

ON THE SMOOTHNESS OF CONVEX ENVELOPES

A. GRIEWANK AND P. J. RABIER

ABSTRACT. We examine differentiability properties of the convex envelope $\text{conv } E$ of a given function $E: \mathbf{R}^n \rightarrow (-\infty, \infty]$ in terms of properties of E . It is shown that C^1 as well as optimal $C^{1,\alpha}$ regularity results, $0 < \alpha \leq 1$, can be obtained under general conditions.

1. INTRODUCTION

Let $E: \mathbf{R}^n \rightarrow (-\infty, \infty]$ be any function. The *convex envelope* (or convex hull) $\text{conv } E$ of E is unambiguously defined as the greatest convex function majorized by E on \mathbf{R}^n . Under appropriate conditions, $\text{conv } E$ will be a proper convex function, hence differentiable almost everywhere on the interior of its domain.

Surprisingly, it does not appear that much more is known about the smoothness of $\text{conv } E$, no matter what assumption is made on E (with the exception of E smooth and convex). For instance, we note that in a recent paper, Raymond [5] has sketched a proof of C^1 regularity for the convex envelope of a smooth enough function of two variables defined on the entire plane and satisfying appropriate growth conditions. Since his purpose was clearly not to establish regularity results for convex envelopes, this confirms that there is both a need for and a lack of available references in this direction. Perhaps this is due to the fact that elementary examples show that $\text{conv } E$ does not in general inherit the smoothness properties of E : even when E is analytic, it is easily seen that $\text{conv } E$ will not even be C^2 under nonexceptional circumstances.

Thus, if a convex envelope has particular smoothness properties not necessarily shared by general convex functions, they do not exceed $C^{1,1}$ regularity. This observation is strongly reminiscent of a result by Brezis [1] on the regularity threshold for solutions to the obstacle problem. Although a convex envelope

Received by the editors December 23, 1988.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 26B05, 26B25; Secondary 26B35.

This work was supported in part by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U. S. Department of Energy, under contract W-31-109-Eng-38.

The submitted manuscript has been authored by a contractor of the U. S. Government under contract No. W-31-109-Eng-38. Accordingly, the U. S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U. S. Government purposes.

is not obtained as a solution to a variational inequality, its very definition makes it appear as a solution to a variant of the obstacle problem. Further analogy can be found in that the solution to the classical obstacle problem satisfies a second order elliptic inequality, while convexity may be thought of as positive semidefiniteness of the Hessian.

Now, the regularity results for the obstacle problem are established in Sobolev spaces and do not exceed $W^{s,p}$ -regularity in general where $s < 2 + \frac{1}{p}$ and $1 < p < \infty$ is arbitrary. From Sobolev's embedding theorem, regularity in the classes $C^{k,\alpha}$ then does not exceed $C^{1,\alpha}$ with $\alpha < 1$.

Put in very broad terms, it will be shown in this paper that a similar property is true for convex envelopes, namely that general conditions can be given under which local $C^{1,\alpha}$ regularity of E ($0 \leq \alpha \leq 1$) is inherited by $\text{conv } E$. Naturally, this statement alone conceals not only the hypotheses to be made on E , but also some limitations as to where $\text{conv } E$ is of class C^1 or $C^{1,\alpha}$. Also, it does not stress that because the construction of $\text{conv } E$ involves global properties of E , appropriate smoothness of the boundary of $\text{dom}(E)$, the domain of E , may have to be required near the points of some subset of $\partial(\text{dom}(E)) \cap \text{dom}(E)$. Finally, saying that $\text{conv } E$ is "as smooth as E " is an understatement, for continuous differentiability of E is not necessary for $\text{conv } E$ to be C^1 , that is, convexification may have a smoothing effect. This phenomenon, easily corroborated by elementary one-dimensional examples, is accounted for in our analysis.

On the practical side, we note that inquiring about smoothness of convex envelopes is not meant only to satisfy mathematical curiosity. For instance, the operation of convexification is fundamental in the mathematical study of thermodynamic phase equilibria: a result of local $C^{1,\alpha}$ regularity for the convex envelope of, say, a Helmholtz energy potential E , yields local $C^{0,\alpha}$ regularity for the corresponding pressure, directly expressed in terms of $\nabla(\text{conv } E)$. Even more important is that differentiability of $\text{conv } E$, by guaranteeing uniqueness of the subgradient, ensures that the pressure is uniquely determined. Applications of some of the results presented here to thermodynamic phase equilibrium are discussed in [4].

Such smoothness results are also relevant regarding the convexification of normal integrands (a finite dimensional problem, indeed), a procedure lying at the bottom of the well-known relaxation technique in the calculus of variations. In these matters, C^1 regularity of the convex envelope is essential to ascertain that the generalized solutions satisfy an associated Euler equation in the sense of distributions. In turn, such a characterization can be used for various purposes, such as determining whether the generalized solutions are or are not solutions to the original problem. This is precisely the way the aforementioned regularity result proved in [5] is used, and was actually used before in several circumstances in problems involving convexification with respect to only one variable (a case when C^1 regularity of the convex envelope is rather obvious). That at least local $C^{1,1}$ regularity is also available should be important regarding derivatives in

the sense of distributions and/or the use of Sobolev's embedding theorems, especially in problems in which Lipschitz continuous solutions are sought.

The paper is organized as follows. General hypotheses are listed and technical preliminaries are developed in §2. §3 contains the continuous differentiability properties and $C^{0,\alpha}$ regularity of the gradient is established in §4 in a somewhat less general framework, sufficient for many applications. Despite the analogy emphasized above, our method is quite different from the one used to prove smoothness of the solutions to variational inequalities. In particular, we shall make no use of Sobolev spaces, and our approach remains quite classical.

All the known properties of convex functions that are used throughout can be found in the book by Rockafellar [6]. For the convenience of the reader, we have made precise reference to the relevant statement in [6] whenever we have felt that it might not be common knowledge. Everywhere in the paper, "int" and "dom" are abbreviations for "interior" and "domain", respectively.

2. PRELIMINARIES

In this section, we shall assume that $E: \mathbf{R}^n \rightarrow (-\infty, \infty]$ satisfies

$$(2.1) \quad \text{dom}(E) = \{\delta \in \mathbf{R}^n, E(\delta) < \infty\} \neq \emptyset,$$

$$(2.2) \quad E \text{ is lower semicontinuous on } \mathbf{R}^n,$$

$$(2.3) \quad \lim_{|\delta| \rightarrow \infty} \frac{E(\delta)}{|\delta|} = \infty,$$

where $|\cdot|$ denotes any norm on \mathbf{R}^n . Other assumptions will be introduced in §3.

Without any of the preceding assumptions, a formula for $\text{conv } E$, as it can be found in [6, Corollary 17.1.5, p. 157], reads

$$(2.4) \quad (\text{conv } E)(d) = \inf \sum_{i=1}^{n+1} \lambda_i E(d_i),$$

where the infimum is taken over all convex combinations with $n+1$ elements $d = \sum_{i=1}^{n+1} \lambda_i d_i$, $\lambda_i \geq 0$, $d_i \in \mathbf{R}^n$, $\sum_{i=1}^{n+1} \lambda_i = 1$. In this statement, it is understood that the d_i 's need not be distinct and that $\lambda_i E(d_i) = 0$ whenever $\lambda_i = 0$ and $E(d_i) = \infty$. Yet another expression for $\text{conv } E$ is [6, p. 36]

$$(2.5) \quad (\text{conv } E)(d) = \inf \sum_{i=1}^q \lambda_i E(d_i),$$

where now the infimum is taken over all convex combinations $d = \sum_{i=1}^q \lambda_i d_i$ (q arbitrary). It is then immediate that a third definition of $\text{conv } E$ is given by (2.5) with the infimum now being taken over all convex combinations with $q \leq n+1$ elements, $\lambda_i > 0$ for $1 \leq i \leq q$ and $d_i \neq d_j$ for $i \neq j$. To see this, it suffices to eliminate from (2.4) all the indices with $\lambda_i = 0$ and collect all the terms with $\lambda_i > 0$ corresponding to the same point d_i .

The preliminary result we need is that assumptions (2.1) to (2.3) are sufficient to guarantee that the infimum in (2.5) is actually a minimum for every $d \in \text{conv}(\text{dom}(E))$, the convex hull of $\text{dom}(E)$. Under a somewhat stronger growth condition at infinity, this result is already in Ekeland and Temam [2, Lemma 3.3, p. 260] when $\text{dom}(E) = \mathbf{R}^n$.

Theorem 2.1. *In the assumptions (2.1) to (2.3), $\text{conv } E$ is a proper convex function and $\text{dom}(\text{conv } E) = \text{conv}(\text{dom}(E))$. Moreover, for $d \in \text{dom}(\text{conv } E)$, one has*

$$(\text{conv } E)(d) = \sum_{i=1}^q \lambda_i E(d_i),$$

for some convex combination $d = \sum_{i=1}^q \lambda_i d_i$ with $q \leq n+1$, $\lambda_i > 0$, $d_i \in \text{dom}(E)$, and $d_i \neq d_j$ for $i \neq j$.

Proof. From assumptions (2.1) to (2.3) it follows that E is bounded from below by a constant $e \in \mathbf{R}$, namely,

$$(2.6) \quad E(\delta) \geq e, \quad \forall \delta \in \mathbf{R}^n.$$

In particular, $\text{conv } E \geq e$, so that $\text{conv } E$ is proper. As $(\text{conv } E)(d) \leq E(d)$ for every $d \in \mathbf{R}^n$, it is clear that $\text{dom}(E) \subset \text{dom}(\text{conv } E)$ and hence $\text{conv}(\text{dom}(E)) \subset \text{dom}(\text{conv } E)$. Now, if a point $d \in \mathbf{R}^n$ cannot be expressed as a convex combination of points of $\text{dom}(E)$, then $\sum_{i=1}^q \lambda_i E(d_i) = \infty$ for every convex combination $d = \sum_{i=1}^q \lambda_i d_i$; hence $(\text{conv } E)(d) = \infty$ from (2.5). This shows that $\text{dom}(\text{conv } E) \subset \text{conv}(\text{dom}(E))$.

The proof of the remaining part of the theorem is more involved. For $d \in \text{dom}(\text{conv } E)$, consider a sequence of convex combinations

$$d = \sum_{i=1}^{q_k} \lambda_i^{(k)} d_i^{(k)}$$

with $1 \leq q_k \leq n+1$, $\lambda_i^{(k)} > 0$, $1 \leq i \leq q_k$, such that

$$(\text{conv } E)(d) = \lim_{k \rightarrow \infty} \sum_{i=1}^{q_k} \lambda_i^{(k)} E(d_i^{(k)}).$$

Existence of such a sequence follows from the comments preceding the theorem. Since $(\text{conv } E)(d) \in \mathbf{R}$, it is not restrictive to assume that $d_i^{(k)} \in \text{dom}(E)$ for $1 \leq i \leq q_k$. Moreover, extracting a subsequence, we may assume that $q_k = \bar{q}$ is constant (as a result of $1 \leq q_k \leq n+1$). Hence, the relations now read

$$(2.7) \quad d = \sum_{i=1}^{\bar{q}} \lambda_i^{(k)} d_i^{(k)}$$

and

$$(2.8) \quad (\text{conv } E)(d) = \lim_{k \rightarrow \infty} \sum_{i=1}^{\bar{q}} \lambda_i^{(k)} E(d_i^{(k)}).$$

For every fixed index $1 \leq i \leq \bar{q}$, the sequence $(d_i^{(k)})$ has accumulation points or does not have any. When it has some, they may belong to $\text{dom}(E)$ or not. After extracting subsequences and rearranging indices, if necessary, we may assume

$$(2.9) \quad \lim_{k \rightarrow \infty} d_i^{(k)} = d_i \in \text{dom}(E), \quad 1 \leq i \leq q,$$

$$(2.10) \quad \lim_{k \rightarrow \infty} d_i^{(k)} = d_i \notin \text{dom}(E), \quad q+1 \leq i \leq q',$$

$$(2.11) \quad \lim_{k \rightarrow \infty} |d_i^{(k)}| = \infty, \quad q'+1 \leq i \leq \bar{q},$$

where $0 \leq q \leq q' \leq \bar{q} (\leq n+1)$. Naturally, the values of q and q' may depend on the choice of the subsequences. Finally, extracting subsequences again (an action that does not affect (2.7)–(2.11)), we may also assume

$$(2.12) \quad \lim_{k \rightarrow \infty} \lambda_i^{(k)} = \lambda_i \geq 0, \quad 1 \leq i \leq \bar{q}.$$

Note that

$$(2.13) \quad E(d_i) \leq \liminf_{k \rightarrow \infty} E(d_i^{(k)}), \quad 1 \leq i \leq q,$$

and

$$(2.14) \quad \lim_{k \rightarrow \infty} E(d_i^{(k)}) = \infty, \quad q+1 \leq i \leq \bar{q}.$$

Indeed, (2.13) and (2.14) for $q+1 \leq i \leq q'$ follow from lower semicontinuity of E whereas (2.14) for $q'+1 \leq i \leq \bar{q}$ is immediate from (2.11) and the growth condition (2.3).

Now, it is clear from (2.14) that $E(d_i^{(k)}) > 0$ for $q+1 \leq i \leq \bar{q}$ and k large enough. With this observation and for every $\varepsilon > 0$ fixed, (2.8) thus yields the basic string of inequalities (recall $(\text{conv } E)(d) \in \mathbf{R}$)

$$(2.15) \quad 0 \leq \sum_{i=q'+1}^{\bar{q}} \lambda_i^{(k)} E(d_i^{(k)}) \leq \sum_{i=q+1}^{\bar{q}} \lambda_i^{(k)} E(d_i^{(k)}) \leq (\text{conv } E)(d) - \sum_{i=1}^q \lambda_i^{(k)} E(d_i^{(k)}) + \varepsilon,$$

for k large enough. With no loss of generality, we can then assume that (2.15) holds for every index k .

With (2.6) and (2.15), we find

$$(2.16) \quad 0 \leq \sum_{i=q'+1}^{\bar{q}} \lambda_i^{(k)} E(d_i^{(k)}) \leq \sum_{i=q+1}^{\bar{q}} \lambda_i^{(k)} E(d_i^{(k)}) \leq (\text{conv } E)(d) + |e| + \varepsilon < \infty.$$

From (2.12), (2.14), and (2.16), it is clear that

$$(2.17) \quad \lambda_i = 0, \quad q+1 \leq i \leq \bar{q}.$$

In particular, this shows that $q \geq 1$, for the relations

$$\sum_{i=1}^{\bar{q}} \lambda_i = \lim_{k \rightarrow \infty} \sum_{i=1}^{\bar{q}} \lambda_i^{(k)} = 1,$$

could not hold if $q = 0$. Hence it follows with (2.17) that

$$(2.18) \quad \sum_{i=1}^q \lambda_i = 1.$$

On the other hand, (2.11) and assumption (2.3) ensure that for every integer $m > 0$

$$E(d_i^{(k)}) > m|d_i^{(k)}|, \quad q' + 1 \leq i \leq \bar{q},$$

provided that k is large enough. Therefore, from (2.16) we find

$$\sum_{i=q'+1}^{\bar{q}} \lambda_i^{(k)} |d_i^{(k)}| \leq \frac{1}{m} [(\text{conv } E)(d) + |e| + \varepsilon],$$

for k large enough. As $m > 0$ is arbitrary, this yields

$$\lim_{k \rightarrow \infty} \sum_{i=q'+1}^{\bar{q}} \lambda_i^{(k)} d_i^{(k)} = 0.$$

But $\lim_{k \rightarrow \infty} \sum_{i=q+1}^{q'} \lambda_i^{(k)} d_i^{(k)} = 0$ is obvious from (2.10) and (2.17) and hence, together with (2.7) and (2.9) we get

$$d = \lim_{k \rightarrow \infty} \sum_{i=1}^{\bar{q}} \lambda_i^{(k)} d_i^{(k)} = \lim_{k \rightarrow \infty} \sum_{i=1}^q \lambda_i^{(k)} d_i^{(k)} = \sum_{i=1}^q \lambda_i d_i.$$

Recalling (2.18), we see that d is a convex combination of d_1, \dots, d_q .

A last and obvious consequence of (2.15) is the inequality

$$\sum_{i=1}^q \lambda_i^{(k)} E(d_i^{(k)}) \leq (\text{conv } E)(d) + \varepsilon.$$

From (2.13) we infer that

$$\sum_{i=1}^q \lambda_i E(d_i) \leq (\text{conv } E)(d) + \varepsilon.$$

Hence

$$\sum_{i=1}^q \lambda_i E(d_i) \leq (\text{conv } E)(d),$$

since $\varepsilon > 0$ is arbitrary. But it was previously shown that d is a convex combination of d_1, \dots, d_q . Therefore

$$(\text{conv } E)(d) \leq \sum_{i=1}^q \lambda_i E(d_i),$$

from the characterization (2.5) of $(\text{conv } E)(d)$. As a result

$$(\text{conv } E)(d) = \sum_{i=1}^q \lambda_i E(d_i).$$

In this formula, the λ_i 's need not be strictly positive (but the d_i 's are in $\text{dom}(E)$ from (2.9)), and the d_i 's need not be distinct. Obviously, both conditions can be met by decreasing q accordingly, and the proof is complete. \square

Remark 2.1. For $d \in \text{dom}(\text{conv } E) = \text{conv}(\text{dom}(E))$, let $d = \sum_{i=1}^q \lambda_i d_i$ be any convex combination with $\lambda_i > 0$ and $d_i \in \text{dom}(E)$, satisfying

$$(\text{conv } E)(d) = \sum_{i=1}^q \lambda_i E(d_i)$$

(in particular, it is not necessary to assume $q \leq n + 1$). Then, $(\text{conv } E)(d_i) = E(d_i)$. Indeed, if $(\text{conv } E)(d_i) < E(d_i)$ for at least one index i , then convexity of $\text{conv } E$ and the inequality $\text{conv } E \leq E$ together imply

$$(\text{conv } E)(d) \leq \sum_{i=1}^q \lambda_i (\text{conv } E)(d_i) < \sum_{i=1}^q \lambda_i E(d_i),$$

a contradiction.

Also, $\text{conv } E$ is affine on $\text{conv}\{d_1, \dots, d_q\}$. Indeed, let $\delta = \sum_{i=1}^q \mu_i d_i$, $\mu_i \geq 0$, $1 \leq i \leq q$, $\sum_{i=1}^q \mu_i = 1$, be any point of $\text{conv}\{d_1, \dots, d_q\}$. From the last inequality and by convexity

$$(\text{conv } E)(\delta) \leq \sum_{i=1}^q \mu_i E(d_i) \equiv \phi(\delta).$$

Obviously, ϕ is affine, and the convex function $(\text{conv } E) - \phi$ achieves its maximum value at the point d of the relative interior of $\text{conv}\{d_1, \dots, d_q\}$, hence is constant on $\text{conv}\{d_1, \dots, d_q\}$. Moreover, $(\text{conv } E)(d) - \phi(d) = 0$, so that $\text{conv } E = \phi$ is affine on $\text{conv}\{d_1, \dots, d_q\}$. \square

Remark 2.2. The growth condition (2.3) cannot be replaced by the weaker one $\lim_{|\delta| \rightarrow \infty} E(\delta) = \infty$ because it is no longer possible to ascertain (in the notation of the proof of Theorem 2.1) that d is a convex combination of d_1, \dots, d_q . \square

3. CONTINUOUS DIFFERENTIABILITY OF CONVEX ENVELOPES

In the remainder of this paper, we shall set

$$\mathbf{D} = \text{int}(\text{dom}(E)).$$

Before we examine smoothness properties, we need the following theorem.

Theorem 3.1. *Retain the assumptions (2.1) to (2.3). Then, for*

$$d \in \text{int}(\text{dom}(\text{conv } E))$$

the subdifferential $\partial(\text{conv } E)(d)$ is nonempty. Moreover, let $d = \sum_{i=1}^q \lambda_i d_i$ be any convex combination with $\lambda_i > 0$ and $d_i \in \text{dom}(E)$ such that

$$(3.1) \quad (\text{conv } E)(d) = \sum_{i=1}^q \lambda_i E(d_i)$$

(the existence of which, with $q \leq n + 1$, follows from Theorem 2.1). Then, for every $g \in \partial(\text{conv } E)(d)$, one has

$$(3.2) \quad E(d_i) - g \cdot d_i = (\text{conv } E)(d) - g \cdot d, \quad 1 \leq i \leq q,$$

and

$$(3.3) \quad E(\delta) - g \cdot \delta \geq E(d_i) - g \cdot d_i, \quad \forall \delta \in \mathbf{R}^n, \quad 1 \leq i \leq q.$$

In addition, suppose that E is differentiable¹ at each point d_i , $1 \leq i \leq q$, and that $\partial(\text{dom}(E)) \cap \text{dom}(E)$ coincides with a differentiable manifold M_i with dimension $n - 1$ near each point $d_i \in \partial(\text{dom}(E))$. Then, for every $g \in \partial(\text{conv } E)(d)$ and each i , one has

$$(3.4) \quad g = \nabla E(d_i) \quad \text{if } d_i \in \mathbf{D},$$

or

$$(3.5) \quad g = \nabla E(d_i) + t_i \nu(d_i) \quad \text{if } d_i \in \partial(\text{dom}(E)) \cap \text{dom}(E),$$

where $\nu(d_i)$ is a nonzero vector normal to M_i at d_i and $t_i \in \mathbf{R}$.

Proof. Nonemptiness of $\partial(\text{conv } E)(d)$ for $d \in \text{int}(\text{dom}(\text{conv } E))$ follows from [6, Theorem 23.4, p. 217]. To prove (3.2), note that since $d_i \in \text{dom}(E)$

$$(3.6) \quad E(d_i) \geq (\text{conv } E)(d_i) \geq (\text{conv } E)(d) + g \cdot (d_i - d), \quad 1 \leq i \leq q,$$

for every $g \in \partial(\text{conv } E)(d)$. But suppose that either of the above inequalities is strict for some index i_0 . Then, multiplying (3.6) by $\lambda_{i_0} > 0$ and adding up the resulting q inequalities, we find (recall $\sum_{i=1}^q \lambda_i = 1$ and $\sum_{i=1}^q \lambda_i d_i = d$)

$$\sum_{i=1}^q \lambda_i E(d_i) > (\text{conv } E)(d),$$

in contradiction with (3.1). Hence, the inequalities in (3.6) are equalities for $1 \leq i \leq q$, which proves (3.2). Relation (3.3) now follows from (3.2) and

$$E(\delta) \geq (\text{conv } E)(\delta) \geq (\text{conv } E)(d) + g \cdot (\delta - d).$$

From (3.2) and (3.3) together, we infer that the mapping

$$(3.7) \quad \delta \in \mathbf{R}^n \rightarrow E(\delta) - g \cdot \delta \in (-\infty, \infty]$$

achieves its minimum value at d_1, \dots, d_q . The assumption that $d_i \in \mathbf{D}$ and E is differentiable at d_i implies relation (3.4). As $d_i \in \text{dom}(E)$, it is clear that $d_i \in \partial(\text{dom}(E)) \cap \text{dom}(E)$ whenever $d_i \notin \mathbf{D}$. If E is differentiable at d_i and $\partial(\text{dom}(E)) \cap \text{dom}(E)$ contains a differentiable manifold M_i with dimension $n - 1$ near d_i , equation (3.5) follows from the fact that the mapping (3.7) achieves its minimum value on M_i at d_i . \square

Remark 3.1. With the restrictions $\text{dom}(E)$ compact, $q \leq n + 1$, and $d \in \mathbf{D}$, Theorem 3.1 can also be obtained as a corollary to a general result in

¹For $d_i \in \partial(\text{dom}(E))$, this means that on $\text{dom}(E)$, E is the restriction of a real-valued function on \mathbf{R}^n which is differentiable at d_i .

semi-infinite linear programming (cf. Glashoff and Gustafson [3, Theorem 1, p. 135]). \square

The subset of $\text{dom}(E)$ of those points at which the subdifferential of E is not empty will play an important role. Accordingly, we shall set

$$(3.8) \quad \text{dom}(E)^* = \{\delta \in \text{dom}(E); \partial E(\delta) \neq \emptyset\}.$$

In the assumption that $\partial(\text{dom}(E)) \cap \text{dom}(E)$ coincides with an $(n-1)$ -dimensional manifold near each point of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$, the normal line to $\partial(\text{dom}(E)) \cap \text{dom}(E)$ at $\delta \in \partial(\text{dom}(E)) \cap \text{dom}(E)^*$ is defined without ambiguity. In this case, we shall denote by $\nu(\delta)$ either of the two unit vectors generating the normal line.

For every integer $q \geq 1$, set

$$(3.9) \quad S_q(E) = \{d \in \text{conv}\{d_1, \dots, d_q\}; d_i \in \partial(\text{dom}(E)) \cap \text{dom}(E)^*, \\ \nu(d_i) = \pm \nu(d_j), \nu(d_i) \cdot (d_i - d_j) = 0, 1 \leq i, j \leq q\}$$

and

$$(3.10) \quad S(E) = \bigcup_{q \geq 1} S_q(E).$$

In what follows, it will sometimes be convenient to refer to $S(E)$ as the singular set of E .

The main result regarding (continuous) differentiability of $\text{conv } E$ is as follows.

Theorem 3.2. *In addition to the assumptions (2.1)–(2.3), suppose that $\partial(\text{dom}(E)) \cap \text{dom}(E)$ coincides with a $(n-1)$ -dimensional manifold near each point of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$. Suppose also that E is differentiable at each point of $\text{dom}(E)^*$ (in particular, this is true if E is differentiable at every point of $\text{dom}(E)$). Then, $\text{conv } E$ is differentiable at every point of the interior of its domain, except possibly at points of the singular set $S(E)$ in (3.10). In particular, $\text{conv } E$ is continuously differentiable on any open subset of its domain not intersecting $S(E)$.*

Note. If $n = 1$, $S(E) = \partial(\text{dom}(E)) \cap \text{dom}(E)$.

Proof. Let $d \in \text{int}(\text{dom}(\text{conv } E))$, and let $d = \sum_{i=1}^q \lambda_i d_i$ be a convex combination satisfying the assumptions of Theorem 3.1. With no loss of generality, assume $d_i \neq d_j$ for $i \neq j$.

To prove that $\text{conv } E$ is differentiable at d , we need only show $\partial(\text{conv } E)(d)$ contains exactly one element. As $\partial(\text{conv } E)(d) \neq \emptyset$ is guaranteed by Theorem 3.1, let $g \in \partial(\text{conv } E)(d)$. From relation (3.3) it is clear that $g \in \partial E(d_i)$; hence $d_i \in \text{dom}(E)^*$, and E is therefore differentiable at d_i , $1 \leq i \leq q$.

If one of the d_i 's is in \mathbf{D} , relation (3.4) yields $g = \nabla E(d_i)$, which specifies g in a unique way. Thus, it remains only to examine the case when all the d_i 's lie on $\partial(\text{dom}(E))$. Assuming that $d \notin S(E)$, one has $d \notin S_1(E)$ and hence $d \notin \partial(\text{dom}(E)) \cap \text{dom}(E)^*$ which, in turn, implies $q \geq 2$.

Using previous notation, suppose first that $\nu(d_i)$ and $\nu(d_j)$ are not collinear for some pair of indices $1 \leq i, j \leq q$. Obviously, one may assume $i = 1$ and $j = 2$ and there is a vector τ_2 tangent to $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$ at d_2 which is not tangent to $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$ at d_1 , i.e., $\nu(d_1) \cdot \tau_2 \neq 0$. From (3.5), one has

$$g = \nabla E(d_1) + t_1 \nu(d_1) = \nabla E(d_2) + t_2 \nu(d_2),$$

for some real numbers t_1 and t_2 . But then

$$\nabla E(d_1) \cdot \tau_2 + t_1 \nu(d_1) \cdot \tau_2 = \nabla E(d_2) \cdot \tau_2,$$

yielding

$$t_1 = (\nabla E(d_2) - \nabla E(d_1)) \cdot \tau_2 / \nu(d_1) \cdot \tau_2,$$

which, again, specifies g in a unique way.

Finally, suppose that $\nu(d_i) = \pm \nu(d_j)$ for every $1 \leq i, j \leq q$. Since $d \notin S(E)$, there is a pair of indices such that $\nu(d_i) \cdot (d_i - d_j) \neq 0$. Again, with no loss of generality, assume $i = 1$ and $j = 2$. From relation (3.2), we get

$$E(d_2) = E(d_1) + g \cdot (d_2 - d_1).$$

But $g = \nabla E(d_1) + t_1 \nu(d_1)$ (cf. (3.5)) and hence

$$E(d_2) = E(d_1) + \nabla E(d_1) \cdot (d_2 - d_1) + t_1 \nu(d_1) \cdot (d_2 - d_1),$$

namely

$$t_1 = [E(d_2) - E(d_1) - \nabla E(d_1) \cdot (d_2 - d_1)] / \nu(d_1) \cdot (d_2 - d_1),$$

and g is thus specified in a unique way.

Above, we have seen that $\text{conv } E$ is differentiable at every point of

$$\text{int}(\text{dom}(\text{conv } E)) - S(E).$$

That $\text{conv } E$ is necessarily *continuously* differentiable on every open subset of its domain not intersecting $S(E)$ follows from [6, Corollary 25.5.1, p. 246]. \square

Remark 3.2. The assumption that E is differentiable at every point of $\text{dom}(E)^*$ can be weakened by requiring only that for $\delta \in \text{dom}(E)^*$, the following alternative holds:

- (i) either $\delta \in \mathbf{D} = \text{int}(\text{dom}(E))$ and $\partial E(\delta)$ contains exactly one element,
- (ii) or $\delta \in \partial(\text{dom}(E))$ and $\partial E(\delta)$ contains exactly one element modulo the normal line to $\partial(\text{dom}(E))$ through δ .

Indeed, in this case, it is easy to obtain a modified version of Theorem 3.1 allowing for a proof of Theorem 3.2 similar to the one given above. An example of when this generalization is relevant will be given later. \square

Many corollaries to Theorem 3.2 may be derived that ensure that $\text{conv } E$ is continuously differentiable on the interior of its domain. For instance, this applies if $\partial(\text{dom}(E)) \cap \text{dom}(E) = \emptyset$ so that $\text{dom}(E)$ is *open* (or, more generally,

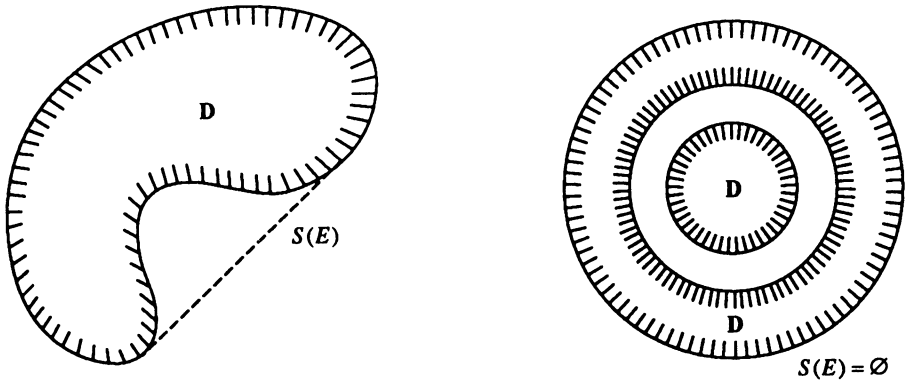


FIGURE 3.1. Two examples when $\mathbf{D} \cap S(E) = \emptyset$ irrespective of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$

if $\partial(\text{dom}(E)) \cap \text{dom}(E)^* = \emptyset$. If so, Theorem 3.2 puts no restriction on the geometry of $\partial(\text{dom}(E))$, and the assertion follows from $S(E) = \emptyset$. Note in particular that $\text{dom}(E)$ needs not be connected (but assumptions (2.1)–(2.3) must, of course, be satisfied). Also, $\text{dom}(E)$ is trivially open when $\text{dom}(E) = \mathbf{R}^n$.

Other corollaries to Theorem 3.2 will guarantee that $\text{conv } E$ is continuously differentiable on certain open subsets of the interior of its domain. For instance, continuous differentiability of $\text{conv } E$ on $\mathbf{D} = \text{int}(\text{dom}(E))$ follows from the assumption $\mathbf{D} \cap S(E) = \emptyset$ which, also, can be realized irrespective of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$ (except for smoothness of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$) without assuming that $\text{dom}(E)$ is convex or connected, as shown on Figure 3.1.

While these observations emphasize that convexity of $\text{dom}(E)$ is not necessary for $\text{conv } E$ to be continuously differentiable, nevertheless it is an important sufficient condition.

Corollary 3.1. *In addition to the assumptions of Theorem 3.1, suppose that $\text{dom}(E)$ is convex. Then, $\text{conv } E$ is continuously differentiable on*

$$\mathbf{D} = \text{int}(\text{dom}(E)) = \text{int}(\text{dom}(\text{conv } E)).$$

Proof. The relation $\text{int}(\text{dom}(\text{conv } E)) = \text{int}(\text{dom}(E))$ follows from $\text{dom}(E)$ convex and Theorem 2.1. Thus, to prove the claim, it suffices to show that $\mathbf{D} \cap S(E) = \emptyset$. We henceforth assume $\mathbf{D} \neq \emptyset$ with no loss of generality.

Let then $q \geq 1$ be a given integer, and pick $d \in S_q(E)$. For some unit vector $\nu \in \mathbf{R}^n$, one then has $\nu(d_j) = \pm \nu$, $1 \leq j \leq q$, and $(d_i - d_j) \cdot \nu = 0$ for $1 \leq i, j \leq q$. Hence

$$(d - d_i) \cdot \nu = (d - d_j) \cdot \nu, \quad 1 \leq i, j \leq q.$$

Writing $d = \sum_{i=1}^q \lambda_i d_i$, $\sum_{i=1}^q \lambda_i = 1$, $\lambda_i \geq 0$, $1 \leq i \leq q$, multiplying the above relation by λ_i , and next summing over $1 \leq i \leq q$ yield

$$(d - d_j) \cdot \nu = 0, \quad 1 \leq j \leq q.$$

In other words, $(d - d_j)$ lies on the tangent space of $\partial(\text{dom}(E)) \cap \text{dom}(E)^*$ at d_j , which is also the tangent space to $\partial(\text{dom}(E))$ at d_j from convexity of $\text{dom}(E)$. Recall that the boundary of a convex set with nonempty interior is locally the graph of a Lipschitz continuous function. Here, near d_j , this graph contains the graph of a C^1 function by hypothesis and hence coincides with it.

As a result, $d = d_j + (d - d_j)$ lies in the (unique) hyperplane supporting $\text{dom}(E)$ at d_j , which does not intersect $\mathbf{D} = \text{int}(\text{dom}(E))$ and hence $d \notin \mathbf{D}$. \square

Remark 3.3. Corollary 3.1 can be extended, with the same proof, to the case when $\text{dom}(E)$ convex is replaced by $\partial(\text{dom}(E)) \subset \partial\mathbf{D}$ with \mathbf{D} open and convex. If so, the conclusion is that $\text{conv } E$ is continuously differentiable on \mathbf{D} . This version allows for $\text{dom}(E) \subset \partial(\text{dom}(E))$; i.e., E is finite only on (part of) the boundary $\partial\mathbf{D}$. \square

To complete this section, we shall now give an example when the generalization of Theorem 3.2 mentioned in Remark 3.2 is relevant. First, consider E as in Theorem 3.2, and suppose also that $\partial(\text{dom}(E)) \cap \text{dom}(E)^* = \emptyset$. For this, it suffices that the graph of E have a “vertical tangent” at every point of $\partial(\text{dom}(E)) \cap \text{dom}(E)$, which can be realized even when $\text{dom}(E)$ is closed. Since $S(E) = \emptyset$ in this case, $\text{conv } E$ is continuously differentiable on the interior of its domain.

Now, consider E_1 and E_2 as above, and set

$$E = \min(E_1, E_2).$$

Then, $\text{dom}(E) = \text{dom}(E_1) \cup \text{dom}(E_2) \neq \emptyset$. Also, E is lower semicontinuous on \mathbf{R}^n and verifies the growth condition (2.3). Unlike differentiability properties at points of $\text{dom}(E)^*$ which, even if true, are hard to check, the conditions (i) and (ii) in Remark 3.2 are easy to verify. Indeed, let $d \in \text{dom}(E)^*$, so that there is $g \in \mathbf{R}^n$ with

$$E(\delta) \geq E(d) + g \cdot (\delta - d), \quad \forall \delta \in \mathbf{R}^n.$$

As $E(d) = E_\alpha(d)$ for $\alpha = 1$ or 2 , say $E(d) = E_1(d)$, one has

$$E_1(\delta) \geq E(\delta) \geq E_1(d) + g \cdot (\delta - d), \quad \forall \delta \in \mathbf{R}^n.$$

In other words, $d \in \text{dom}(E_1)^*$. But $\partial(\text{dom}(E_1)) \cap \text{dom}(E_1)^* = \emptyset$, so that $d \in \text{int}(\text{dom}(E_1)) \subset \text{int}(\text{dom}(E))$. Since E_1 is differentiable on the interior of its domain, $g = \nabla E(d_1)$ is uniquely determined. It follows that $\partial(\text{dom}(E)) \cap \text{dom}(E)^* = \emptyset$ and that $\partial E(\delta)$ contains exactly one element whenever $\partial E(\delta) \neq \emptyset$. The hypotheses of Theorem 3.2 are satisfied with $S(E) = \emptyset$, so that $\text{conv } E$ is continuously differentiable on the interior of its domain.

4. HÖLDER CONTINUITY OF THE GRADIENT

Since only $C^{1,1}$ regularity of $\text{conv } E$ may be expected no matter how smooth E is, it is quite natural to investigate whether general conditions can be found

that do guarantee that $\text{conv } E$ is (locally at least) of class $C^{1,1}$, or more generally of class $C^{1,\alpha}$, $0 < \alpha \leq 1$.

Existence of such conditions will be established in this section. We have found technical obstacles that have prevented us from being able to investigate these matters with as much generality as the continuous differentiability property. For both brevity and clarity, and also because the current applications do not seem to require a more general framework, we have limited our exposition to the case when

$$(4.1) \quad \partial(\text{dom}(E)) \cap \text{dom}(E)^* = \emptyset,$$

or, equivalently,

$$(4.1') \quad \text{dom}(E)^* \subset \mathbf{D}.$$

Of course, we shall retain the hypotheses (2.1) to (2.3) and require some smoothness of E . Precisely, we shall assume

$$(4.2) \quad E \text{ is locally of class } C^{1,\alpha}, \quad 0 < \alpha \leq 1, \text{ on some open neighborhood of } \text{dom}(E)^*,$$

a condition that makes sense in view of (4.1').

As a result of these hypotheses, Theorem 3.2 yields C^1 regularity of $\text{conv } E$. Moreover, since $S(E) = \emptyset$ from (4.1), one has

$$(4.3) \quad \text{conv } E \in C^1(\text{conv } \mathbf{D}).$$

To avoid endless repetition, we shall once and for all decide that $d \in \text{conv } \mathbf{D}$ is written in the form

$$(4.4) \quad d = \sum_{i=1}^q \lambda_i d_i,$$

with $\lambda_i > 0$, $\sum_{i=1}^q \lambda_i = 1$, $1 \leq q \leq n+1$, $d_i \in \text{dom}(E)$, $d_i \neq d_j$ for $i \neq j$ and

$$(4.5) \quad (\text{conv } E)(d) = \sum_{i=1}^q \lambda_i E(d_i).$$

Existence of such a convex combination follows from Theorem 2.1. But since its uniqueness is not true in general, the λ_i 's and d_i 's will henceforth refer to *any* such combination.

As observed in the proof of Theorem 3.2, one has

$$(4.6) \quad d_i \in \text{dom}(E)^*, \quad 1 \leq i \leq q.$$

Thus, from (4.1'), $d_i \in \mathbf{D}$ and hence

$$(4.7) \quad \nabla(\text{conv } E)(d) = \nabla E(d_i), \quad 1 \leq i \leq q,$$

from Theorem 3.1.

Remark 4.1. Note that (4.1) holds if $\text{dom } E = \mathbf{R}^n$ and (4.2) is fulfilled as soon as E is locally $C^{1,\alpha}$ on \mathbf{R}^n , an important case in practice. These hypotheses are also satisfied by many thermodynamic potentials E with $\text{dom}(E) \neq \mathbf{R}^n$. \square

Our aim is to prove that $\text{conv } E$ is locally $C^{1,\alpha}$ on $\text{conv } \mathbf{D}$ (global properties will be examined, too) namely that $\nabla(\text{conv } E)$ is locally $C^{0,\alpha}$ on $\text{conv } \mathbf{D}$. This result relies on a sequence of preliminary lemmas.

Lemma 4.1. *Let M and G be arbitrary constants. Let*

$$\Gamma = \Gamma(M, G) = \{\delta \in \text{dom}(E)^* ; E(\delta) \leq M, |\nabla E(\delta)| \leq G\}.$$

Then, Γ is a compact subset of $\text{dom}(E)^$.*

Proof. With no loss of generality, assume $\Gamma \neq \emptyset$. Since Γ is bounded because of the growth condition (2.3), it suffices to prove that Γ is closed in \mathbf{R}^n .

Let then (δ_k) be a sequence of points in Γ and $\delta \in \mathbf{R}^n$ such that $\lim_{k \rightarrow \infty} \delta_k = \delta$. From E being lower semicontinuous in \mathbf{R}^n

$$(4.8) \quad E(\delta) \leq \liminf_{k \rightarrow \infty} E(\delta_k) \leq M.$$

Next, from (4.2), it is clear that $\partial E(\delta_k) = \{\nabla E(\delta_k)\}$. Therefore,

$$(4.9) \quad E(\xi) \geq E(\delta_k) + \nabla E(\delta_k) \cdot (\xi - \delta_k), \quad \forall \xi \in \mathbf{R}^n, \forall k.$$

Because $|\nabla E(\delta_k)| \leq G$ by hypothesis, and after extracting a subsequence, one may assume that $\lim_{k \rightarrow \infty} \nabla E(\delta_k) = g$. Obviously,

$$(4.10) \quad |g| \leq G.$$

Using again lower semicontinuity of E , it follows from (4.9) that

$$E(\xi) \geq E(\delta) + g \cdot (\xi - \delta), \quad \forall \xi \in \mathbf{R}^n.$$

This shows that $\delta \in \text{dom}(E)^*$ and that $g \in \partial E(\delta)$. But, from (4.2) again, $\partial E(\delta) = \{\nabla E(\delta)\}$ so that $g = \nabla E(\delta)$ and (4.10) reads $|\nabla E(\delta)| \leq G$. Together with (4.8) and $\delta \in \text{dom}(E)^*$, this shows that $\delta \in \Gamma$. \square

Remark 4.2. Clearly, the proof goes through with $E \in C^1(\text{dom}(E)^*)$ instead of (4.2). Also, note that the set $\{\delta \in \text{dom}(E)^*, E(\delta) \leq M\}$ is not compact in general, so that involving ∇E in the definition of Γ is essential. \square

The next lemma exhibits a property of a different nature which plays a fundamental role in obtaining some uniform estimates.

Lemma 4.2. *Let $\kappa \subset \text{conv } \mathbf{D}$ be compact. Then, there are constants M and G such that (in the notation (4.4)–(4.5)) for every $d \in \kappa$, there is $1 \leq i_0 \leq q$ such that $\lambda_{i_0} \geq 1/(n+1)$ and $d_{i_0} \in \Gamma(M, G)$, with $\Gamma(M, G)$ as in Lemma 4.1.*

Proof. Since $\text{conv } E \in C^1(\text{conv } \mathbf{D})$ (see (4.3)), there is a constant G such that

$$|\nabla(\text{conv } E)(d)| \leq G, \quad \forall d \in \kappa.$$

Thus, with (4.6) and (4.7), one has $d_i \in \text{dom}(E)^*$ and $|\nabla E(d_i)| \leq G$ for $1 \leq i \leq q$, and it remains to find a constant M such that $\lambda_{i_0} \geq 1/(n+1)$ and $E(d_{i_0}) \leq M$ for some index $1 \leq i_0 \leq q$. Of course, the index i_0 may depend on d , but not the constant M .

Since $\text{conv } E$ is continuous on $\text{conv } \mathbf{D}$, there is a constant $M_0 \geq 0$ such that

$$(4.12) \quad (\text{conv } E)(d) \leq M_0, \quad \forall d \in \kappa.$$

On the other hand, recall that E is bounded from below by a constant e as a result of (2.2) and (2.3). Then, one may take

$$(4.13) \quad M = (n+1)(M_0 + |e|).$$

Indeed, since $q \leq n+1$, there is an index $1 \leq i_0 \leq q$ such that $\lambda_{i_0} \geq 1/(n+1)$ since the λ_i 's sum to one. For convenience of notation, suppose $i_0 = 1$. Now, rewrite (4.5) in the form

$$\lambda_1 E(d_1) = \text{conv } E(d) - \sum_{i=2}^q \lambda_i E(d_i) \leq M_0 - \sum_{i=2}^q \lambda_i E(d_i).$$

For each index $2 \leq i \leq q$ such that $E(d_i) \geq 0$, zero is a majorant of $-\lambda_i E(d_i)$. If $E(d_i) < 0$, then $e \leq E(d_i) < 0$ and $-\lambda_i E(d_i) \leq -\lambda_i e = \lambda_i |e|$. Thus, in any case, one has $-\lambda_i E(d_i) \leq \lambda_i |e|$ and hence

$$\lambda_1 E(d_1) \leq M_0 + \sum_{i=2}^q \lambda_i |e| \leq M_0 + |e|.$$

As $\lambda_1 \geq 1/(n+1)$ by hypothesis, it follows that $E(d_1) \leq M$ with M as in (4.13). \square

The third lemma contains nothing of a new nature. We give its proof for completeness.

Lemma 4.3. *Let $L \subset \text{dom}(E)^*$ be compact. Then, there are constants $C = C(L) \geq 0$ and $\varepsilon > 0$ such that, for every $\xi_0 \in L$, the open ball $B(\xi_0, \varepsilon)$ with center ξ_0 and radius ε is contained in \mathbf{D} , ∇E is defined on $B(\xi_0, \varepsilon)$, and*

$$(4.14) \quad E(\xi') \leq E(\xi) + \nabla E(\xi) \cdot (\xi' - \xi) + C|\xi' - \xi|^{1+\alpha},$$

for every $\xi, \xi' \in B(\xi_0, \varepsilon)$.

Proof. Given $\xi_0 \in \text{dom}(E)^*$, it follows from (4.1') and (4.2) that there is $\rho_0 > 0$ such that the open ball $B(\xi_0, \rho_0)$ with center ξ_0 and radius ρ_0 is contained in \mathbf{D} , ∇E is defined on $B(\xi_0, \rho_0)$ and

$$|\nabla E(\xi') - \nabla E(\xi)| \leq C(\xi_0)|\xi' - \xi|^\alpha, \quad \forall \xi, \xi' \in B(\xi_0, \rho_0),$$

where $C(\xi_0) \geq 0$ is a constant.

As L is compact, one may cover L by finitely many such balls $B(\xi_l, \rho_l)$. Call $C \geq 0$ the largest corresponding constant $C(\xi_l)$. It is a standard result

that $\varepsilon > 0$ can be found such that for every $\xi_0 \in L$, $B(\xi_0, \varepsilon) \subset B(\xi_l, \rho_l)$ for some l . With such a choice, it is obvious that

$$|\nabla E(\xi') - \nabla E(\xi)| \leq C|\xi' - \xi|^\alpha, \quad \forall \xi, \xi' \in B(\xi_0, \varepsilon), \quad \forall \xi_0 \in L.$$

That inequality (4.14) holds, with the same constant C as above, follows from a straightforward application of the mean value theorem. \square

We are now ready to establish our fourth and last preliminary lemma, whose proof relies on the results proved in the first three.

Lemma 4.4. *Let $\kappa \subset \text{conv } \mathbf{D}$ be compact. Then, there is a constant $C = C(\kappa) \geq 0$ and $\eta > 0$ such that, for every $d \in \kappa$, the open ball $B(d, \eta)$ with center d and radius η is contained in $\text{conv } \mathbf{D}$ and*

$$(\text{conv } E)(\delta) \leq (\text{conv } E)(d) + \nabla(\text{conv } E)(d) \cdot (\delta - d) + C|\delta - d|^{1+\alpha},$$

for every $\delta \in B(d, \eta)$.

Proof. First, use Lemma 4.2 to obtain constants M and G such that, for every $d \in \kappa$ and in the notation (4.4)–(4.5), there is an index $1 \leq i_0 \leq q$ such that $\lambda_{i_0} \geq 1/(n+1)$ and $d_{i_0} \in \Gamma(M, G)$ with $\Gamma(M, G)$ as in Lemma 4.1, hence compact. In what follows, we set $\Gamma(M, G) = L$ and, for notational convenience, we assume that the ordering of the d_i 's is consistent with the choice $i_0 = 1$ for every $d \in \kappa$, namely

$$(4.15) \quad d_1 \in L, \quad \lambda_1 \geq 1/(n+1).$$

Now, pick a ε as in Lemma 4.3. With $d \in \kappa$ being arbitrary, it follows from (4.15) that

$$|\delta - d| < \frac{\varepsilon}{n+1} \Rightarrow d_1 + \frac{\delta - d}{\lambda_1} \in B(d_1, \varepsilon) \subset \mathbf{D}.$$

Since κ is compact, one can choose $\eta > 0$ independent of d such that $\delta \in \text{conv } \mathbf{D}$ whenever $|\delta - d| < \eta$. In addition, taking $\eta < \varepsilon/(n+1)$, one finds

$$(4.16) \quad |\delta - d| < \eta \Rightarrow \delta \in \text{conv } \mathbf{D} \quad \text{and} \quad d_1 + \frac{\delta - d}{\lambda_1} \in B(d_1, \varepsilon) \subset \mathbf{D}.$$

For $\delta \in B(d, \eta)$ write, consistently with (4.4)

$$\delta = \lambda_1 \left(d_1 + \frac{\delta - d}{\lambda_1} \right) + \lambda_2 d_2 + \cdots + \lambda_q d_q,$$

so that

$$(4.17) \quad (\text{conv } E)(\delta) \leq \lambda_1 E \left(d_1 + \frac{\delta - d}{\lambda_1} \right) + \lambda_2 E(d_2) + \cdots + \lambda_q E(d_q).$$

Now, from (4.15) and (4.16), one may use inequality (4.14) of Lemma 4.3 with $\xi_0 = \xi = d_1$ and $\xi' = d_1 + (\delta - d)/\lambda_1$ to get (with $C = C(L)$)

$$E \left(d_1 + \frac{\delta - d}{\lambda_1} \right) \leq E(d_1) + \nabla E(d_1) \cdot \frac{\delta - d}{\lambda_1} + C \left| \frac{\delta - d}{\lambda_1} \right|^{1+\alpha}.$$

Upon substituting this into (4.17), one arrives at

$$(\operatorname{conv} E)(\delta) \leq \sum_{i=1}^q \lambda_i E(d_i) + \nabla E(d_1) \cdot (\delta - d) + \frac{C}{\lambda_1^{1+\alpha}} |\delta - d|^{1+\alpha}.$$

But $\sum_{i=1}^q \lambda_i E(d_i) = \operatorname{conv} E(d)$, $\nabla E(d_1) = \nabla(\operatorname{conv} E)(d)$ (see (4.7)), and $\lambda_1 \geq 1/(n+1)$. Thus, the assertion follows upon changing $C = C(L)$ above into $C(\kappa) = (n+1)^{1+\alpha} C(L)$. \square

After all these preliminaries, we are in a position to prove the main result of this section.

Theorem 4.2. *In addition to assumptions (2.1)–(2.3), suppose that conditions (4.1) and (4.2) hold. Then, $\operatorname{conv} E$ is locally of class $C^{1,\alpha}$ on $\operatorname{conv} \mathbf{D}$.*

Proof. Let $d_0 \in \operatorname{conv} \mathbf{D}$ be fixed and let $R > 0$ be such that the closed ball $\overline{B}(d_0, R)$ is contained in $\operatorname{conv} \mathbf{D}$. For short, set $\kappa = \overline{B}(d_0, R)$ and choose $\eta > 0$ and $C \geq 0$ as in Lemma 4.4. Thus, for every $d \in \kappa$ and every $\delta \in B(d, \eta)$, one has

$$(4.18) \quad (\operatorname{conv} E)(\delta) - (\operatorname{conv} E)(d) - \nabla(\operatorname{conv} E)(d) \cdot (\delta - d) \leq C |\delta - d|^{1+\alpha}.$$

Since this inequality holds for every $d \in \kappa$, it holds for every d in any subset of κ . In particular, it is true for $d \in \overline{B}(d_0, 3r)$ where $r < \min(R/3, \eta/6)$. With this choice, inequality (4.18) holds with arbitrary pairs $d, \delta \in \overline{B}(d_0, 3r)$ since $\operatorname{diam} \overline{B}(d_0, 3r) \leq 6r < \eta$.

We shall achieve our goal by combining inequality (4.18) with an argument already used by Stoer and Witzgall [7] for showing that a convex function differentiable at every point is automatically continuously differentiable.

Fix $d \in \overline{B}(d_0, r)$, and pick ξ with $|\xi - d| \leq r$. Denoting by S_ξ the sphere with center d and radius $2|\xi - d| \leq 2r$, one has $S_\xi \subset \overline{B}(d_0, 3r)$ and hence (4.18) holds with $\delta \in S_\xi$. On the other hand,

$$(\operatorname{conv} E)(\delta) \geq (\operatorname{conv} E)(\xi) + \nabla(\operatorname{conv} E)(\xi) \cdot (\delta - \xi).$$

Hence, subtracting $\nabla(\operatorname{conv} E)(d) \cdot (\delta - \xi)$ from both sides, one easily obtains

$$\begin{aligned} (\operatorname{conv} E)(\delta) - (\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d) \cdot (\delta - \xi) \\ \geq [\nabla(\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d)] \cdot (\delta - \xi). \end{aligned}$$

Writing $\delta - \xi = \delta - d + d - \xi$ on the left-hand side and using

$$\nabla(\operatorname{conv} E)(d) \cdot (\xi - d) \leq (\operatorname{conv} E)(\xi) - (\operatorname{conv} E)(d),$$

one arrives at

$$\begin{aligned} & [\nabla(\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d)] \cdot (\delta - \xi) \\ & \leq (\operatorname{conv} E)(\delta) - (\operatorname{conv} E)(d) - \nabla(\operatorname{conv} E)(d) \cdot (\delta - d) \leq C |\delta - d|^{1+\alpha}, \end{aligned}$$

where the last inequality follows from (4.18). Dividing both sides by $|\delta - \xi|$ and since

$$|\delta - \xi| \geq |\delta - d| - |d - \xi| = |d - \xi|,$$

one finds

$$[\nabla(\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d)] \cdot \frac{\delta - \xi}{|\delta - \xi|} \leq 2^{1+\alpha} C |\xi - d|^\alpha.$$

This inequality holds with $\delta \in S_\xi$ arbitrary. Taking $\delta \in S_\xi$ such that $\delta - \xi$ is the appropriate positive multiple of $\nabla(\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d)$ yields²

$$|\nabla(\operatorname{conv} E)(\xi) - \nabla(\operatorname{conv} E)(d)| \leq 2^{1+\alpha} C |\xi - d|^\alpha,$$

for every ξ with $|d - \xi| \leq r$ and every $d \in \overline{B}(d_0, r)$. In particular, one may take d and ξ arbitrary in the ball $\overline{B}(d_0, r/2)$ and the proof is complete. \square

We shall conclude with a few remarks. First, suppose that ∇E satisfies a *uniform* Hölder condition in the sense that for every ξ_0 in the open neighborhood of $\operatorname{dom}(E)^*$ referred to in condition (4.2), one has

$$(4.19) \quad |\nabla E(\xi') - \nabla E(\xi)| \leq C |\xi' - \xi|^\alpha,$$

for ξ' and ξ close enough to ξ_0 (possibly depending on ξ_0), where $C \geq 0$ is independent of ξ_0 . Then, the same constant C can be taken as $C(L)$ in Lemma 4.3, namely independently of L (although ε may depend on L). In turn, the constant $C(\kappa) = (n+1)^{1+\alpha} C(L)$ in Lemma 4.4 can be taken independently of κ (although η may depend on κ).

From the proof of Theorem 4.1 and the above observations, one finds that for every $d_0 \in \operatorname{conv} \mathbf{D}$, one has

$$(4.20) \quad |\nabla(\operatorname{conv} E)(\delta) - \nabla(\operatorname{conv} E)(d)| \leq [2(n+1)]^{1+\alpha} C |\delta - d|^\alpha,$$

for δ and d close enough to d_0 (possibly depending on d_0), with C as in (4.19). In other words, a uniform choice can be made for the Hölder constants associated with $\nabla(\operatorname{conv} E)$.

More particularly, if $\alpha = 1$, then (4.20) holds for every pair $\delta, d \in \operatorname{conv} \mathbf{D}$ (so that $\operatorname{conv} E$ is Lipschitz continuous with constant $4(n+1)^2 C$ on $\operatorname{conv} \mathbf{D}$). Indeed, it suffices to cover the line segment $[d, \delta] \subset \operatorname{conv} \mathbf{D}$ with finitely many open subsets in which (4.20) holds with $\alpha = 1$ and consider a subdivision (ξ_l) , $0 \leq l \leq N+1$, of $[d, \delta]$ such that $\xi_0 = d$, $\xi_{N+1} = \delta$, and

$$|\nabla(\operatorname{conv} E)(\xi_{l+1}) - \nabla(\operatorname{conv} E)(\xi_l)| \leq 4(n+1)^2 C |\xi_{l+1} - \xi_l|,$$

for $0 \leq l \leq N$. The result follows from $|\delta - d| = \sum_{l=0}^N |\xi_{l+1} - \xi_l|$ since the ξ_l 's are aligned.

As an example, when this situation is encountered, take $\mathbf{D} = \operatorname{dom}(E) = \mathbf{R}^n$ with $E \in C^{1,1}(\mathbf{R}^n)$ (in practice, E asymptotically quadratic at infinity).

Remark 4.3. In any case, with nothing more than the assumptions of Theorem 4.1, one has for every compact subset $L \subset \operatorname{conv} \mathbf{D}$ that there is a constant $C(L) \geq 0$ such that

$$(4.21) \quad |\nabla(\operatorname{conv} E)(\delta) - \nabla(\operatorname{conv} E)(d)| \leq C(L) |\delta - d|^\alpha,$$

²That such a multiple exists is straightforward, although S_ξ is centered at d , not ξ .

for every pair δ, d in L . To see this, one may assume that L is convex by replacing L by $\text{conv } L$ (compact) if necessary. One may then cover L by finitely many, say N , open subsets in which the analog of (4.21) holds, and define a constant $C_0(L)$ as the largest constant involved in the corresponding inequalities. As above, one may consider a subdivision of the line segment $[d, \delta]$. Observing that the number of points in this subdivision can be uniformly bounded by N , and through repeated use of the inequality $a^\alpha + b^\alpha \leq 2^{1-\alpha}(a+b)^\alpha$ for $a, b \geq 0$ and $0 < \alpha \leq 1$ (a number of times bounded in terms N), one obtains $C(L)$ from α and $C_0(L)$. Note that even if $C_0(L)$ is independent of L , this procedure does not yield $C(L)$ independent of L , unless $\alpha = 1$. \square

Finally, let us mention that despite local $C^{1,\alpha}$ regularity of $\text{conv } E$ can be proved in a variety of other situations when condition (4.1) is not fulfilled, it is not clear that a genuine extension of Theorem 3.2 exists. The corresponding proof of such a result would definitely require going through considerable technicalities to classify all the possible boundary phenomena. Rather fortunately, it seems that the simplifying assumption (4.1) is sufficient in most applications, at least for the time being.

REFERENCES

1. H. Brezis, *Seuil de régularité pour certains problèmes unilatéraux*, C. R. Acad. Sci. Paris **273** (1971), 35–37.
2. I. Ekeland and R. Temam, *Analyse convexe et problèmes variationnels*, Dunod, 1974.
3. K. Glashoff and S. A. Gustafson, *Linear optimization and approximation*, Appl. Math. Sci. **45**, Springer, 1983.
4. A. Griewank and P. J. Rabier, *Generic aspects of phase equilibria* (in preparation).
5. J. P. Raymond, *Hamiltoniens, relaxation et existence de solutions en calcul des variations*, J. Differential Equations **70** (1987), 226–274.
6. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.
7. J. Stoer and C. Witzgall, *Convexity and optimization in finite dimension I*, Grundlehren Math. Wiss. **163**, Springer, 1970.

MATHEMATICS AND COMPUTER SCIENCE DIVISION, ARGONNE NATIONAL LABORATORY, 9700 SOUTH CASS AVENUE, ARGONNE, ILLINOIS 60439-4844

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PENNSYLVANIA 15260