 
  \leq |x'|. \]

  Consider first the contribution made to \( J_1(x) \) by those \( s \in \mathcal{R}_1 \) which satisfy the first case in \((8.6)\). There exists at least one index \( i \in \{1, 2, \cdots, d - 1\} \) such that \(|x_i| \geq |x'|/\sqrt{d - 1}\). Our problem is invariant with respect to rotations of \( \mathbb{R}^{d-1} \), which leave the coordinate \( x_d \) unchanged. Therefore without loss of generality, we may assume throughout the remainder of the proof of the Lemma that
  \[(8.7) \quad |x_1| \geq |x'|/\sqrt{d - 1}. \]

  We are working in the situation where \( 1 \ll \max(|x'|, |x_d + |x'|^2|) = |x'| \lesssim |x_1| \) by \((8.1)\) and the definition of \( \mathcal{R}_{1,1} \), so \( |x_1| \gg 1 \).

  Introduce the notations
  \[ \tilde{s} = (s_2, \cdots, s_{d-1}) \text{ and } \phi(x, \tilde{s}) = x_d + |x'|^2 - 2 \sum_{i=2}^{d-1} x_is_i. \]

  Thus \( x_d + |x'|^2 - 2x' \cdot s = \phi(x, \tilde{s}) - 2x_1s_1 \), so \( |2x_1s_1 - \phi(x, \tilde{s})| \leq 1 \) by definition of \( \mathcal{R}_1(x) \) and the first case of \((8.6)\). The following fact will be used repeatedly throughout the analysis: If \((x, \tilde{s})\) is fixed, then an inequality \( |2x_1s_1 - \phi(x, \tilde{s})| \leq \delta \) forces \( s_1 \) to lie in an interval of length \( \delta|x_1|^{-1} \).

  Now the contribution made by those \( s \) belonging to the first subcase of \((8.6)\) to \( \int_{\mathcal{R}_1(x)} \langle s \rangle^{-d} \, ds \) is
  \[ \lesssim \int_{\mathcal{R}_1(x)} \langle s \rangle^{-d} \, ds \lesssim \int_{\mathbb{R}^{d-2}} \int_{|s_1 - \phi(x, \tilde{s})| \leq |x_1|^{-1}} \langle \tilde{s} \rangle^{-d} \, ds_1 \, d\tilde{s} \lesssim |x_1|^{-1} \int_{\mathbb{R}^{d-2}} \langle \tilde{s} \rangle^{-d} \, d\tilde{s} \lesssim |x_1|^{-1}. \]

  This is the required bound, for \( |x_1|^{-1} \lesssim v_*(x)^{1/d} \) because we are working in the case where \(|x_1| \gtrsim |x'| \gtrsim |x_d + |x'|^2| \gg 1 \).

  Next we consider the contribution of those \( s \in \mathcal{R}_1(x) \) which satisfy the second case in \((8.6)\), still under the assumption that \(|x_d + |x'|^2| \leq |x'|\). For \( j, k \geq 1 \) define
  \[ E_{1}^{j,k}(x) = \{ s \in \mathcal{R}_1(x) : 2^{-k} < \frac{|x_d + |x'|^2 - 2x' \cdot s|}{|x'|} \leq 2^{-k+1} \text{ and } 2^{-j} < \frac{|s|}{|x'|} \leq 2^{-j+1} \}. \]

  For any \( s \in E_{1}^{j,k}(x) \),
  \[ 2^{-k} |x'| \leq |x_d + |x'|^2 - 2x' \cdot s| \leq |s| \leq 2^{-j+1} |x'|. \]
Thus \( j \leq k + 1 \).

If \( \tilde{s} \) remains fixed and \( s \in E_{1}^{j,k} \), then \( s_{1} \) lies in an interval of length \( \asymp |x_{1}|^{-1} 2^{-k} |x'| \asymp 2^{-k} \). Since \( |\tilde{s}| \leq |s| \leq 2^{-j+1} |x'| \),

\[
|E_{1}^{j,k}(x)| \lesssim (2^{-j} |x'|)^{d-2} \cdot 2^{-k}.
\]

Therefore the contribution made to \( J_{1}(x) \) by all \( s \) belonging to this subcase is

\[
\leq \sum_{k=1}^{\lceil \log_{2} |x'|+1 \rceil} \sum_{j=1}^{k+1} \int_{E_{1}^{j,k}(x)} (s)^{-d} ds \leq \sum_{k=1}^{\lceil \log_{2} |x'|+1 \rceil} \sum_{j=1}^{k+1} (2^{-j} |x'|)^{-d} |E_{1}^{j,k}(x)| \lesssim \sum_{k=1}^{\lceil \log_{2} |x'|+1 \rceil} \sum_{j=1}^{k+1} 2^{2j} 2^{-k} \lesssim |x'|^{-1}.
\]

This completes the analysis of \( J_{1}(x) \) in the case where \( |x'| \geq |x_{d} + |x'|^2| \).

**Estimate for \( J_{1}(x) \) in the case \( |x'| \leq |x_{d} + |x'|^2| \).** The definition \((8.2)\) of \( \mathcal{R}_{1}(x) \) becomes

\[
(8.8) \quad |x_{d} + |x'|^2 - 2x' \cdot s| \leq |s| \leq |x_{d} + |x'|^2| \quad \text{for all} \quad s \in \mathcal{R}_{1}(x).
\]

If \( |x'| \leq 1 \), then since \( \max\{|x'|, |x_{d} + |x'|^2|\} \gg 1 \), necessarily \( |x_{d}| \gg 1 \). Since \( |x_{d} + |x'|^2 - 2x' \cdot s| \leq |s| \), it follows that \( |x_{d}| \leq C|s| \). On the other hand, by \((8.8)\) again, \( |s| \leq |x_{d} + |x'|^2| \leq 2|x_{d}| \). So \( |s| \asymp |x_{d}| \). Therefore

\[
J_{1}(x) \lesssim \int_{|s|\leq 2|x_{d}|} |x_{d}|-d ds \lesssim |x_{d}|^{-1} \sim |x_{d} + |x'|^2|^{-1} \sim v_{s}(x)^{1/d}.
\]

Suppose now that \( |x'| \geq 1 \). Recall our standing assumption that \( |x_{1}| \geq |x'|/\sqrt{d-1} \). Since \( |x_{d} + |x'|^2 - 2x' \cdot s| \leq |s| \),

\[
|x_{d} + |x'|^2| \leq |s| + 2|x'| \cdot |s| \leq 3|x'| \cdot |s|.
\]

Suppose that there exists some \( s \in \mathcal{R}_{1}(x) \) satisfying \( |s| \leq 1 \), and consider the contribution to \( J_{1}(x) \) made by all such \( s \). For fixed \( \tilde{s} \), according to \((8.8)\), \( s_{1} \) lies in an interval of length \( \lesssim |x'|^{-1} |s| \leq |x'|^{-1} \). Since \( |\tilde{s}| \leq |s| \leq 1 \), the intersection of \( \mathcal{R}_{1}(x) \) with \( \{ s : |s_{1}| \leq 1 \} \) has measure \( \lesssim |x'|^{-1} \). Therefore the contribution made to \( J_{1}(x) \) by all \( s \in \mathcal{R}_{1}(x) \) satisfying \( |s| \leq 1 \) is \( \lesssim |x'|^{-1} \). This is the required bound, for since \( |x_{d} + |x'|^2| \lesssim 3|x'| \cdot |s| \) and \( |s| \leq 1 \),

\[
|x'|^{-1} \lesssim |x_{d} + |x'|^2|^{-1}.
\]

Continuing to assume that \( |x'| \geq 1 \), consider next the contribution of all \( s \in \mathcal{R}_{1}(x) \) satisfying \( |s| > 1 \). Now \( s \in \mathcal{R}_{1} \) implies that \( |x_{d} + |x'|^2| \leq |s| + 2|s||x'| \leq 3|s| \cdot |x'| \), that is,

\[
(8.9) \quad |s| \geq 3^{-1} |x'|^{-1} |x_{d} + |x'|^2|.
\]

We will consider two subcases, (i) \( |x_{d} + |x'|^2 - 2x' \cdot s| > 3^{-1} |x'|^{-1} |x_{d} + |x'|^2| \) and (ii) \( |x_{d} + |x'|^2 - 2x' \cdot s| \leq 3^{-1} |x'|^{-1} |x_{d} + |x'|^2| \).
First, to treat the contribution of those \( s \in \mathcal{R}_1(x) \) which satisfy \(|x_d + |x'|^2 - 2x' \cdot s| > 3^{-1}|x'|^{-1}|x_d + |x'|^2|\), define \( E_2^{j,k}(x) \) to be the set of all \( s \in \mathcal{R}_1(x) \) which satisfy both of

\[
2^k 3^{-1} |x'|^{-1} |x_d + |x'|^2| < |x_d + |x'|^2 - 2x' \cdot s| \leq 2^{k+1} 3^{-1} |x'|^{-1} |x_d + |x'|^2|
\]
\[
2^j 3^{-1} |x'|^{-1} |x_d + |x'|^2| \leq |s| < 2^{j+1} 3^{-1} |x'|^{-1} |x_d + |x'|^2|.
\]

If \( s \in E_2^{j,k}(x) \) then \( k \leq j + 1 \).

Now

\[
|E_2^{j,k}(x)| \lesssim (2^{j+1} |x'|^{-1} |x_d + |x'|^2|)^{d-2} 2^{k+1} |x_d + |x'|^2| \cdot |x'|^{-2}.
\]

Indeed, \( |\tilde{s}| \leq |s| \leq 2^{j+1} 3^{-1} |x'|^{-1} |x_d + |x'|^2| \), while for fixed \( \tilde{s} \), \( s_1 \) lies in an interval of length \( \sim 2^{k+1} |x_d + |x'|^2| \cdot |x'|^{-2} \).

Together with (8.8), this implies that the total contribution made to \( J_1(x) \) by all \( s \in \mathcal{R}_1(x) \) which satisfy \(|x_d + |x'|^2 - 2x' \cdot s| > 3^{-1} |x'|^{-1} |x_d + |x'|^2| \) is

\[
\leq \sum_{j=1}^{\lceil \log_2 |x| \rceil + 1} \frac{(j+1)}{j} \int_{E_2^{j,k}(x)} |s|^{-d} ds
\]
\[
\lesssim \sum_{j=1}^\infty \sum_{k=1}^{j+1} (2^j |x'|^{-1} |x_d + |x'|^2|)^{-d} |E_2^{j,k}(x)|
\]
\[
\lesssim \sum_{j=1}^\infty \sum_{k=1}^{j+1} (2^j |x_d + |x'|^2| \cdot |x'|^{-1})^{-d} (2^{j+1} |x'|^{-1} |x_d + |x'|^2|)^{d-2} 2^{k+1} |x'|^{-2} |x_d + |x'|^2|
\]
\[
\lesssim (|x_d + |x'|^2| \cdot |x'|^{-1})^{-d} \cdot (|x'|^{-1} |x_d + |x'|^2|)^{d-2} \cdot |x'|^{-2} |x_d + |x'|^2| \cdot \sum_{j=1}^\infty 2^{-j}
\]
\[
= |x_d + |x'|^2|^{-1}.
\]

Secondly, to treat the contribution of those \( s \in \mathcal{R}_1(x) \) which satisfy \(|x_d + |x'|^2 - 2x' \cdot s| \leq 3^{-1} |x'|^{-1} |x_d + |x'|^2| \), define \( E_3^j(x) \) to be the set of all \( s \in \mathcal{R}_1(x) \) which satisfy both of

\[
|x_d + |x'|^2 - 2x' \cdot s| \leq 3^{-1} |x'|^{-1} |x_d + |x'|^2|
\]
\[
2^j |x'|^{-1} |x_d + |x'|^2| < |s| \leq 2^{j+1} |x'|^{-1} |x_d + |x'|^2|.
\]

If \( s \in E_3^j(x) \) and \( \tilde{s} \) is fixed, then \( s_1 \) lies in an interval of length \( \lesssim |x'|^{-2} |x_d + |x'|^2| \), while \( |\tilde{s}| \leq |s| \leq 2^{j+1} |x'|^{-1} |x_d + |x'|^2| \). Thus

\[
|E_3^j(x)| \lesssim |x'|^{-2} |x_d + |x'|^2| \cdot (2^j |x'|^{-1} |x_d + |x'|^2|)^{d-2} = 2^{(d-2)j} |x'|^{-d} |x_d + |x'|^2|^{d-1}.
\]
Consequently the contribution made by all \( s \in \mathcal{R}_1(x) \) which satisfy \( |x_d + |x'|^2 - 2x' \cdot s| \leq 3^{-1}|x'|^{-1}|x_d + |x'|^2| \) is
\[
\lesssim \sum_{j=1}^{[\log_2 |x_1|+1]} \int_{E_j^2(x)} \langle s \rangle^{-d} ds
\]
\[
\lesssim \sum_{j=1}^{\infty} (2^j |x'|^{-1}|x_d + |x'|^2| - |x'|^{d-2})^j |x'|^{-d} |x_d + |x'|^2|^d - 1 \leq |x_d + |x'|^2|^{-1}.
\]
This completes the discussion of \( J_1(x) \).

We turn to the discussion of \( J_2(x) = \int_{\mathcal{R}_2(x)} \langle x_d + |x'|^2 - 2x' \cdot s \rangle^{-d} ds \). As already noted, we may continue to assume that \( \max(|x'|, |x_d + |x'|^2|) \gg 1 \).

- **Estimate for \( J_2(x) \) in the case \( |x'| \leq |x_d + |x'|^2| \) and \( |x'| > 1 \).** If \( |x_d + |x'|^2 - 2x' \cdot s| \leq \frac{|x_d + |x'|^2|}{4|x'|} \), then \( |2x' \cdot s| \geq |x_d + |x'|^2|(1 - \frac{1}{4|x'|}) \). Therefore
\[
(8.12) \quad |s| \geq \frac{|x_d + |x'|^2|}{2|x'|}(1 - \frac{1}{4|x'|}) \geq \frac{3}{4} \frac{|x_d + |x'|^2|}{|x'|} \geq \frac{3}{2} |x_d + |x'|^2| - 2x' \cdot s|.
\]
This contradicts the definition of \( \mathcal{R}_2(x) \). We conclude that if \( |x'| \leq |x_d + |x'|^2| \) and \( |x'| > 1 \), then
\[
|x_d + |x'|^2 - 2x' \cdot s| > \frac{|x_d + |x'|^2|}{4|x'|}.
\]
Define
\[
E_j^3(x) = \{ s \in \mathcal{R}_2(x) : 2^j |x'|^{-1}|x_d + |x'|^2| < |x_d + |x'|^2| - 2x' \cdot s| \leq 2^{j+1}|x'|^{-1}|x_d + |x'|^2| \}.
\]
If \( s \in E_j^3(x) \) and \( \bar{s} \) is fixed, then since \( |x_1| \sim |x'| \), \( s_1 \) lies in an interval of length \( \sim \frac{2^j|x_d + |x'|^2|}{|x'|} \). From the bound
\[
|\bar{s}| \leq |s| \leq |x_d + |x'|^2| - 2x' \cdot s| \leq 2^{j+1}|x_d + |x'|^2||x'|^{-1},
\]
it now follows that
\[
|E_j^3(x)| \lesssim |x'|^{-2} |x_d + |x'|^2| (2^j |x'|^{-1}|x_d + |x'|^2|) d-2 = 2^{j+1}|x'|^{-d} |x_d + |x'|^2|^{d-1}.
\]
Therefore if \( |x'| \leq |x_d + |x'|^2| \) and \( |x'| > 1 \), the contribution of \( \mathcal{R}_2(x) \) to \( J_2(x) \) is
\[
\sum_{j=0}^{\infty} \int_{E_j^3(x)} \langle x_d + |x'|^2 - 2x' \cdot s \rangle^{-d} ds
\]
\[
\lesssim \sum_{j=0}^{\infty} (2^j \frac{|x_d + |x'|^2|}{|x'|})^{-d} 2^j (d-1)|x'|^{-d} |x_d + |x'|^2|^{d-1} = |x_d + |x'|^2|^{-1}.
\]

- **Estimate for \( J_2(x) \) in the case \( |x'| \leq |x_d + |x'|^2| \) and \( |x'| \leq 1 \).** Let \( s \in \mathcal{R}_2(x) \). Then
\[
|x_d + |x'|^2 - 2x' \cdot s| \geq |x_d + |x'|^2| - 2|s \cdot x'|
\]
\[
\geq |x_d + |x'|^2| - 2|s|
\]
\[
\geq |x_d + |x'|^2| - 2|x_d + |x'|^2 - 2x' \cdot s|,
\]
where the definition of $\mathcal{R}_2(x)$ was invoked to obtain the last inequality. This implies that

$$|x_d + |x'|^2 - 2x' \cdot s| \geq \frac{1}{3}|x_d + |x'|^2|.$$  

Define

$$E_3^j(x) = \{ s \in \mathcal{R}_2(x) : 2^j|x_d + |x'|^2| \leq 3|x_d + |x'|^2 - 2x' \cdot s| < 2^{j+1}|x_d + |x'|^2| \}.$$  

Any $s \in E_3^j(x)$ satisfies

$$|s| \leq |x_d + |x'|^2| - 2|x_d + |x'|^2 - 2x' \cdot s| \sim 2^j|x_d + |x'|^2|,$$

so $|E_3^j(x)| \lesssim (2^j|x_d + |x'|^2|)^{d-1}$, whence

$$J_2(x) = \int_{\mathcal{R}_2(x)} \langle x_d + |x'|^2 - 2x' \cdot s \rangle^{-d} ds \lesssim \sum_{j=0}^{\infty} (2^j|x_d + |x'|^2|)^{-d}|E_3^j(x)|$$

$$\lesssim \sum_{j=0}^{\infty} (2^j|x_d + |x'|^2|)^{-d}(2^j|x_d + |x'|^2|)^{d-1} = C|x_d + |x'|^2|^{-1}.$$  

This concludes the analysis of $J_2(x)$, in the case in which $|x'| \leq |x_d + |x'|^2|$.  

- **Estimate for $J_2(x)$ in the case $|x'| \geq |x_d + |x'|^2|$.** We may continue to assume that $|x_1| \geq |x'|/\sqrt{d} - 1$. Partition $\mathcal{R}_2(x)$ into the following three subregions:

$$|x_d + |x'|^2 - 2x' \cdot s| > 2|x'|,$$

$$1 < |x_d + |x'|^2 - 2x' \cdot s| \leq 2|x'|,$$

$$|x_d + |x'|^2 - 2x' \cdot s| \leq 1.$$  

To analyze the contribution of the subregion in which $|x_d + |x'|^2 - 2x' \cdot s| > 2|x'|$, for each integer $j \geq 1$ define

$$E_4^j(x) = \{ s \in \mathcal{R}_2(x) : 2^j|x'| < |x_d + |x'|^2 - 2x' \cdot s| \leq 2^{j+1}|x'| \}.$$  

For fixed $\bar{s}$, $s_1$ lies in a subinterval of length $\sim 2^j$, while $|\bar{s}| \leq |s| \leq |x_d + |x'|^2 - 2x' \cdot s| \lesssim 2^j|x'|$, so $|E_4^j(x)| \lesssim 2^{(d-1)|x'|^{d-2}}$. Therefore the contribution of this subset of $\mathcal{R}_2(x)$ to $J_2(x)$ is

$$\lesssim \sum_{j=1}^{\infty} (2^j|x'|)^{-d}|E_4^j(x)| \lesssim \sum_{j=1}^{\infty} (2^j|x'|)^{-d}2^{j(d-1)|x'|^{d-2}} = |x'|^{-2} \ll |x'|^{-1}.$$  

To analyze the contribution of the subregion in which $1 < |x_d + |x'|^2 - 2x' \cdot s| \leq 2|x'|$, partition further into subregions in which $|x_d + |x'|^2 - 2x' \cdot s| \leq 2|x'| \sim 2^{-j}|x'|$, where $1 \leq 2^j \lesssim |x'|$. Such a sub-subregion has measure $\lesssim 2^{-j}(2^{-j}|x'|)^{d-2} = 2^{-j(d-1)}|x'|^{d-2}$, giving a total contribution to $J_2(x)$ which is

$$\lesssim \sum_{1 \leq 2^j \lesssim |x'|} (2^{-j}|x'|)^{-d}2^{-j(d-1)|x'|^{d-2}} = \sum_{1 \leq 2^j \lesssim |x'|} 2^j|x'|^{-2} \lesssim |x'|^{-1},$$

as required.

It remains only to analyze the contribution made to $J_2(x)$ by the subregion in which $|x_d + |x'|^2 - 2x' \cdot s| \leq 1$, assuming still that $|x'| \geq |x_d + |x'|^2|$. Since $|s| \leq$
\[ |x_d + |x'|^2 - 2x' \cdot s| \leq 1, \text{ and since } s_1 \text{ lies in an interval of length } \lesssim |x_1|^{-1} \lesssim |x'|^{-1} \text{ so long as } \tilde{s} \text{ remains fixed, the measure of this subregion is } \lesssim |x'|^{-1}. \text{ The integrand } \langle x_d + |x'|^2 - 2x' \cdot s \rangle^{-d} \text{ is } \lesssim 1, \text{ so its integral over this subregion is } \lesssim |x'|^{-1}. \]

The proof of Lemma 2.1 is complete. □

Corollary 8.1. Let \( f_0 = \chi_{|x| \leq 1} \). Define \( f_{n+1} \) to be \((T(f_n))^d\) if \( n \) is even, and \((T^*(f_n))^d\) if \( n \) is odd. Then for any even \( n > 0 \),

\[ f_n \leq C^n v_* . \]

This follows at once from \( n \) applications of Lemma 2.1 since \( 0 \leq f_0 \leq v_* \).

9. PROOF OF LEMMA 7.1

The following argument is essentially taken from [12].

Proof. Fix a smooth, compactly supported cutoff function \( \eta \in C_0^\infty(\mathbb{R}^d) \) satisfying \( \eta(\xi) \equiv 1 \) for all \( |\xi| \leq 1 \), and \( \eta(\xi) = 0 \) for all \( |\xi| \geq 2 \). For each \( k \in \{0, 1, 2, \ldots\} \) introduce the Fourier multiplier \( P_k \) defined by \( \hat{P}_k f(\xi) = \hat{f}(\xi)\eta(2^{-k}\xi) \). For \( k \geq 1 \) define \( Q_k = P_k - P_{k-1} \). Observe that \( \hat{Q}_k f(\xi) \) is supported in \( \{\xi : 2^{k-1} \leq |\xi| \leq 2^{k+1}\} \). Fix \( K \) such that \( 2^K \geq \Lambda \geq 2^{k-1} \). Set \( R_K f = f - P_K f \). \( \hat{R}_K f(\xi) \) is supported in \( \{\xi : |\xi| \geq 2^K\} \).

Decompose

\[ f = P_0 f + \sum_{k=1}^K Q_k f + R_K f, \]

and decompose \( g \) in the same manner. By expanding the product \( fg \) in terms of these summands and recombining terms, one obtains

\begin{align}
(9.1) \quad fg &= \sum_{k=3}^K Q_k f \cdot P_{k-3} g + \sum_{k=3}^K Q_k g \cdot P_{k-3} f \\
(9.2) &\quad + \sum_{k=2}^K Q_k f (Q_{k-2} g + Q_{k-1} g + Q_k g) + \sum_{k=2}^K Q_k g (Q_{k-2} f + Q_{k-1} f) \\
(9.3) &\quad + R_K f \cdot P_{K-2} g + Q_{K} g \cdot P_{K-2} f \\
&\quad + R_K f (Q_{K-1} g + Q_K g) + R_K g (Q_{K-1} f + Q_K f) + R_K f \cdot R_K g \\
&\quad + \mathcal{R}(f, g)
\end{align}

where \( \mathcal{R}(f, g) \) is a constant-coefficient finite linear combination of twofold products of the factors \( P_0 f, P_0 g, Q_1 f, Q_1 g \).

Consider the contribution made to \( D_\Lambda^*(fg) \) by the first term made on the right-hand side in this equation. The Fourier transform of \( D_\Lambda^*(Q_k f \cdot P_{k-3} g) \) is supported in \( \{\xi : 2^{k-2} \leq |\xi| \leq 2^{k+2}\} \). Therefore by weighted Littlewood-Paley theory \[19\], since \( u \in L^r(\mathbb{R}^d) \),

\[ \left\| \sum_{k=2}^K D_\Lambda^*(Q_k f \cdot P_{k-3} g) \right\|_{L^r(u)} \lesssim \left\| \left\{ D_\Lambda^*(Q_k f \cdot P_{k-3} g) \right\}_{k=2}^K \right\|_{L^r(\hat{\mathbb{R}}^d)(u)} \]
where
\[ \| \{ h_k \} \|_{L^r(\ell^2)(u)} = \int_{\mathbb{R}^d} \left( \sum_k |h_k(x)|^2 \right)^{r/2} u(x) \, dx.\]

Writing \( \hat{D}_A^s f(\xi) = \lambda_{s,\Lambda}(\xi) \hat{f}(\xi) \), let \( m_k(\xi) = 2^{-ks} \lambda_{s,\Lambda}(\xi) \zeta(2^{-k}\xi) \) where \( \zeta(\xi) \equiv 1 \) whenever \( \frac{1}{8} \leq |\xi| \leq 4 \), and \( \zeta(\xi) \equiv 0 \) whenever \( |\xi| \leq \frac{1}{8} \) or \( |\xi| \geq 8 \).

Define \( M_k f(\xi) = m_k(\xi) \hat{f}(\xi) \). Then
\[ \| \{ D_k^s (Q_k f \cdot P_{k-3} g) \} \|_{L^r(\ell^2)(u)} = \| \{ M_k (2^{ks} Q_k f \cdot P_{k-3} g) \} \|_{L^r(\ell^2)(u)}.\]

Because \( u \in A_r \) and the operator \( \tilde{M} \{ h_k \} = \{ M_k h_k \} \) is a vector-valued Calderón-Zygmund operator, \( \tilde{M} \) is bounded on \( L^r(\ell^2)(u) \) \( \|$\| \\

\[ \| D_k^s \sum_{k=2}^K Q_k f \cdot P_{k-3} g \|_{L^r(u)} \lesssim \| \{ 2^{ks} Q_k f \cdot P_{k-3} g \} \|_{L^r(\ell^2)(u)}.\]

Now
\[ |2^{ks} Q_k f \cdot P_{k-3} g| \leq 2^{ks} |Q_k f| \cdot |M g|\]

where \( M \) denotes the Hardy-Littlewood maximal function. Therefore by Hölder’s inequality and the factorization \( u = u_1 v_1 \),
\[ \| \{ 2^{ks} Q_k f \cdot P_{k-3} g \} \|_{L^r(\ell^2)(u)} \leq C \| M g \| \| \sum_{k=2}^K |2^{ks} Q_k f|^2 \|^{1/2}_{L^r(u)} \]
\[ \leq C \| M g \|_{L^{q_1/q_1}(u_1^{q_1/r})} \| \{ 2^{ks} Q_k f \} \|_{L^{p_1}(\ell^2)(u_1^{p_1/r})}.\]

Since \( u_1^{q_1/r} \in A_{q_1} \) and \( M \) is bounded on \( L^{q_1} \) with respect to any weight in \( A_{q_1} \) \( [19] \), this is majorized by
\[ \| g \|_{L^{q_1/q_1}(u_1^{q_1/r})} \| \{ 2^{ks} Q_k f \} \|_{L^{p_1}(\ell^2)(u_1^{p_1/r})}.\]

Again by weighted vector-valued Calderón-Zygmund theory \( \|$\| \\

The contributions of the second term on the right in \( \|$\| \\

We discuss next the contribution of \( \sum_{k=2}^K Q_k f \cdot Q_k g \). The summand \( Q_k f \cdot Q_k g \) has Fourier transform supported in \( \{ \xi : |\xi| \leq 2^{k+2} \} \) and therefore
\[ D_k^s (\sum_{k=2}^K Q_k f \cdot Q_k g) = \sum_{k=2}^K M_k (2^{ks} Q_k f \cdot Q_k g)\]

where \( M_k \) is the Fourier multiplier operator with multiplier
\[ m_k(\xi) = \lambda_{s,\Lambda}(\xi) 2^{-ks} \eta(2^{-k-2}\xi).\]
It is routine to verify, using the hypothesis that $s \geq 0$, that $|M_k h| \leq C M(h)$ for any function $h$, uniformly in $k, \Lambda$ for $0 \leq k \leq K$. Therefore
\[
\|D^s_{\Lambda} \left( \sum_{k=2}^{K} Q_k f \cdot Q_k g \right) \|_{L^r(u)} \leq C \| \sum_{k=2}^{K} Q_k f \cdot Q_k g \|_{L^r(u)} \\
\leq C \| \sum_{k=2}^{K} 2^{ks} Q_k f \cdot Q_k g \|_{L^r(u)} \\
\leq C \left( \sum_{k=2}^{K} |2^{ks} Q_k f|^2 \right)^{1/2} \left( \sum_{k=2}^{K} |Q_k g|^2 \right)^{1/2} \|_{L^r(u)} \\
\leq C \left\| \left\{ 2^{ks} Q_k f \right\}_{k=2}^{K} \right\|_{L^{p_1}(u_1^{r_1/\ell})} \cdot \left\| \{ Q_k g \}_{k=2}^{K} \right\|_{L^{p_1}(u_2^{q_1/\ell})} \\
\leq C \| D^s_{\Lambda} f \|_{L^{p_1}(u_1^{r_1/\ell})} \| g \|_{L^{p_1}(u_2^{q_1/\ell})} ,
\]
as desired.
All remaining terms can be treated in the same way as we have done for $\sum_{k=2}^{K} Q_k f \cdot Q_k g$. □

**References**

[1] K. F. Andersen and R. T. John, *Weighted inequalities for vector-valued maximal functions and singular integrals*, Studia Math. 69 (1980/81), no. 1, 19-31.
[2] W. Beckner, *Inequalities in Fourier analysis*, Ann. of Math. (2) 102 (1975), no. 1, 159-182.
[3] J. Bennett, N. Bez, A. Carbery, and D. Hundertmark, *Heat-flow monotonicity of Strichartz norms*, preprint, arXiv:0809.4783.
[4] J. Bennett, A. Carbery, M. Christ, and T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure, and extremals*, The Brascamp-Lieb inequalities: finiteness, structure and extremals. Geom. Funct. Anal. 17 (2008), no. 5, 1343-1415.
[5] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. 12 (1984), no. 3, 647-702.
[6] M. Christ, *Quasiextremals for a Radon-like transform*, preprint.
[7] , *Extremals for a Radon-like transform*, preprint.
[8] , *Regularity of solutions of certain Euler-Lagrange equations*, in preparation.
[9] M. Christ and R. Quiñodr´an, *Gaussians rarely extremize adjoint Fourier restriction inequalities for paraboloids*, preprint, math.CA, arXiv:1012.1346.
[10] M. Christ and S. Shao, *Existence of Extremals for a Fourier Restriction Inequality*, to appear, Analysis and Partial Differential Equations, arXiv:1006.4319.
[11] , *On the extremizers of an adjoint Fourier restriction inequality*, preprint, arXiv:1006.4318
[12] F. M. Christ and M. Weinstein, *Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation*, J. Funct. Anal. 100 (1991), no. 1, 87-109.
[13] D. Foschi, *Maximizers for the Strichartz inequality*, J. Eur. Math. Soc. (JEMS) 9 (2007), no. 4, 739-774.
[14] D. Hundertmark and V. Zharnitsky, *On sharp Strichartz inequalities in low dimensions*, Int. Math. Res. Not. 2006, Art. ID 34080.
[15] M. Kunze, *On the existence of a maximizer for the Strichartz inequality*, Comm. Math. Phys. 243 (2003), 137-162.
[16] E. Lieb, *Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities*, Ann. of Math. (2) 118 (1983), no. 2, 349-374.
[17] Gaussian kernels have only Gaussian maximizers, Invent. Math. 102 (1990), no. 1, 179-208.

[18] S. K. Pichorides, On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov, Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, II. Studia Math. 44 (1972), 165-179.

[19] E. M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy, Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993

[20] B. Stovall, Quasi-extremals for convolution with surface measure on the sphere, Illinois J. Math. 53 (2009), no. 2, 391-412.

Michael Christ, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA
E-mail address: mchrist@math.berkeley.edu

Qingying Xue, School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing 100875, People’s Republic of China
E-mail address: qyxue@bnu.edu.cn