Vanishing mean oscillation and regularity
in the calculus of variations

by

Roger Moser

Preprint no.: 96 2001
Vanishing mean oscillation and regularity in the calculus of variations

Roger Moser
MPI for Mathematics in the Sciences
D-04103 Leipzig (Germany)

November 26, 2001

Abstract

We consider critical points and solutions of the gradient flow for variational integrals with integrands satisfying a Legendre-Hadamard condition. We show that if a solution of the Euler-Lagrange system or the gradient flow has first (spatial) derivatives which are bounded and of vanishing mean oscillation, higher regularity follows.

1 Introduction

Let $\Omega \subseteq \mathbb{R}^n$ be a bounded open domain and $f \in C^2(\mathbb{R}^{m \times n})$ a function which satisfies the Legendre-Hadamard condition

$$\sum_{i,j=1}^m \sum_{\alpha,\beta=1}^n \frac{\partial^2 f}{\partial \xi_i^\alpha \partial \xi_j^\beta} (\xi) \partial^j \zeta^i \eta_\alpha \eta_\beta \geq \lambda |\zeta|^2 |\eta|^2$$

(1)

for some $\lambda > 0$ and for all $\xi \in \mathbb{R}^{m \times n}$, $\zeta \in \mathbb{R}^m$, and $\eta \in \mathbb{R}^n$. Moreover, we require that the second derivatives of $f$ are bounded, that is

$$|D^2 f| \leq \Lambda$$

for some $\Lambda > 0$. We consider the functional

$$F(u) = \int_\Omega f(\nabla u) \, dx$$

for $u \in H^1(\Omega, \mathbb{R}^m)$. We study critical points thereof, i.e. solutions of the Euler-Lagrange system

$$\text{div} Df(\nabla u) = 0$$

(2)

in $\Omega$, which is short for

$$\sum_{\alpha=1}^n \frac{\partial}{\partial x_\alpha} \left( \frac{\partial f}{\partial \xi_\alpha^i} (\nabla u) \right) = 0, \quad i = 1, \ldots, m,$$

and solutions of the corresponding gradient flow,

$$\frac{\partial u}{\partial t} - \text{div} Df(\nabla u) = 0$$

(3)

1
in $\Omega \times (0, T)$ for some $T > 0$. It is easy to see that smooth solutions of (3) satisfy

$$\frac{d}{dt} \int_{\Omega \times \{t\}} f(\nabla u) \eta \, dx$$

$$= \int_{\Omega \times \{t\}} \left( f(\nabla u) \frac{\partial \eta}{\partial t} - \left| \frac{\partial u}{\partial t} \right|^2 \eta - \sum_{i=1}^{m} \sum_{\alpha=1}^{n} \frac{\partial^2 f}{\partial \xi^i \partial x_\alpha} (\nabla u)^i \frac{\partial \eta}{\partial t} \frac{\partial \eta}{\partial x_\alpha} \right) \, dx$$

for all $\eta \in C^\infty_0(\Omega \times (0, T))$ and $0 < t < T$. For weak solutions, we will impose this as an additional condition.

Under the stronger assumption that $f$ is uniformly strictly quasiconvex, Evans [3] proved the following partial regularity result for minimizers $u$ of the functional $F$ (see also Evans–Gariepy [5]): There exists an open set $\Omega_0 \subset \Omega$ with $|\Omega_0| > 0$, such that $u$ is $C^{1,\alpha}$-regular in $\Omega_0$ for some $\alpha > 0$. This result has been improved by Acerbi–Fusco [1], Fusco–Hutchinson [6], Giaquinta–Modica [7], and Kristensen–Taheri [9], among others. In all these papers it is crucial that $u$ satisfies some minimality condition. Indeed there is an example by Miller–Sverdlov [11] of a function $f$ and a critical point $u$ of the corresponding functional that is Lipschitz continuous; but not $C^1$-smooth in any open subset of $\Omega$. Hence the Euler-Lagrange system (2) alone cannot give any Hölder continuity of the gradient.

On the other hand, if we do have $C^{1,\alpha}$-regularity for a solution of (2), and if the function $f$ is smooth, then it is possible to prove also $C^\infty$-regularity, and the same is true for solutions of (3). The question that we study in this note is, how much initial regularity is necessary to obtain higher regularity? One needs more than Lipschitz regularity, as the example mentioned above shows, but possibly less than $C^1$-regularity is sufficient. Results by Chiarenza–Frasca–Longo [2] and Ragusa [12] on solutions of elliptic equations with coefficients in the space VMO (see definition below) suggest that the condition $\nabla u \in \text{VMO}(\Omega, \mathbb{R}^{m \times n})$ for a solution of (2) or a similar condition for a solution of (3) might imply higher regularity. This turns out to be the case.

The space $\text{VMO}(\Omega, \mathbb{R}^{m \times n})$ of functions of vanishing mean oscillation in $\Omega$ with values in $\mathbb{R}^{m \times n}$ is a subspace of the John–Nirenberg space $\text{BMO}(\Omega, \mathbb{R}^{m \times n})$. These spaces can be defined as follows (cf. [8] and [13]).

For $x_0 \in \mathbb{R}^n$ and $r > 0$, we write $B_r(x_0)$ for the open ball in $\mathbb{R}^n$ with center $x_0$ and radius $r$. Let $\phi \in L^1(\Omega, \mathbb{R}^{m \times n})$. We define

$$[\phi]_{\text{BMO}(\Omega)} = \sup_{B_r(x_0)} \int_{B_r(x_0)} |\phi - (\phi)_{B_r(x_0)}| \, dx,$$

where we use the notation

$$(\phi)_{B_r(x_0)} = \int_{B_r(x_0)} \phi \, dx = \frac{1}{|B_r(x_0)|} \int_{B_r(x_0)} \phi \, dx.$$

We set

$$\text{BMO}(\Omega, \mathbb{R}^{m \times n}) = \{ \phi \in L^1(\Omega, \mathbb{R}^{m \times n}) : [\phi]_{\text{BMO}(\Omega)} < \infty \}.$$
and $\text{VMO}(\Omega, \mathbb{R}^{m \times n})$ the space of functions $\phi \in \text{BMO}(\Omega, \mathbb{R}^{m \times n})$ such that

$$
\sup_{r \leq \rho} \sup_{B_r(x_0) \subset \Omega} \int_{B_r(x_0)} |\phi - (\phi)_{B_r(x_0)}| \, dx \to 0 \quad \text{as } \rho \to 0.
$$

For the parabolic problem (3), it is natural to work with parabolic cylinders $P_r(z_0) = B_r(z_0) \times (t_0 - r^2, t_0 + r^2)$, where $z_0 = (x_0, t_0)$. Instead of balls, replacing everywhere $B_r(z_0)$ by $P_r(z_0)$ and $\Omega$ by an open set $\Omega' \subset \Omega \times \mathbb{R}$ in the definitions above, we obtain the spaces $\text{BMO}'(\Omega', \mathbb{R}^{m \times n})$ and $\text{VMO}'(\Omega', \mathbb{R}^{m \times n})$.

We have the following results.

**Theorem 1.1** For any $L > 0$ there exist constants $\epsilon > 0$ and $\alpha > 0$ with the following properties.

(i) If $u \in H^1(\Omega, \mathbb{R}^m)$ is a weak solution of (2) satisfying $\|\nabla u\|_{L^\infty(\Omega)} \leq L$ and $\|\nabla u\|_{\text{BMO}(\Omega)} \leq \epsilon$, then $u \in C^{1,\alpha}_{\text{loc}}(\Omega, \mathbb{R}^m)$.

(ii) If $u \in H^1(\Omega \times (0,T), \mathbb{R}^m)$ is a weak solution of (3) satisfying (4) and the estimates $\|\nabla u\|_{L^\infty(\Omega \times 0,T)} \leq L$ and $\|\nabla u\|_{\text{BMO}(\Omega \times 0,T)} \leq \epsilon$, then $u \in C^{0,\alpha}_{\text{loc}}(\Omega \times (0,T), \mathbb{R}^{m \times n})$.

**Corollary 1.1**

(i) Let $u \in H^1(\Omega, \mathbb{R}^m)$ be a weak solution of (2). If $\nabla u \in L^\infty(\Omega, \mathbb{R}^{m \times n}) \cap \text{VMO}(\Omega, \mathbb{R}^{m \times n})$, then $u \in C^{1,\alpha}_{\text{loc}}(\Omega, \mathbb{R}^m)$ for some $\alpha > 0$.

(ii) Let $u \in H^1(\Omega \times (0,T), \mathbb{R}^m)$ be a weak solution of (3) satisfying (4). If $\nabla u \in L^\infty(\Omega \times (0,T), \mathbb{R}^{m \times n}) \cap \text{VMO}'(\Omega \times (0,T), \mathbb{R}^{m \times n})$, then $u \in C^{0,\alpha}_{\text{loc}}(\Omega \times (0,T), \mathbb{R}^{m \times n})$ for some $\alpha > 0$.

The corollary follows immediately from the theorem, and (i) follows from (ii) by considering time-independent solutions of (3). Thus we concentrate on the evolution equation (3) in the rest of the paper.

### 2 Blow-up

The following arguments are due to Evans [3]. We sketch them briefly for the convenience of the reader.

We use the abbreviation

$$
E(u, z_0, r) = \int_{P_r(z_0)} |\nabla u - (\nabla u)_{P_r(z_0)}|^2 \, dz
$$

for $P_r(z_0) \subset \mathbb{R}^n \times \mathbb{R}$ and $u \in H^1(P_r(z_0), \mathbb{R}^m)$.

**Lemma 2.1** For any $L > 0$ and any $\delta > 0$ there exist $\epsilon > 0$ and $\theta \in (0, 1)$, such that for any $P_r(z_0) \subset \mathbb{R}^n \times \mathbb{R}$ and any weak solution $u \in H^1(P_r(z_0), \mathbb{R}^m)$ of (3) and (4) with the properties $|(\nabla u)_{P_r(z_0)}| \leq L$ and $E(u, z_0, r) \leq \epsilon^2$, we have

$$(\theta r)^{-\frac{1}{2}} \int_{P_r(z_0)} |u(x) - u)_{P_r(z_0)} - (\nabla u)_{P_r(z_0)} \cdot x|^2 \, dz \leq \delta E(u, z_0, r).$$
Proof. Fix $L > 0$ and $\delta > 0$. If the claim were false for this particular choice, then for any fixed $\theta \in (0, 1)$, we could find a sequence of parabolic cylinders $P_{r_k}(z_k) \subset \mathbb{R}^n \times \mathbb{R}$ and a sequence of solutions $u_k \in H^1(P_{r_k}(z_k), \mathbb{R}^m)$ of (3) and (4), such that $\|\nabla u_k\|_{P_{r_k}(z_k)} \leq L$ and

$$E(u_k, z_k, r_k) = \varepsilon_k^2 \to 0 \quad \text{as} \quad k \to \infty,$$

but

$$(\theta r_k)^{-2} \int_{P_{r_k}(z_k)} |u_k(x) - (u_k)_{P_{r_k}(z_k)} - (\nabla u_k)_{P_{r_k}(z_k)} \cdot x|^2 \, dz \geq \delta \varepsilon_k^2.$$ 

Write $z_k = (x_k, t_k)$, and define $\xi_k = (\nabla u_k)_{P_{r_k}(z_k)}$. Set

$$v_k(x, t) = \frac{1}{\varepsilon_k r_k} (u(x_k + r_k x, t_k + r_k^2 t) - (u_k)_{P_{r_k}(z_k)} - r_k \xi_k \cdot x)$$

and

$$f_k(\xi) = \frac{1}{\varepsilon_k} (f(\xi_k + \varepsilon_k \xi) - f(\xi_k) - Df(\xi_k) \xi).$$

Then we have

$$(v_k)_{P_{1}(0)} = 0, \quad (\nabla v_k)_{P_{1}(0)} = 0, \quad \int_{P_{1}(0)} |\nabla v_k|^2 \, dz = 1,$$

and

$$(\theta^{-2} \int_{P_{r_k}(z_k)} |v_k - (v_k)_{P_{r_k}(z_k)} - (\nabla v_k)_{P_{r_k}(z_k)} \cdot x|^2 \, dz \geq \delta.) \quad (5)$$

Furthermore,

$$f_k(0) = 0, \quad Df_k(0) = 0, \quad |D^2 f_k| \leq \Lambda.$$ 

It is easy to check that the $v_k$ still satisfy (3) and (4) with $f$ replaced by $f_k$. From (4) we obtain a uniform bound on $\|v_k\|_{H^1(P_{1/2}(0), \mathbb{R}^m)}$ weakly in $H^1$ and strongly in $L^2$. Furthermore we may assume that $\xi_k \to \xi_0$ as $k \to \infty$. Note that we have for $\phi \in C_0^\infty(P_{1/2}(0), \mathbb{R}^m)$ the identity

$$0 = \int_{P_{1/2}(0)} \left( \frac{\partial v_k}{\partial t} \cdot \phi + Df_k(\nabla v_k) \nabla \phi \right) \, dz$$

$$= \int_{P_{1/2}(0)} \left( \frac{\partial v_k}{\partial t} \cdot \phi + \int_0^1 D^2 f_s(\nabla v_k)(\nabla v_k, \nabla \phi) \, ds \right) \, dz.$$ 

From the fact that $D^2 f_k(\xi) = D^2 f(\xi_k + \varepsilon_k \xi)$, we infer that

$$\int_0^1 D^2 f_k(s \nabla v_k) \, ds \to D^2 f(\xi_0)$$

strongly in $L^2$. Hence $v$ satisfies

$$\frac{\partial v_i}{\partial t} - \sum_{j=1}^m \sum_{\alpha, \beta = 1}^n \frac{\partial}{\partial x_\alpha} \left( \frac{\partial^2 f}{\partial \xi_i \partial \xi_\beta} (\xi_0) \frac{\partial v_i}{\partial x_\beta} \right) = 0, \quad i = 1, \ldots, m.$$
Since (1) holds, we can apply linear regularity theory to \( v \). We obtain the estimate
\[
\int_{P_r(0)} |v - (v)_{P_r(0)} - (\nabla v)_{P_r(0)} \cdot x|^2 \, dz \leq C \theta^4
\]
for a constant \( C = C(n, m, \lambda, \Lambda) \). If we choose \( \theta < C^{-1/4} \delta^{1/4} \), then this contradicts (5) by the strong convergence of \( v_k \) to \( v \) in \( L^2 \).

\[\square\]

3 Estimates involving the BMO-norm

In this section we will prove Theorem 1.1, using certain ideas of Kristensen-Taheri [9]. For this we will need the following inequality for the parabolic Hardy–Littlewood and Fefferman–Stein maximal functions
\[
\phi^*(z_0) = \sup_{P_r(z_1) \ni z_0} \int_{P_r(z_1)} |\phi| \, dz
\]
and
\[
\phi^\#(z_0) = \sup_{P_r(z_1) \ni z_0} \int_{P_r(z_1)} |\phi - (\phi)_{P_r(z_1)}| \, dz
\]
for \( \phi \in L^1_{\text{loc}}(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^{m \times n}) \). It is a version of Lemma 6.1 in [9], and its proof is the same, except that instead of cubes in \( \mathbb{R}^n \), one uses sets of the form \([x_1 - r, x_1 + r] \times \ldots \times [x_n - r, x_n + r] \times [t - r^2, t + r^2] \).

**Lemma 3.1** Let \( \Phi : [0, \infty) \to [0, \infty) \) be a continuous and increasing function with \( \Phi(0) = 0 \). There exist constants \( C = C(n) > 0 \) and \( A = A(n) \geq 1 \), such that for any \( \delta > 0 \) and any \( \phi \in L^\infty(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^{m \times n}) \) with compact support, the inequality
\[
\int_{\mathbb{R}^n \times \mathbb{R}} \Phi(\phi^*) \, dz \leq \frac{C}{\delta} \int_{\mathbb{R}^n \times \mathbb{R}} \Phi\left( \frac{\phi^\#}{\delta} \right) \, dz + C \delta \int_{\mathbb{R}^n \times \mathbb{R}} \Phi(A \phi^*) \, dz
\]
is satisfied.

This allows to prove the following estimate.

**Lemma 3.2** For any \( L > 0 \) there exist constants \( \epsilon > 0 \) and \( \theta \in (0, 1) \) and a continuous function \( \omega : [0, \infty) \to [0, \infty) \) with \( \omega(0) = 0 \), such that for any \( P_r(z_0) \subset \mathbb{R}^n \times \mathbb{R} \) and any solution \( u \in H^1(P_r(z_0), \mathbb{R}^m) \) of (3) and (4) satisfying \( ||\nabla u||_{L^\infty} \leq L \) and \( E(u, z_0, r) \leq \epsilon^2 \), the inequality
\[
E(u, z_0, \theta r) + (\theta r)^2 \int_{P_{\theta r}(z_0)} \left| \frac{\partial u}{\partial r} \right|^2 \, dz
\leq \left( \frac{1}{2} + \omega(||\nabla u||_{\text{BMO}(P_r(z_0))}) \right) \left( E(u, z_0, r) + r^2 \int_{P_r(z_0)} \left| \frac{\partial u}{\partial r} \right|^2 \, dz \right)
\]
holds.
Proof. Let \( \rho : [0, 2L] \to [0, \infty] \) be the continuous function
\[
\rho(\alpha) = \sup_{k_1, k_2 \leq L} \sup_{k_0, k_1 \leq \alpha} |D^2f(\xi_1) - D^2f(\xi_2)|,
\]
and define \( \sigma : [0, \infty) \to [0, \infty) \) to be the concave and non-decreasing function
\[
\sigma(\alpha) = \inf \{ t(\alpha) : t : \mathbb{R} \to \mathbb{R} \text{ is linear and increasing with } t \geq \rho \text{ in } [0, 2L] \}.
\]
Then clearly \( \sigma \geq \rho \) and \( \sigma(0) = 0 \). Moreover, \( \sigma \) has the property
\[
\sigma(A\alpha) \leq A\sigma(\alpha)
\]
for all \( A \geq 1 \), since it is concave.

We prove that for any \( \delta > 0 \), we have
\[
E(u, z_0, r/2) \leq \left( \delta + C_0\sigma(\|\nabla u\|_{BMO(P_r(z_0))}) \right) \left( E(u, z_0, r) + r^2 \int_{P_r(z_0)} \left| \frac{\partial u}{\partial t} \right|^2 dz \right) + C_0 r^{-2} \int_{P_r(z_0)} |u - (u)_{P_r(z_0)} - (\nabla u)_{P_r(z_0)} \cdot x|^2 dz \tag{6}
\]
for a constant \( C_0 = C_0(n, \lambda, \Lambda, \delta) \). Then the claim follows from (4) and Lemma 2.1.

By the same transformations as in the proof of Lemma 2.1, we may assume that \( P_r(z_0) = P_1(0) \) and
\[
(u)_{P_1(0)} = 0, \quad (\nabla u)_{P_1(0)} = 0,
\]
as well as
\[
f(0) = 0, \quad Df(0) = 0.
\]

Note that (1) implies
\[
\lambda \int_{B_1(0)} |\nabla \phi|^2 dx \leq \int_{B_1(0)} D^2f(0)(\nabla \phi, \nabla \phi) dx
\]
for all \( \phi \in H^1_0(B_1(0), \mathbb{R}^m) \). Choose a cut-off function \( \gamma \in C^\infty_0(P_1(0)) \) with \( 0 \leq \gamma \leq 1 \) and \( \gamma \equiv 1 \) in \( P_{1/2}(0) \). Set \( \eta = \gamma^2 \). We find that
\[
\lambda \int_{P_{1/2}(0)} |\nabla u|^2 dz \leq \lambda \int_{P_1(0)} |\nabla (\eta u)|^2 dz
\]
\[
\leq \int_{P_1(0)} \eta^2 D^2f(0)(\nabla u, \nabla u) dz + \lambda \delta \int_{P_1(0)} |\nabla u|^2 dz
\]
\[
+ C_1 \int_{P_1(0)} |u|^2 dz
\]
for a constant \( C_1 = C_1(\eta, \lambda, \Lambda, \delta) \). Furthermore, we have
\[
\int_{P_1(0)} \eta^2 \left< \frac{\partial u}{\partial t}, u \right> dz = \int_{P_1(0)} Df(\nabla u) \nabla (\eta^2 u) dz
\]
\[
= \int_{P_1(0)} \int_0^1 D^2f(s \nabla u)(\nabla u, \nabla (\eta^2 u)) ds dz
\]
\[
\geq \int_0^1 \int_{P_1(0)} \eta^2 D^2f(s \nabla u)(\nabla u, \nabla u) ds dz
\]
\[
- \lambda \delta \int_{P_1(0)} |\nabla u|^2 dz - C_2 \int_{P_1(0)} |u|^2 dz,
\]
where
\[
\eta^2(\theta) = \begin{cases} 1 & \text{if } |\theta| < \epsilon, \\
0 & \text{if } |\theta| > 1, \\
1 - |\theta|^2 & \text{if } \epsilon \leq |\theta| \leq 1,
\end{cases}
\]
and \( \epsilon, \lambda, \Lambda, \delta \) are fixed constants.
where $C_2 = C_2(\eta, \lambda, \Lambda, \delta)$. Finally,
\[
\int_{\Omega} \eta^2 \left( |D^2 f(s \nabla u) - D^2 f(0)| (\nabla u, \nabla u) \right) dz \leq \int_{\Omega} \eta^2 \sigma(\nabla u) |\nabla u|^2 \, dz \\
\leq \int_{\Omega} \sigma((\gamma^2 \nabla u)^*) |(\gamma^2 \nabla u) |^2 \, dz
\]
for $0 \leq s \leq 1$ by the definition of $\sigma$. It is easy to see that $\|\gamma^2 \nabla u\|_{\text{BMO}(\mathbb{R}^n)} \leq C_3 \|\nabla u\|_{\text{BMO}(\Omega)}$ for a constant $C_3 = C_3(n, m, \gamma)$ (cf. [4]). Applying Lemma 3.1 with $\Phi(\alpha) = \sigma(\alpha) \alpha^2$ and the Hardy–Littlewood–Wiener maximal inequality, we find that
\[
\int_{\Omega} \sigma((\gamma^2 \nabla u)^*) |(\gamma^2 \nabla u) |^2 \, dz \leq C_4 \sigma([\nabla u]_{\text{BMO}(\Omega)}) \int_{\Omega} |\nabla u|^2 \, dz,
\]
where the constant $C_4$ depends only on $m$, and $\gamma$. Combining these estimates and replacing $\delta$ by a smaller constant, we obtain (6).

**Proof of Theorem 1.1.** All that remains to be done is to estimate the expression
\[
\Phi(u, z_0, r) = \sup_{P_{z}(z) \subseteq P_{\delta}(z_0)} \int_{P_{z}(z)} \left( |\nabla u - (\nabla u)_{P_{z}(z)}|^2 + s^2 \left| \frac{\partial u}{\partial r} \right|^2 \right) \, dz
\]
with the help of Lemma 3.2, to obtain
\[
\Phi(u, z_0, \theta r) \leq \left( \frac{1}{2} + \omega ([\nabla u]_{\text{BMO}(\Omega)}) \right) \Phi(u, z_0, r),
\]
where $\theta$ and $\omega$ are determined by Lemma 3.2. If $\|\nabla u\|_{\text{BMO}(\Omega \times (0, T))}$ is sufficiently small, then we conclude that
\[
\Phi(u, z_0, r) \leq C_0 \left( \frac{r}{r_0} \right)^{\alpha} \Phi(u, z_0, r_0)
\]
for $P_{\delta}(z_0) \subseteq \Omega \times (0, T)$ and $0 < r \leq r_0$, and for a constant $C_0 = C_0(n, m, L, f)$.

The Hölder continuity of $\nabla u$ follows from this and Theorem 4.6.1 in [10]. □

**Acknowledgements.** I am grateful to S. Müller for drawing my attention to this problem, and to J. Kristensen for his helpful comments. This work was supported by a fellowship of the Swiss National Science Foundation.

**References**


