Vol. 51, No. 3 DUKE MATHEMATICAL JOURNAL © September 1984

DIFFERENTIABILITY PROPERTIES OF SYMMETRIC AND ISOTROPIC FUNCTIONS

J. M. BALL

1. Introduction. Let $\mathbf{G}$ be a group of linear transformations $T: \mathbb{R}^n \to \mathbb{R}^n$. A function $f: \mathbb{R}^n \to \mathbb{R}$ is *invariant* under $\mathbf{G}$ if $f(Tx) = f(x)$ for all $x \in \mathbb{R}^n$, $T \in \mathbf{G}$. Many well-known representation theorems take the following form: for a particular group $\mathbf{G}$, a function f is invariant if and only if there exists a function F such that

$$f(x) = F(y(x)) \qquad \text{for all} \quad x \in \mathbb{R}^n,$$

where $y(x)$ is a vector of preferred new variables. The object of this article is to study, in certain special cases, how the differentiability properties of f and F are related. In the examples we consider $y(\cdot)$ takes values in some subset of $\mathbb{R}^m$, $m \leqslant n$, and F may be required to have its own invariance properties. Also f may only be defined on an invariant subset of $\mathbb{R}^n$. In general there are many different possible representations F of f in terms of different sets of new variables $y(x)$.

Perhaps the best known result of this type, due to Whitney [1942], concerns the case $n = 1$, $\mathbf{G} = \{1, -1\}$. Clearly f is invariant under $\mathbf{G}$ if and only if $f(t) = f(-t)$ for all $t \in \mathbb{R}$ (i.e., f is *even*), and this holds if and only if $f(t) = F(t^2)$ for all $t \in \mathbb{R}$ and some $F: [0, \infty) \to \mathbb{R}$. (An example of another representation is $f(t) = \tilde{F}(t^4)$.) Whitney's result is that if $f \in C^{2r}(\mathbb{R})$ then $F \in C^r([0, \infty))$, and that in general $F \notin C^{r+1}([0, \infty))$, so that there is a loss of derivatives in going from f to its representation F.

In Section 3 we study the case $\mathbf{G} = \mathbf{S}_n = \{\text{permutations on } n \text{ symbols}\}$. A function $f: \mathbb{R}^n \to \mathbb{R}$ is invariant under $\mathbf{S}_n$ if and only if f is *symmetric*, i.e.,

$$f(x_{P1}, \ldots, x_{Pn}) = f(x_1, \ldots, x_n) \qquad \text{for all} \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \quad P \in \mathbf{S}_n,$$

and, as is well known, this holds if and only if

$$f(x) = F(S(x)) \qquad \text{for all} \quad x \in \mathbb{R}^n,$$

for some F, where $S(x) = (S_1(x), \ldots, S_n(x))$ denotes the n-vector of elementary symmetric functions of x, i.e.,

$$S_1(x) = -(x_1 + \cdots + x_n), \qquad S_2(x) = x_1x_2 + \cdots, \ldots,$$
$$S_n(x) = (-1)^n x_1 x_2 \ldots x_n .$$

Received November 1, 1982. Revision received December 19, 1983.

We give a proof (Theorem 3.2) of a result of Barbançon [1972] (for developments see Barbançon & Rais [1982]), that if $f \in C^{nr}(\mathsf{R}^n)$ then $F = C^r(\bar{D})$, where $D = \operatorname{int} S(\mathsf{R}^n)$. The same result for $r = \infty$ was proved earlier by Glaeser [1963]. We also show (Theorem 3.8) that if f has the form

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \phi(x_i)$$

then $f \in C^{nr}(\mathsf{R}^n)$ if and only if $F \in C^r(\bar{D})$, so that the loss of derivatives given by Theorem 3.2 is optimal. To see that a loss of derivatives occurs, consider the example $n = 2$, $r = 1$. Then

$$f(x_1, x_2) = F(-(x_1 + x_2), x_1 x_2).$$

Differentiating formally we obtain

$$\frac{\partial f}{\partial x_1} = -\frac{\partial F}{\partial S_1} + x_2 \frac{\partial F}{\partial S_2},$$

$$\frac{\partial f}{\partial x_2} = -\frac{\partial F}{\partial S_1} + x_1 \frac{\partial F}{\partial S_2},$$

so that

$$-\frac{\partial F}{\partial S_1} = \frac{x_1 \dfrac{\partial f}{\partial x_1} - x_2 \dfrac{\partial f}{\partial x_2}}{x_1 - x_2}, \qquad -\frac{\partial F}{\partial S_2} = \frac{\dfrac{\partial f}{\partial x_1} - \dfrac{\partial f}{\partial x_2}}{x_1 - x_2}.$$

To evaluate these derivatives as $x_1 \to x_2$ requires the existence of certain second derivatives of f. Of course f and F belong to the same differentiability class away from $x_1 = x_2$. The result of Barbançon is stronger than that proved here because he uses a different definition of $C^r(\bar{D})$. Throughout this paper we mean by $C^r(\bar{D})$, where $D \subset \mathsf{R}^n$ is open, the set of r-times continuously differentiable functions $u : D \to \mathsf{R}$ which together with their derivatives of order $\leqslant r$ have continuous extensions to $\bar{D}$. Barbançon uses the definition $\tilde{C}^r(\bar{D}) = \{u : \bar{D} \to \mathsf{R} : u$ is the restriction to $\bar{D}$ of a function $\tilde{u} \in C^r(\mathsf{R}^n)\}$. We always have $\tilde{C}^r(\bar{D}) \subset C^r(\bar{D})$, but if ∂D is irregular then equality may not hold, and it is not known if equality holds for $D = \operatorname{int} S(\mathsf{R}^n)$ if $n > 5$. (If $n \leqslant 4$ equality is proved in Barbançon [1969], who states that he can prove equality also for $n = 5$.) Barbançon's proof of Theorem 3.2 is neither elementary nor simple, and avoids proving equality by means of a complexification technique. Our proof, though giving a weaker result, is completely elementary and if a proof of equality for $D = \operatorname{int} S(\mathsf{R}^n)$ were found would probably provide the quickest route to Barbançon's theorem.

In Section 4 we discuss the case $\mathbf{G} = \mathbf{O}(n)$ = orthogonal group, an example of a group that is compact but not finite. Then $f : \mathsf{R}^n \to \mathsf{R}$ is invariant if and only if

$f(Qx) = f(x)$ for all $x \in \mathbf{R}^n$, $Q \in \mathbf{O}(n)$, which holds if and only if f is *radial*, i.e.,

$$f(x) = F(|x|), \qquad x \in \mathbf{R}^n,$$

for some $F:[0,\infty) \to \mathbf{R}$, where $|\cdot|$ denotes the Euclidean norm. We extend F to $\mathbf{R}$ by making F even. Then (Theorem 4.1) $f \in C^r(\mathbf{R}^n)$ if and only if $F \in C^r(\mathbf{R})$, so that there is no loss of derivatives. A similar result (Theorem 4.2) holds in Hölder spaces. The proofs of these results are a useful warm up for those of Section 5.

In Section 5 we consider functions $h: S^{n\times n} \to \mathbf{R}$, where $S^{n\times n} \cong \mathbf{R}^{n(n-1)/2}$ denotes the space of real symmetric $n \times n$ matrices. We let $\mathbf{G} = \{A \mapsto QAQ^T; Q \in \mathbf{SO}(n)\}$, where $\mathbf{SO}(n)$ denotes the special orthogonal group. Thus f is invariant under $\mathbf{G}$ if and only if f is *isotropic*, i.e.,

$$h(QAQ^T) = h(A), \qquad A \in S^{n\times n}, \quad Q \in \mathbf{SO}(n),$$

and it is well known that this holds if and only if

$$h(A) = H(a_1, \ldots, a_n),$$

for some symmetric function $H: \mathbf{R}^n \to \mathbf{R}$, where the a_i are the eigenvalues of A. We conjecture that $h \in C^r(S^{n\times n})$ if and only if $H \in C^r(\mathbf{R}^n)$, and prove this (Theorem 5.5) in the cases $r = 0$, 1 or 2. We also show (Theorem 5.7) that if r is an arbitrary nonnegative integer and if $0 < \alpha < 1$ then $h \in C^{r,\alpha}(S^{n\times n})$ if and only if $H \in C^{r,\alpha}(\mathbf{R}^n)$. These results are quite surprising since h is a composition of the two mappings $H(\cdot)$ and $A \mapsto (a_1, \ldots, a_n)$, and the second of these is only Lipschitz; nevertheless the composition retains the same differentiability as H. Of course $h \in C^r(S^{n\times n})$ (resp. $h \in C^{r,\alpha}(S^{n\times n})$) implies $H \in C^r(\mathbf{R}^n)$ (resp. $H \in C^{r,\alpha}(\mathbf{R}^n)$) trivially, for all r. It would be very interesting to prove a more general chain rule in which symmetry conditions compensate for reduced differentiability at certain points of one or more maps in a composition. Another way of interpreting Theorem 5.5 and 5.7 is that the set of eigenvalues $\{a_1, \ldots, a_n\}$ of A behaves with respect to symmetry preserving compositions *as if it were a* C^∞ *function of* A.

At several points in our analysis we are able to compute derivatives of functions at points not belonging to some small exceptional set, but need to draw conclusions concerning differentiability everywhere. A useful technical lemma enabling this to be done is stated and proved in Section 3.

The article was motivated by the example of isotropic stored-energy functions in nonlinear elasticity, and we apply our results to this case in Section 6. There are several different representations commonly used for the stored-energy function W of a homogeneous isotropic material in n space dimensions. For instance, we can write

$$W(F) = \Phi(v_1, \ldots, v_n),$$

where Φ is a symmetric function of the singular values $v_1, \ldots, v_n$ of the deformation gradient F, or

$$W(F) = \tilde{H}(I_1(B), \ldots, I_n(B)),$$

where the $I_i(B)$ denote the invariants of $B = FF^T$, that is the elementary symmetric functions of the eigenvalues of B. Our results show that $W \in C^\infty$ if and only if $\Phi, \tilde{H} \in C^\infty$, but that in general $\tilde{H}$ is less differentiable than W or Φ. A loss of differentiability in a different but related situation is discussed by Serrin [1959]. Explicit expressions in terms of Φ obtained for the second derivatives of W are applied to recover known results concerning convexity and strong ellipticity.

Notation. In this article, except where stated otherwise, C denotes a generic constant whose value may vary from line to line. We do *not* use the summation convention for repeated suffices.

2. Extension of differentiable functions. Let $U \subset \mathbf{R}^n$ be open and let $f \in C(U; \mathbf{R}^m)$, the space of r times continuously differentiable real valued functions on U with values in $\mathbf{R}^m$. We will frequently encounter the following situation: we know that $f \in C^r(U \backslash K; \mathbf{R}^m)$ for some 'small' set K, that the derivatives of f up to order r extend continuously to U, and we wish to conclude that $f \in C^r(U; \mathbf{R}^m)$. For this purpose it is not sufficient that K have measure zero or be of first category. For example, let $U = (0, 1)$ and let K be the Cantor set. Let $f \in C([0, 1])$ be a 'Cantor function', that is a function satisfying $f(0) = 0$, $f(1) = 1$ and $f'(t) = 0$ for $t \notin K$; then f is not C^1 but all derivatives of f extend continuously.

Definition 2.1. A set $K \subset \mathbf{R}^n$ is *sparse* if given any $x \in K$ and any nonzero $\zeta \in \mathbf{R}^n$ there exist sequences $x_{(j)} \to x$, $\zeta_{(j)} \to \zeta$ and a number $\epsilon > 0$ such that for each $j = 1, 2, \ldots$ the line segment $\{x_{(j)} + t\zeta_{(j)} : t \in [0, \epsilon]\}$ intersects K at most countably often.

Clearly any sparse set K has empty interior. Note that if $K = \{x \in \mathbf{R}^n; p(x) = 0\}$, where p is a nonconstant polynomial in the coordinates of x, then K is sparse. Indeed, let $x \in K$, $0 \neq \zeta \in \mathbf{R}^n$, and suppose for contradiction that $\{x + t(\zeta + y) : t \in [0, \epsilon]\} \cap K$ contains many points for all sufficiently small $\epsilon > 0$ and $|y|$. Then since $p(x + t(\zeta + y))$ is analytic in t it vanishes identically for sufficiently small $t > 0$, $|y|$. Thus K contains a ball, which is impossible since p is a polynomial.

PROPOSITION 2.2. *Let $U \subset \mathbf{R}^n$ be open and let $K \subset \mathbf{R}^n$ be closed and sparse. Let $f \in C(U; \mathbf{R}^m) \cap C^r(U \backslash K; \mathbf{R}^m)$ be such that for each j with $0 < |j| \leqslant r$ and each $y \in U \cap K$ the limit of $D^j f(x)$ as $x \to y$ with $x \in U \backslash K$ exists. Then $f \in C^r(U; \mathbf{R}^m)$.*

Proof. By induction it is sufficient to prove the result for $r = 1$. We first suppose also that $n = 1$. Clearly we may assume U to be connected. For $n = 1$ $K \subset \mathsf{R}$ is itself countable and it follows from our assumption that $Df: U\backslash K \to \mathsf{R}^m$ has a continuous extension $g \in C(U; \mathsf{R}^m)$. Let $t_0 \in U$ and define for $t \in U$

$$h(t) = f(t) - f(t_0) - \int_{t_0}^{t} g(s)\, ds.$$

Then $Dh(t)$ exists and is zero except for at most countably many $t \in U$ and it follows from Dieudonné [1960 Theorem 8.7.1] that $h \equiv 0$, so that $f \in C^1(U; \mathsf{R}^m)$.

Suppose $n > 1$ and let f satisfy the hypotheses of the proposition. It suffices to prove that each partial derivative $D_i f$ exists and is continuous on U, and we are given that $D_i f$ has a continuous extension $g_i: U \to \mathsf{R}^m$. Let $x \in U \cap K$, let $0 \neq \zeta \in \mathsf{R}^n$, and let $x_{(j)}$, $\zeta_{(j)}$ and ϵ be as in Definition 2.1. We can suppose that $x_{(j)} + t\zeta_{(j)} \in U$ for all $t \in [0, \epsilon]$ and all j. For each j the set $K_j \overset{\text{def}}{=} \{t \in [0, \epsilon];\ x_{(j)} + t\zeta_{(j)} \in K\}$ is closed and countable. Applying the case $n = 1$ to the function $f_j(t) \overset{\text{def}}{=} f(x_{(j)} + t\zeta_{(j)})$ we deduce that $f_j \in C^1(0, \epsilon; \mathsf{R}^m)$ and that for $0 < \tau \leqslant t < \epsilon$

$$f(x_{(j)} + t\zeta_{(j)}) - f(x_{(j)} + \tau\zeta_{(j)}) = \int_{\tau}^{t} \sum_{i=1}^{n} g_i(x_{(j)} + s\zeta_{(j)})(J_{(j)})_i\, ds. \tag{2.1}$$

By continuity (2.1) holds also for $\tau = 0$. Writing

$$f(x + t\zeta) - f(x) = f(x + te_i) - f(x_{(j)} + t\zeta_{(j)})$$
$$+ f(x_{(j)} + t\zeta_{(j)}) - f(x_{(j)}) + f(x_{(j)}) - f(x)$$

and letting $j \to \infty$ we thus obtain

$$f(x + t\zeta) - f(x) = \int_0^t \sum_{i=1}^{n} g_i(x + s\zeta)\zeta_i\, ds \tag{2.2}$$

for $0 < t < \epsilon$. Setting $\zeta = \pm e_i$, where e_i denotes the ith basis vector, we see that the partial derivatives $D_i f$ exist at x and $D_i f(x) = g_i(x)$. The result follows. □

Remark 2.3. A similar result to Proposition 2.2 is given by Dieudonné [1960 p. 159 Exercise 6]. (The statement of Dieudonné's result in certain editions accidentally omits an essential hypothesis, that f be continuous.)

3. Symmetric functions in R^n. For $x = (x_1, \ldots, x_n) \in \mathsf{R}^n$ the *elementary* *symmetric functions* are defined by

$$\left.\begin{aligned} &S_0(x) = 1, \\ &S_j(x) = (-1)^j \sum_{1 \leqslant i_1 < \cdots < i_j \leqslant n} x_{i_1} \ldots x_{i_j}, \ 1 \leqslant j \leqslant n \end{aligned}\right\}. \tag{3.1}$$

We write $S(x) = (S_1(x), \ldots, S_n(x))$, so that $S: \mathbb{R}^n \to \mathbb{R}^n$.

Let $E \subset \mathbb{R}^n$ be open and *symmetric*; i.e., $PE = E$ for every permutation P of $x_1, \ldots, x_n$. Let K_n denote the open cone consisting of those points $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with $x_1 > x_2 > \cdots > x_n$. Let $\Omega_E = S(E \cap K_n)$.

LEMMA 3.2.
(i) $\Omega_E = \operatorname{int} S(E)$,
(ii) $\partial\Omega_E = S(\partial E) \cup S(E \cap \partial K_n)$,
(iii) $\overline{\Omega}_E = S(\overline{E})$.

Proof. Let $x = (x_1, \ldots, x_n) \in \Omega_E$ and $w = (w_1, \ldots, w_n) = S(x)$. Since the x_i are the distinct roots of the equation $\lambda^n + w_1\lambda^{n-1} + \cdots + w_{n-1}\lambda + w_n = 0$, it follows from the implicit function theorem that if $|z - w|$ is sufficiently small the equation $S(y) = z$ has a solution $y \in E \cap K_n$. This proves that Ω_E is open and hence $\Omega_E \subset \operatorname{int} S(E)$. Suppose $\theta \in \operatorname{int} S(E)$ but $\theta \notin \Omega_E$. Then $\theta = S(x)$ for some $x \in E \cap \partial K_n$, and hence $x_i = x_j$ for some $i \neq j$. Therefore there is a sequence $\theta_{(i)} \to \theta$ such that the polynomial with coefficients given by $\theta_{(i)}$ has two complex conjugate roots, and hence $\theta \notin \operatorname{int} S(E)$, a contradiction. This proves (i).

Let $\theta \in \partial\Omega_E$. Then there exists a sequence $\theta_{(j)} = S(x_{(j)}) \to \theta$ with $x_{(j)} \in E \cap K_n$. Since the roots $x_{(j)}$ of $\sum_{r=0}^n S_r(x_{(j)})t^r = 0$ can be bounded in terms of the coefficients we may suppose that $x_{(j)} \to x \in \overline{E \cap K_n}$. Since S is continuous, $\theta = S(x)$. Since $\theta \notin \Omega_E$ it follows that $x \notin E \cap K_n$ and hence that $x \in \partial(E \cap K_n)$. If $x \in \partial E$ then $\theta \in S(\partial E)$. If $x \notin \partial E$ then $x \in E$, and since E, K_n are open we obtain $x \in E \cap \partial K_n$ and $\theta \in S(E \cap \partial K_n)$. Thus $\partial\Omega_E \subset S(\partial E) \cup S(E \cap \partial K_n)$. To prove the inverse inclusion, note first since each S_i is symmetric, $S(A) = S(A \cap \overline{K}_n)$ for any symmetric set A. Let $\theta \in S(\partial E) \cup S(E \cap \partial K_n) = S(\partial E \cap \overline{K}_n) \cup S(E \cap \partial K_n)$. Then $\theta = S(x)$ for some $x \in (\partial E \cap \overline{K}_n) \cup (E \cap \partial K_n)$. Since E is open and symmetric there exists a sequence $x_{(j)} \to x$ with $x_{(j)} \in E \cap K_n$. Therefore $\theta = \lim S(x_{(j)}) \in \overline{S(\Omega_E)}$. But $S|_{\overline{K}_n}$ is one-to-one, and so $S(E \cap K_n) \cap S((\partial E \cap \overline{K}_n) \cup (E \cap \partial K_n))$ is empty. Hence $\theta \in \partial\Omega_E$ and $S(\partial E) \cup S(E \cap \partial K_n) \subset \partial\Omega_E$.

To prove (iii) we use (ii) to obtain

$$\begin{aligned} \overline{\Omega}_E &= S(E \cap K_n) \cup S(\partial E) \cup S(E \cap \partial K_n) \\ &= S(E \cap \overline{K}_n) \cup S(\partial E) \\ &= S(E) \cup S(\partial E) = S(\overline{E}). \quad \square \end{aligned}$$

If $r \geq 0$ and $\Omega \subset \mathbf{R}^n$ is open, then $C^r(\overline{\Omega}; \mathbf{R}^m)$ denotes the space of functions $f \in C^r(\Omega; \mathbf{R}^m)$ which together with their derivatives of order $\leq r$ have continuous extensions to $\overline{\Omega}$. If Ω is bounded then $C^r(\overline{\Omega}; \mathbf{R}^m)$ is a Banach space with norm $\|f\| = \sum_{|j| \leq r} \sup_{x \in \Omega} |D^j f(x)|$. We write $C^r(\overline{\Omega}) = C^r(\overline{\Omega}; \mathbf{R})$ and $C^\infty(\overline{\Omega}) = \bigcap_{r=0}^\infty C^r(\overline{\Omega})$. (Note that if $\Omega_1 \subset \Omega$ is open with $\overline{\Omega}_1 = \overline{\Omega}$ then in general $C^r(\overline{\Omega}_1) \neq C^r(\overline{\Omega})$; for example, if $\Omega = (0, 1)$, K is the Cantor set and $\Omega_1 = \Omega \backslash K$ then any Cantor function $f \in C^\infty(\overline{\Omega}_1) \backslash C^1(\overline{\Omega})$. However, if in addition $\overline{\Omega \backslash \Omega_1}$ is sparse it follows easily from Proposition 2.2 that $C^r(\overline{\Omega}_1) = C^r(\overline{\Omega})$.)

Given any symmetric function $f: E \to \mathbf{R}$ there exists a unique function $F: S(E) \to \mathbf{R}$ such that

$$f(x) = F(S(x)) \qquad \text{for all} \quad x \in E. \tag{3.2}$$

Our main result relates the differentiability properties of f and F. For simplicity we assume from now on that E is *convex*; this hypothesis can be weakened (see below).

THEOREM 3.2 (Barbançon [1972]). *If $f \in C^{nr}(\overline{E})$, then $F \in C^r(\overline{\Omega}_E)$, $r = 0, 1, 2, \ldots$.*

COROLLARY 3.3 (Glaeser [1963]). *If $f \in C^\infty(\overline{E})$, then $F \in C^\infty(\overline{\Omega}_E)$.*

Both Barbançon and Glaeser proved their results for the case $E = \mathbf{R}^n$, but our statements can easily be deduced from theirs by first extending f to a function $\tilde{f} \in C^{nr}(\mathbf{R}^n)$ by means of the following result of Whitney (which is a consequence of the more familiar version of the Whitney extension theorem (Whitney [1934a], Federer [1969])).

THEOREM 3.4 (Whitney [1934b]). *Let $\Omega \subset \mathbf{R}^n$ be open and satisfy the following regularity condition: for any $M > 0$ there is a constant C_M such that if $x, y \in \Omega$ with $|x| \leq M$, $|y| \leq M$ then x may be joined to y by a rectifiable curve lying in Ω with length $l \leq C_M |x - y|$. Let $g \in C^s(\overline{\Omega})$. Then there exists $\tilde{g} \in C^s(\mathbf{R}^n)$ with $\tilde{g}|_{\overline{\Omega}} = g|_{\overline{\Omega}}$.*

(If the boundary of Ω is not regular, the extension in Theorem 3.4 is not in general possible. A one-dimensional example is furnished by a Cantor function, as described above. An example with Ω a two-dimensional domain with an inward pointing cusp is given in Gilbarg & Trudinger [1977 p. 52] and is discussed by Fraenkel [1982].)

Thus Theorem 3.2 and Corollary 3.3 hold if E is symmetric and satisfies the regularity condition in Theorem 3.4, and even this regularity condition is only needed near the set $\{x : x_i = x_j \text{ for some } i \neq j\}$. The proof given below also works if E is symmetric and such that $x \in E$ implies $(tx_1 + (1-t)x_2, tx_2 + (1-t)x_1, x_3, \ldots, x_n) \in E$ for all $t \in [0, 1]$.

As explained in the introduction, Barbançon proves the stronger result that $f \in C^{nr}(\mathbf{R}^n)$ implies $F \in \tilde{C}^r(\overline{\Omega}_{\mathbf{R}^n})$. This would follow from Theorem 3.2 if it were true that $\Omega_{\mathbb{R}^n} = \operatorname{int} S(\mathbb{R}^n)$ satisfies the regularity condition in Theorem 3.4, this being unknown for $n > 5$. The first nontrivial case is $n = 3$, when $\Omega_{\mathbb{R}^n}$ is the region lying between the two surfaces which are the images under S of the sets $\{x_1 = x_2 > x_3\}$ and $\{x_1 > x_2 = x_3\}$, these surfaces being joined along the line $t \mapsto (-3t, 3t^2, -3t^3)$ corresponding to polynomials with three equal roots. The regularity condition is then an easily proved generalization of the statement that 'the cusp points outwards'; cf. Arnold [1983 p. 256] and the principle that 'everything good is fragile'. Despite some efforts and seeking of advice the author has not been able to find a proof for arbitrary n. The most obvious attempt at a proof fails. This consists of taking as the path joining $S(x)$ and $S(y)$, where $x_1 > x_2 > \cdots > x_n$, $y_1 > y_2 > \cdots > y_n$, the path $S(tx + (1-t)y)$, $t \in [0,1]$. Even in the case $n = 3$, when the desired result holds, this path does not satisfy the inequality $l \leqslant C|x - y|$, as can be seen by taking

$$x = \epsilon\left(\frac{1+\sqrt{5}}{4}, \frac{1}{2}, \frac{1-\sqrt{5}}{4}\right), \qquad y = \epsilon(1,0,0)$$

and letting $\epsilon \to 0$. Some possibly relevant information on the structure of $A(\mathbb{R}^n)$, where $A(x) = (A_1(x), \ldots, A_n(x))$ and $A_j(x) = \sum_{i=1}^n x_i^j$ is the jth Newton polynomial, is given in Ursell [1959], and on account of the diffeomorphism between the A_i and the S_i given by Newton's formulae it is clear that $S(\mathbb{R}^n)$ is regular if and only if $A(\mathbb{R}^n)$ is. $S(\mathbb{R}^n)$ and $A(\mathbb{R}^n)$ may be characterized by finitely many polynomial inequalities (Gantmacher [1960 p. 203]), but it is not obvious how to exploit this to give a proof.

Before proving Theorem 3.2 we introduce some notation. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. If $n > 1$ we write $x^{[i]} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$, $i = 1, 2, \ldots, n$. Note that by (3.1)

$$\frac{\partial S_j(x)}{\partial x_i} = (-1)^j \sum_{\substack{1 \leqslant i_1 < \cdots < i_{j-1} \leqslant n \\ i_k \neq i}} x_{i_1} \cdots x_{i_{j-1}} = -S_{j-1}(x^{[i]}), \qquad 1 \leqslant i, j \leqslant n, \tag{3.3}$$

with the obvious convention if $n = 1$.

By a simple and well-known application of the implicit function theorem the roots of a polynomial depend smoothly on its coefficients in any domain in which the roots remain distinct. Thus, by the chain rule, if $f \in C^k(\overline{E})$ then $F \in C^k(\Omega_E)$. In particular, if $f \in C^1(\overline{E})$ then

$$\begin{bmatrix} f_{,1}(x) \\ \vdots \\ f_{,n}(x) \end{bmatrix} = -\begin{bmatrix} S_0(x^{[1]}) & \cdots & S_{n-1}(x^{[1]}) \\ \vdots & & \vdots \\ S_0(x^{[n]}) & \cdots & S_{n-1}(x^{[n]}) \end{bmatrix} \begin{bmatrix} F_{,1}(S(x)) \\ \vdots \\ F_{,n}(S(x)) \end{bmatrix} \tag{3.4}$$

for all $x = (x_1, \ldots, x_n) \in E$ with x_i distinct, where the commas denote partial differentiation. We may write (3.4) in the abbreviated form

$$Df(x) = -A(x)Dg(S(x)), \tag{3.5}$$

where $A(x)$ is the $n \times n$ matrix with components $A_{ij}(x) = S_{j-1}(x^{[i]})$.

LEMMA 3.5.

$$A(x)^{-1} = \begin{bmatrix} x_1^{n-1} & \cdots & x_n^{n-1} \\ \vdots & & \vdots \\ x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \Pi_1(x)^{-1} & & 0 \\ & \ddots & \\ 0 & & \Pi_n(x)^{-1} \end{bmatrix}$$

for all $x = (x_1, \ldots, x_n)$ *with the* x_i *distinct, where* $\Pi_j(x) \stackrel{\text{def}}{=} \prod_{i \neq j}(x_j - x_i)$. *(We set* $\Pi_1(x) = 1$ *if* $n = 1$.)

Proof. The equation $S_0(x^{[i]})\lambda^{n-1} + \cdots + S_{n-2}(x^{[i]})\lambda + S_{n-1}(x^{[i]}) = 0$ has roots $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$. Therefore

$$S_0(x^{[i]})x_j^{n-1} + \cdots + S_{n-2}(x^{[i]})x_j + S_{n-1}(x^{[i]}) = \prod_{k \neq i}(x_j - x_k)$$

for $1 \leqslant i, j \leqslant n$. The result follows immediately. □

By Lemma 3.5 and (3.4) we have that for $i = 1, 2, \ldots, n$

$$F_{,i}(S(x)) = -\left[x_1^{n-i}\Pi_1(x)^{-1}f_{,1}(x) + \cdots + x_n^{n-i}\Pi_n(x)^{-1}f_{,n}(x)\right] \tag{3.6}$$

provided the x_i are distinct. From (3.6) we see that $F_{,i}$ has the form

$$F_{,i}(S(x)) = h(x_1, x_2, \ldots, x_n)\Pi_1(x)^{-1} + h(x_2, x_3, \ldots, x_n, x_1)\Pi_2(x)^{-1}$$
$$+ \cdots + h(x_n, x_1, \ldots, x_{n-1})\Pi_n(x)^{-1},$$

where h is symmetric with respect to permutations of its last $n-1$ arguments.

LEMMA 3.6. *Let* $m \geqslant 0$, $s \geqslant 0$. *Let* $A \subset \mathbf{R}^{n+m}$ *be open, convex, and such that if* $(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m) \in A$ *so does* (Px, y) *for any permutation* P *of* $x_1, \ldots, x_n$. *Let* $h = h(x; y) = h(x_1, \ldots, x_n; y_1, \ldots, y_n)$ *belong to* $C^{n+s-1}(\overline{A})$ *and be symmetric with respect to permutations of* $x_2, \ldots, x_n$.

Then

$$I(x, y) \overset{\text{def}}{=} \frac{h(x_1, \ldots, x_n; y)}{\Pi_1(x)} + \frac{h(x_2, \ldots, x_n, x_1; y)}{\Pi_2(x)} + \cdots + \frac{h(x_n, x_1, \ldots, x_{n-1}; y)}{\Pi_n(x)},$$

defined for $(x, y) \in A$ *with the* x_i *distinct, has an extension belonging to* $C^s(\bar{A})$.

Remark 3.7. We will use Lemma 3.5 only in the case $m = 0$. The extra variables y are introduced only to facilitate the proof.

Proof of Lemma 3.6. The proof is by induction on n. When $n = 1$ the result is trivial. Let $n = 2$, $h \in C^{s+1}(\bar{A})$. We have to prove that

$$I(x, y) = \frac{h(x_1, x_2; y)}{x_1 - x_2} + \frac{h(x_2, x_1; y)}{x_2 - x_1}$$

has an extension in $C^s(\bar{A})$. But since A is convex,

$$I(x, y) = \frac{1}{x_1 - x_2}\int_0^1 \frac{d}{dt} h(x_2 + t(x_1 - x_2), x_1 + t(x_2 - x_1); y)\, dt$$

$$= \int_0^1 (h_{,1} - h_{,2})(x_2 + t(x_1 - x_2), x_1 + t(x_2 - x_1); y)\, dt \qquad (3.7)$$

when $x_1 \neq x_2$. Let $K = \{(x, y) \in \mathbb{R}^{2+m} : x_1 = x_2\}$. K is a closed, sparse subset of $\mathbb{R}^{2+m}$. By (3.7), for each $(\bar{x}, \bar{y}) \in \bar{A}$ and each multi-index j with $0 \leqslant |j| \leqslant s$ the limit of $D^j I(x, y)$ as $(x, y) \to (\bar{x}, \bar{y})$ with $(x, y) \in A \setminus K$ exists. If $(x, y) \in \partial A \cup (A \cap K)$ we define $I(x, y)$ by continuity, so that $I \in C(\bar{A}) \cap C^s(A \setminus K)$. It follows from Proposition 2.2 that $I \in C^s(A)$ and hence that $I \in C^s(\bar{A})$ as required.

Suppose now that the result holds for $n - 1$ and for all $m \geqslant 0$, $s \geqslant 0$ and A. We prove it holds for n. Let $h \in C^{n+s-1}(\bar{A})$. We first remark that by Lemma 3.5 and (3.5) with $f(x) = x_1 + \cdots + x_n$ we have that

$$\sum_{j=1}^{n} \Pi_j(x)^{-1} = 0,$$

provided the x_i are distinct. Hence, when the x_i are distinct,

$$I(x, y) = \frac{h(x_2, \ldots, x_n, x_1; y) - h(x_1, x_2, \ldots, x_n; y)}{\Pi_2(x)} + \cdots + \frac{h(x_n, x_1, \ldots, x_{n-1}; y) - h(x_1, x_2, \ldots, x_n; y)}{\Pi_n(x)}.$$

Let

$$H(x_2, x_3, \ldots, x_n; x_1, y) \stackrel{\text{def}}{=} \frac{h(x_2, x_3, \ldots, x_n, x_1; y) - h(x_1, x_2, x_3, \ldots, x_n; y)}{x_2 - x_1}.$$

We apply the case $n = 2$ to H, absorbing the variables $x_3, \ldots, x_n$ into a new $y \in \mathbf{R}^{m+n-2}$. Thus H has an extension, again denoted by H, belonging to $C^{n+s-2}(\bar{A}) = C^{(n-1)+s-1}(\bar{A})$. We now note that

$$I(x, y) = \frac{H(x_2, \ldots, x_n; x_1, y)}{\Pi_1(x^{[1]})} + \frac{H(x_3, x_4, \ldots, x_n, x_2; x_1, y)}{\Pi_2(x^{[1]})} + \cdots + \frac{H(x_n, x_2, \ldots, x_{n-1}; x_1, y)}{\Pi_n(x^{[1]})},$$

and apply the induction hypothesis with $(n-1, m+1)$ for (n, m). It follows that I has an extension in $C^s(\bar{A})$, which completes the proof. □

Proof of Theorem 3.2. We use induction on r. When $r = 0$ the result is a consequence of the symmetry of f and the fact that the set of roots of a polynomial varies continuously with the coefficients.

Suppose the result holds for $r - 1$. Let $f \in C^{nr}(\bar{E})$. It suffices to show that $F \in C^1(\bar{\Omega}_E)$ and $F_{,i} \in C^{r-1}(\bar{\Omega}_E)$ for each $i = 1, \ldots, n$. But $F_{,i}(S(\cdot))$ is given by (3.6) and satisfies the hypotheses of Lemma 3.5 with $m = 0$, $A = E$ and $h \in C^{nr-1}(\bar{E}) = C^{n+n(r-1)-1}(\bar{E})$. Thus $F_{,i}(S(\cdot))$ has an extension belonging to $C^{n(r-1)}(\bar{E})$. The extension is clearly symmetric in $x_1, \ldots, x_n$ and thus can be written as a function $\tilde{F}_i : S(E) \to \mathbf{R}$. By the induction hypothesis $\tilde{F}_i \in C^{r-1}(\bar{\Omega}_E)$. But $F \in C^{nr}(\Omega_E) \cap C^0(\bar{\Omega}_E)$ and $F_{,i}(\theta) = \tilde{F}_i(\theta)$ for all $\theta \in \Omega_E$. Thus $F \in C^1(\bar{\Omega}_E)$ and $F_{,i} \in C^{r-1}(\bar{\Omega}_E)$ as required. □

We now show that the loss of derivatives given by Theorem 3.2 is optimal; thus in general $f \in C^{nr-1}(\bar{E})$ does not imply $F \in C^r(\bar{\Omega}_E)$, and $f \in C^{nr+1}(\bar{E})$ does not imply $F \in C^{r+1}(\bar{\Omega}_E)$ unless $n = 1$. Consider the special case when

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \phi(x_i), \tag{3.8}$$

where $\phi : I \to \mathbf{R}$ and I is an open interval (possibly unbounded). Thus $f : E = I^n \to \mathbf{R}$. Let $F = F_\phi$ be given by (3.2) and (3.8).

THEOREM 3.8.

$$\phi \in C^{nr}(\bar{I}) \quad \textit{if and only if} \quad F \in C^r(\bar{\Omega}_E).$$

The proof of Theorem 3.8 is based on the following lemmas.

LEMMA 3.9. *If ϕ is a polynomial then for all $t \in \bar{I}$ and $r = 0, 1, 2, \dots$,*

$$\frac{d^{nr}\phi(t)}{dt^{nr}} = \frac{(-1)^r(nr)!}{n \cdot r!} \frac{\partial^r F}{\partial S_n^r}\left(nt, \frac{n(n-1)}{2!} t^2, \dots, t^n\right). \tag{3.9}$$

Proof. Suppose that $\phi(t) = \sum_{j=0}^N c_j t^j$ is a polynomial. We use Newton's equations

$$\begin{aligned}
&a_1(x) + S_1(x) = 0,\\
&a_2(x) + S_1(x)a_1(x) + 2S_2(x) = 0,\\
&\vdots\\
&a_n(x) + S_1(x)a_{n-1}(x) + \cdots + nS_n(x) = 0,\\
&a_{n+k}(x) + S_1(x)a_{n+k-1}(x) + \cdots + S_n(x)a_k(x) = 0,
\end{aligned}$$

where $a_j(x) \overset{\text{def}}{=} \sum_{i=1}^n x_i^j$. It follows from these equations that $F(S_1, \dots, S_n)$ is a polynomial and that only the sum $c_{nr}\sum_{i=1}^n x_i^{nr}$ contributes a term in S_n^r, namely $(-1)^r nc_{nr}$. Therefore

$$\frac{d^{nr}\phi(0)}{dt^{nr}} = \frac{(-1)^r(nr)!}{n \cdot r!} \frac{\partial^r F}{\partial S_n^r}(0, 0, \dots, 0). \tag{3.10}$$

Now let $t \neq 0$, $\theta(\tau) = \phi(t + \tau)$, $e = (1, 1, \dots, 1)$. Then

$$\begin{aligned}
\sum_{i=1}^n \theta(x_i) &= \sum_{i=1}^n \phi((x + te)_i)\\
&= F(S_1(x + te), \dots, S_n(x + te))\\
&= F(S_1(x) + tS_1(e), \dots, S_n(x) + \cdots + t^n).
\end{aligned}$$

Applying (3.10) we obtain (3.9). □

LEMMA 3.10. *Let I be a bounded open interval. Then the map $T : \phi \mapsto F_\phi$ from $C^{nr}(\bar{I})$ into $C^r(\bar{\Omega}_{I^n})$ is continuous.*

Proof. T is clearly linear. We show that T has closed graph. Let $\phi_j \to \phi$ in $C^{nr}(\bar{I})$ and $F_{\phi_j} \to F$ in $C^r(\bar{\Omega}_{I^n})$. Let $\sigma_{(j)} \to \sigma$ in $\bar{\Omega}_{I^n}$.

Then $\sigma_{(j)} = S(x_{(j)})$ for some $x_{(j)} \in \bar{I}^n$, and we can suppose that $x_{(j)} \to x$ in $\bar{I}^n$, and thus that $\sigma = S(x)$. Hence

$$F(\sigma) = \lim_{j\to\infty} F_{\phi_j}(\sigma_j) = \lim_{j\to\infty} \sum_{i=1}^n \phi_j(x_{(j)i}) = \sum_{i=1}^n \phi(x_i) = F_\phi(\sigma).$$

Hence $F = F_\phi$ and T is continuous by the closed graph theorem. □

LEMMA 3.11. *If $\phi \in C^{nr}(\bar{I})$ then (3.9) holds for all $t \in \bar{I}$.*

Proof. Without loss of generality we can suppose I is bounded. Let $\phi \in C^{\infty}(\bar{I})$ and let $\{\phi_j\}$ be a sequence of polynomials converging to ϕ in $C^{nr}(\bar{I})$. By Lemma 3.9 ϕ_j satisfies (3.9). Passing to the limit using Lemma 3.10 we obtain (3.9) for ϕ. □

Proof of Theorem 3.8. Let $\phi : \bar{I} \to \mathsf{R}$ with $F = F_{\phi} \in C^{r}(\bar{\Omega}_E)$. Since

$$n\phi(t) = F(nt, \ldots, t^n)$$

it follows that $\phi \in C(\bar{I})$. Let ρ be a mollifier; i.e., $\rho \in C^{\infty}(\mathsf{R})$, $\rho \geqslant 0$, $\int_{\mathsf{R}} \rho(t)\,dt = 1$, $\operatorname{supp}\rho \subset (-1, 1)$. Define $\rho_{\epsilon}(t) = \epsilon^{-1}\rho(t/\epsilon)$ for $\epsilon > 0$. Let J be a bounded open interval with $\bar{J} \subset I$. For $\epsilon > 0$ sufficiently small and $t \in \bar{J}$ let

$$(\rho_{\epsilon} * \phi)(t) \stackrel{\text{def}}{=} \int_{\mathsf{R}} \rho_{\epsilon}(\tau)\phi(t - \tau)\,d\tau.$$

Then if $x \in \bar{J}^n$

$$\begin{aligned} F_{\rho_{\epsilon}*\phi}(S(x)) &= \sum_{i=1}^{n} \int_{\mathsf{R}} \rho_{\epsilon}(\tau)\phi(x_i - \tau)\,d\tau \\ &= \int_{\mathsf{R}} \rho_{\epsilon}(\tau) F_{\phi}\Big(\sum_{i} (x_i - \tau), \sum_{i \neq j} (x_i - \tau)(x_j - \tau), \ldots \Big)\,d\tau \\ &= \int_{\mathsf{R}} \rho_{\epsilon}(\tau) F_{\phi}(S_1(x) - n\tau, \ldots, S_n(x) + \cdots)\,d\tau. \end{aligned}$$

Thus $F_{\rho_{\epsilon}*\phi} \in C^{r}(\bar{\Omega}_{J^n})$ and

$$\frac{\partial^r F_{\rho_{\epsilon}*\phi}}{\partial S_n^r}(S_1, \ldots, S_n) = \int_{\mathsf{R}} \rho_{\epsilon}(\tau) \frac{\partial^r F_{\phi}}{\partial S_n^r}(S_1 - n\tau, \ldots, S_n + \cdots)\,d\tau$$

for all $S = (S_1, \ldots, S_n) \in \bar{\Omega}_{J^n}$. Let $v \in C_0^{\infty}(J)$. Then from Lemma 3.11 we deduce that

$$\begin{aligned} &\frac{(-1)^r n \cdot r!}{(nr)!} \int_{\mathsf{R}} \frac{d^{nr}}{dt^{nr}} (\rho_{\epsilon} * \phi)(t) v(t)\,dt \\ &\qquad = \int_{\mathsf{R}}\int_{\mathsf{R}} \rho_{\epsilon}(\tau) \frac{\partial^r F_{\phi}}{\partial S_n^r}\big(n(t - \tau), \ldots, (t - \tau)^n\big) v(t)\,d\tau\,dt, \end{aligned}$$

and hence that

$$\begin{aligned} &\frac{(-1)^{(n+1)r} n \cdot r!}{(nr)!} \int_{\mathsf{R}}\int_{\mathsf{R}} \rho_{\epsilon}(t - \tau)\phi(\tau) \frac{d^{nr}v(t)}{dt^{nr}}\,d\tau\,dt \\ &\qquad = \int_{\mathsf{R}}\int_{\mathsf{R}} \rho_{\epsilon}(t - \tau) \frac{\partial^r F_{\phi}}{\partial S_n^r}(n\tau, \ldots, \tau^n) v(t)\,d\tau\,dt. \end{aligned}$$

Passing to the limit using standard results for mollifiers we obtain

$$\frac{(-1)^{(n+1)r}n\cdot r!}{(nr)!}\int_{\mathsf{R}}\phi(t)\frac{d^{nr}v(t)}{dt^{nr}}\,dt=\int_{\mathsf{R}}\frac{\partial^r F_\phi}{\partial S_n^r}(nt,\ldots,t^n)\cdot v(t)\,dt.$$

Since v is arbitrary we deduce that

$$\frac{d^{nr}}{dt^{nr}}\phi(t)=\frac{(-1)^r(nr)!}{n\cdot r!}\frac{\partial^r F}{\partial S_n^r}(nt,\ldots,t^n)\qquad\text{in } \mathscr{D}'(J).\tag{3.11}$$

Since J is arbitrary and the right-hand side of (3.11) is continuous on $\bar{I}$ it follows that $\phi\in C^{nr}(\bar{I})$ as required. □

4. Radial functions. In this section we consider functions $g:\mathsf{R}^n\to\mathsf{R}$ which are invariant under the orthogonal group $\mathbf{O}(n)$; i.e., $g(Qx)=g(x)$ for all $x\in\mathsf{R}^n$, $Q\in\mathbf{O}(n)$. Any such g can be written in the form

$$g(x)=G(|x|)\tag{4.1}$$

for some $G:[0,\infty)\to\mathsf{R}$. We extend the domain of G to R by requiring that G be *even*; i.e., $G(\rho)=G(-\rho)$ for all $\rho\in\mathsf{R}$.

THEOREM 4.1. *Let $r\geqslant 0$. Then $g\in C^r(\mathsf{R}^n)$ if and only if $G\in C^r(\mathsf{R})$.*

Proof. Let $g\in C^r(\mathsf{R}^n)$. Let $e\in\mathsf{R}^n$, $|e|=1$. Then $G(\rho)=g(\rho e)$ for all $\rho\in\mathsf{R}$, and hence $G\in C^r(\mathsf{R})$. Conversely, let $G\in C^r(\mathsf{R})$ be even. Then $G_r(\rho)\stackrel{\text{def}}{=}\sum_{j=0}^r D^jG(0)(\rho^j/j!)$ contains no odd powers and consequently $g_r(x)\stackrel{\text{def}}{=}G_r(|x|)$ is smooth. Therefore by subtracting G_r from G we can without loss of generality suppose that $D^jG(0)=0$ for $0\leqslant j\leqslant r$. Since $|x|$ is smooth for $x\neq 0$ it follows from the chain rule that $g\in C^r(\mathsf{R}^n\backslash\{0\})$. It is therefore sufficient (using Proposition 2.2 or a direct argument) to show that $D^\alpha g(x)=o(|x|^{r-|\alpha|})$ as $x\to 0$ for $0\leqslant|\alpha|\leqslant r$ for any function $g(x)=G(|x|)$ with $G\in C^r(\mathsf{R})$ and $D^jG(0)=0$ for $0\leqslant j\leqslant r$. We prove this by induction on r. The case $r=0$ is trivial. Suppose that the result is true for $r-1$ and that $G\in C^r(\mathsf{R})$, $D^jG(0)=0$ for $0\leqslant j\leqslant r$. By the induction hypothesis $D^\beta G'(|x|)=o(|x|^{r-1-|\beta|})$, $0\leqslant|\beta|\leqslant r-1$. Also, it is easily proved that $D^\gamma(x/|x|)=O(|x|^{-|\gamma|})$. Since $Dg(x)=G'(|x|)(x/|x|)$, it follows that if $0<|\alpha|\leqslant r$ then $D^\alpha g(x)$ is a finite sum of terms $D^\beta G'(|x|)$ $D^\gamma(x/|x|)$ with $|\beta|+|\gamma|=|\alpha|-1$. Thus $D^\alpha g(x)=o(|x|^{r-|\alpha|})$ as $x\to 0$. Finally $D^0g(x)=G(|x|)=o(|x|^r)$ by Taylor's theorem, which completes the induction. □

Next we consider Hölder continuity. If $\Omega\subset\mathsf{R}^n$ is open, $k\geqslant 0$, $0<\alpha<1$, we define $C^{k,\alpha}(\Omega)$ to be the subspace of $C^k(\Omega)$ consisting of functions whose kth order partial derivatives are locally Hölder continuous with exponent α in Ω.

THEOREM 4.2. *Let $r\geqslant 0$, $0<\alpha<1$. Then $g\in C^{r,\alpha}(\mathsf{R}^n)$ if and only if $G\in C^{r,\alpha}(\mathsf{R})$.*

Proof. That $g \in C^{r,\alpha}(\mathbb{R}^n)$ implies $G \in C^{r,\alpha}(\mathbb{R})$ follows as in Theorem 4.1. Let $G \in C^{0,\alpha}(\mathbb{R})$. Then if $D \subset \mathbb{R}^n$ is compact and $x, y \in D$,

$$|g(x) - g(y)| = |G(|x|) - G(|y|)| \leqslant C||x| - |y||^\alpha \leqslant C|x - y|^\alpha,$$

so that $g \in C^{0,\alpha}(\mathbb{R}^n)$. If $G \in C^{1,\alpha}(\mathbb{R})$ and $x, y \in D$ with $|x| \leqslant |y|$ then

$$\begin{aligned} |Dg(x) - Dg(y)| &= \left|G'(|x|)\frac{x}{|x|} - G'(|y|)\frac{y}{|y|}\right| \\ &\leqslant |G'(|x|) - G'(|y|)| + |G'(|y|)|\left|\frac{x}{|x|} - \frac{y}{|y|}\right| \\ &\leqslant C|x-y|^\alpha + C|y|^\alpha \frac{|x-y|}{|y|} \\ &\leqslant C|x-y|^\alpha, \end{aligned}$$

where we have used the fact that $G'(0) = 0$. Hence $g \in C^{1,\alpha}(\mathbb{R}^n)$. We complete the proof by induction. Suppose the result is true for $r - 1$, where $r \geqslant 2$. Let $G \in C^{r,\alpha}(\mathbb{R}^n)$. By Theorem 4.1, $g \in C^2(\mathbb{R}^n)$ and

$$\Delta g(x) = G''(|x|) + (n-1)\frac{G'(|x|)}{|x|} \overset{\text{def}}{=} \Gamma(|x|).$$

Since

$$\Gamma(|x|) = G''(|x|) + (n-1)\int_0^{|x|} G'(t|x|)\,dt$$

we have that $\Gamma \in C^{r-2,\alpha}(\mathbb{R})$, that Γ is even, and hence $\Gamma(|\cdot|) \in C^{r-2,\alpha}(\mathbb{R}^n)$ by the induction hypothesis. By the regularity theory for Poisson's equation (see Gilbarg and Trudinger [1977]) it follows that $g \in C^{r,\alpha}(\mathbb{R}^n)$, as required. □

5. Isotropic functions. Let $S^{n\times n}$ denote the $n(n-1)/2$-dimensional vector space of real, symmetric $n \times n$ matrices with inner product $\langle A, B\rangle = \sum_{i,j=1}^n A_{ij}B_{ij}$. Let $E \subset S^{n\times n}$ be open and *invariant* under $\mathbf{SO}(n)$; i.e., if $A \in E$ and $Q \in \mathbf{SO}(n)$ then $QAQ^T \in E$. Let Γ_E denote the set of diagonal matrices belonging to E. We regard Γ_E as a subset of $\mathbb{R}^n$. Clearly Γ_E is open. A function $h : E \to \mathbb{R}$ is said to be *isotropic* if

$$h(QAQ^T) = h(A) \qquad \text{for all} \quad A \in E, \quad Q \in \mathbf{SO}(n). \tag{5.1}$$

It is well known that h is isotropic if and only if there exists a symmetric function $H : \Gamma_E \to \mathbb{R}$ such that

$$h(A) = H(v_1(A), \ldots, v_n(A)), \tag{5.2}$$

for all $A \in E$, where the $v_i(A)$ are the eigenvalues of A. Of course $H(x_1, \ldots, x_n) = h(\operatorname{diag}(x_1, \ldots, x_n))$ for all $x = (x_1, \ldots, x_n) \in \Gamma_E$, and this shows that $h \in C^r(E)$ (resp. $h \in C^{r,\alpha}(E)$) implies that $H \in C^r(\Gamma_E)$ (resp. $H \in C^{r,\alpha}(\Gamma_E)$). To obtain results in the reverse direction we first note that if $H \in C^r(\Gamma_E)$ then $h \in C^r(E_1)$, where E_1 is the open set consisting of those $A \in E$ whose eigenvalues are all different; this is because the eigenvalues $v_i(A)$ are smooth functions of A in E_1. The subset $M = \{A \in S^{n\times n} : v_i(A) = v_j(A)$ for some $i \neq j\}$ is closed and sparse, since it is the zero set of the discriminant

$$\left[\prod_{1 \leq i < j \leq n} (v_i(A) - v_j(A))\right]^2,$$

which is a symmetric polynomial function of the v_i and is thus expressible as a polynomial in the entries of A. To prove that $h \in C^r(E)$ it therefore suffices by Proposition 2.2 to show that $\lim_{A\to B,\ A\in E_1} D^j h(A)$ exists for all $B \in E \cap M = E \setminus E_1$ whenever $0 \leq |j| \leq r$.

Let $H \in C(\Gamma_E)$. The set of eigenvalues of A varies continuously with A, and since H is symmetric it follows that $h \in C(E)$. The same argument shows more, namely that if a sequence of symmetric functions H_j converges uniformly to H on compact subsets of Γ_E then the corresponding h_j converge uniformly to h on compact subsets of E.

Let $H \in C^1(\Gamma_E)$. For $A \in E_1$ we have that

$$Dh(A) = \sum_{i=1}^{n} H_{,i}(v(A))\, Dv_i(A), \tag{5.3}$$

where $v(A) \stackrel{\text{def}}{=} (v_1(A), \ldots, v_n(A))$. We use the following well-known lemma.

LEMMA 5.1.

$$Dv_i(A) = P_i(A) \tag{5.4}$$

for all $A \in S^{n\times n} \setminus M$ *and* $i = 1, \ldots, n$, *where* $P_i(A)$ *denotes the projection onto the* i^{th} *eigenspace of* A.

Remark 5.2. We regard $P_i(A)$ as an element of $S^{n\times n}$, so that $P_i(A)x = (x, e_i(A))e_i(A)$ for $x \in \mathbb{R}^n$, where $e_i(A)$ denotes the ith unit eigenvector of A and $(\ ,\ )$ the inner product in $\mathbb{R}^n$. Equivalently, $P_i(A) = e_i(A) \otimes e_i(A)$. We identify $\mathscr{L}(S^{n\times n}, \mathbb{R})$ with $S^{n\times n}$, so that $Dv_i(A)$ is the unique $n \times n$ symmetric matrix satisfying

$$\frac{d}{dt} v_i(A + tB)\Big|_{t=0} = \langle Dv_i(A), B\rangle.$$

Proof of Lemma 5.1 (cf. Serrin [1959]). Let $A \in S^{n\times n}\setminus M$, $B \in S^{n\times n}$, $Q \in \mathbf{O}(n)$. If the eigenvalues are ordered by magnitude then

$$\frac{d}{dt} v_i(QAQ^T + tB)\Big|_{t=0} = \frac{d}{dt} v_i(A + tQ^TBQ)\Big|_{t=0}$$

and so $Dv_i(QAQ^T) = QDv_i(A)Q^T$. Also we have that $P_i(QAQ^T) = QP_i(A)Q^T$. It is therefore sufficient to prove (5.4) when $A = \text{diag}(a_1, \ldots, a_n)$ and $a_i \neq a_j$ if $i \neq j$, and we can take $\{e_i(A)\}$ to be the standard basis $\{e_i\}$ of $\mathbf{R}^n$. Fix j and let Q satisfy $Qe_j = -e_j$, $Qe_k = e_k$ for $k \neq j$. Then $QAQ^T = A$ and hence $(QDv_i(A)Q^T)_{jk} = -(Dv_i(A))_{jk} = (Dv_i(A))_{jk}$ when $j \neq k$. Therefore $Dv_i(A)$ is diagonal. But if $B_{kl} = \delta_{jk}\delta_{jl}$ then

$$(Dv_i(A))_{jj} = \frac{d}{dt} v_i(A + tB)\Big|_{t=0}$$

$$= \frac{d}{dt}(a_i + tB_{ii})\Big|_{t=0} = \delta_{ij}.$$

Hence $Dv_i(A) = P_i(A)$ as required. □

It follows from (5.3) and Lemma 5.1 that if $\Lambda = \text{diag}(a) \in S^{n\times n}\backslash M$ then

$$Dh(\Lambda) = \text{diag}(H_{,1}(a), \ldots, H_{,n}(a)). \tag{5.5}$$

Let $D \subset S^{n\times n}$ be open with $\overline{D} \subset E$. Let $\{H_{(j)}\}$ be a sequence of symmetric polynomials converging to H in $C^1(\overline{Z})$ for every open Z with $\overline{Z} \subset \Gamma_E$. (Without the requirement of symmetry the existence of such a sequence $\phi_{(j)}(v_1, \ldots, v_n)$ is standard. A symmetric sequence is obtained by defining

$$H_{(j)}(v_1, \ldots, v_n) = \frac{1}{n!}\sum_{\pi} \phi_{(j)}(v_{\pi(1)}, \ldots, v_{\pi(n)})$$

where the sum is over all permutations π of $(1, \ldots, n)$.) Define $h_{(j)}(A) = H_{(j)}(v_1(A), \ldots, v_n(A))$. Then $h_{(j)}$ is a polynomial. By (5.5)

$$Dh_{(j)}(A) = Q\,\text{diag}(H_{(j),1}(v(A)), \ldots, H_{(j),n}(v(A)))\,Q^T,$$

where $A = Q\,\text{diag}(v(A))Q^T$ and $Q \in \mathbf{SO}(n)$, and hence (since M is nowhere dense)

$$\|Dh_{(j)} - Dh_{(k)}\|_{C(\overline{D};S^{n\times n})} \leqslant C\|H_{(j)} - H_{(k)}\|_{C(\overline{\Gamma}_D)}.$$

But $h_{(j)} \to h$ uniformly on $\overline{D}$. Therefore $h_{(j)}$ is a Cauchy sequence in $C^1(\overline{D})$ and hence $h \in C^1(\overline{D})$. Since D was arbitrary we have shown that $h \in C^1(E)$.

Let $H \in C^2(\Gamma_E)$. For $A \in E_1$ and $B \in S^{n\times n}$ we have that

$$\frac{d^2}{dt^2} h(A + tB)\Big|_{t=0} = \sum_{i=1}^{n} H_{,i}(v(A))\frac{d^2 v_i}{dt^2}(A + tB)\Big|_{t=0}$$

$$+ \sum_{i,j=1}^{n} H_{,ij}(v(A))\frac{dv_i}{dt}(A + tB)\Big|_{t=0}\frac{dv_j}{dt}(A + tB)\Big|_{t=0}. \tag{5.6}$$

LEMMA 5.3. *Let $A = \operatorname{diag}(a_1, \ldots, a_n) \in E_1$, $B \in S^{n\times n}$. Then*

$$\frac{d^2 v_i}{dt^2}(A + tB)\bigg|_{t=0} = 2\sum_{j\neq i} \frac{B_{ij}B_{ij}}{a_i - a_j}. \tag{5.7}$$

Proof. Fix i and A. We have that

$$\frac{d^2 v_i}{dt^2}(A + tB)\bigg|_{t=0} = \sum_{j,k,r,s=1}^{n} a_{jkrs} B_{jk} B_{rs}$$

for certain coefficients a_{jkrs} with $a_{jkrs} = a_{kjrs} = a_{rsjk}$. By choosing Q as in the proof of Lemma 5.1 it is easily shown that $a_{jkrs} = 0$ unless $\{j, k, r, s\}$ consists of two, possibly equal, pairs of integers. One of these pairs is $\{i, i\}$, since if $B_{lm} = 0$ when $l = i$ or $m = i$ then $v_i(A + tB)$ is an eigenvalue of A. Choosing B diagonal shows that $a_{iirr} = 0$, and thus

$$\frac{d^2 v_i}{dt^2}(A + tB)\bigg|_{t=0} = \sum_{j\neq i} c_j B_{ij} B_{ij}$$

for some constants c_j. To compute c_j we can without loss of generality take $n = 2$, $i = 1$, $j = 2$ and $B_{11} = B_{22} = 0$, $B_{12} = B_{21} = 1$. Since $v_1 + v_2 = a_1 + a_2$, $v_1 v_2 = a_1 a_2 - t^2$, we obtain easily that $c_2 = 2/(a_1 - a_2)$, as required. □

Remark 5.4. Formulae (5.4) and (5.7) are well-known results of perturbation theory (see, for example, Kato [1966 Chapter II]), but we have given elementary proofs to make the exposition self-contained.

Substituting (5.4) and (5.7) into (5.6) we deduce that if $A = \operatorname{diag}(a_1, \ldots, a_n) \in E_1$ and $B \in S^{n\times n}$ then

$$\frac{d^2 h}{dt^2}(A + tB)\bigg|_{t=0} = \sum_{i,j=1}^{n} H_{,ij}(a) B_{ii} B_{jj} + 2\sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j} B_{ij} B_{ij}. \tag{5.8}$$

To prove that $h \in C^2(E)$ it suffices to show that, given any $A_0 \in E$, there exists $\epsilon = \epsilon(A_0) > 0$ such that $h \in C^2(\overline{B_\epsilon(A_0)})$, where $B_\epsilon(A_0)$ denotes the open ball in $S^{n\times n}$ with centre A_0 and radius ϵ. Let A_0 have eigenvalues $\alpha_1 \geqslant \cdots \geqslant \alpha_n$. Let N be a closed ball contained in Γ_E with centre $\alpha = (\alpha_1, \ldots, \alpha_n)$ and positive radius. Choose $\epsilon > 0$ small enough so that if $A = Q \operatorname{diag}(a_1, \ldots, a_n) Q^T \in \overline{B_\epsilon(A_0)} \setminus M$ with $Q \in \mathbf{SO}(n)$ and $a_1 \geqslant \cdots \geqslant a_n$ then $a = (a_1, \ldots, a_n) \in N$ and $|a_i - a_j| > \epsilon$ whenever $\alpha_i \neq \alpha_j$. We deduce from (5.8) that for such an A

$$\frac{d^2 h}{dt^2}(A + tB)\bigg|_{t=0} = \sum_{i,j=1}^{n} H_{,ij}(a)(Q^T B Q)_{ii}(Q^T B Q)_{jj}$$
$$+ 2\sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}(Q^T B Q)_{ij}(Q^T B Q)_{ij}$$

for all $B \in S^{n\times n}$. If $\alpha_i \neq \alpha_j$ then

$$\left|\frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}\right| \leqslant C\epsilon \|H\|_{C^1(N)},$$

while if $\alpha_i = \alpha_j$, $i > j$, then

$$\left|\frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}\right|$$

$$= \left|\int_0^1 (H_{,ii} - H_{,ij})(a_1, \ldots, a_{j-1}, ta_j + (1-t)a_i, a_{j+1},\right.$$

$$\left. \times \ldots, a_{i-1}, ta_i + (1-t)a_j, a_{i+1}, \ldots, a_n)\, dt\right|$$

$$\leqslant C\|H\|_{C^2(N)}.$$

Hence

$$\|D^2h(A)(B,B)\| \leqslant C\|H\|_{C^2(N)}\|B\|^2. \tag{5.9}$$

Now let $\{H_{(k)}\}$ be a sequence of symmetric polynomials converging to H in $C^2(N)$. Applying (5.9) to $H_{(k)} - H_{(l)}$ we deduce that the corresponding sequence $h_{(k)}$ is Cauchy in $C^2(\overline{B_\epsilon(A_0)})$, and hence that $h \in C^2(\overline{B_\epsilon(A_0)})$. This completes the proof that $h \in C^2(E)$.

We sum up our results so far.

THEOREM 5.5. *Let $r = 0$, 1 or 2. Then $h \in C^r(E)$ if and only if $H \in C^r(\Gamma_E)$. If $A = Q\,\mathrm{diag}(a_1, \ldots, a_n)Q^T \in E$ with $Q \in \mathbf{SO}(n)$ and $a_i \neq a_j$ for $i \neq j$ then*

$$Dh(A)(B) = \sum_{i=1}^{n} H_{,i}(a)(Q^TBQ^T)_{ii} \tag{5.10}$$

for all $B \in S^{n\times n}$ if $H \in C^1(\Gamma_E)$, and

$$D^2h(A)(B,B) = \sum_{i,j=1}^{n} H_{,ij}(a)(Q^TBQ)_{ii}(Q^TBQ)_{jj}$$

$$+ 2\sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}(Q^TBQ)_{ij}(Q^TBQ)_{ij} \tag{5.11}$$

for all $B \in S^{n\times n}$ if $H \in C^2(\Gamma_E)$.

Remark 5.6. It is natural to conjecture that $h \in C^r(E)$ if and only if $H \in C^r(\Gamma_E)$ for any r, but a proof of this has eluded the author. We have already noted that $h \in C^r(E)$ implies $H \in C^r(\Gamma_E)$ for any r. On the other hand, if $H \in C^r(\Gamma_E)$ then given any $A, B \in S^{n\times n}$ the map $t \mapsto h(A + tB)$ is C^r for sufficiently small $|t|$; this follows from the fact that the eigenvalues $v_i(t)$ of $A + tB$ can be ordered so as to be smooth in t (see Rellich [1969], Kato [1966]), and hence $t \mapsto H(v_1(t), \ldots, v_n(t))$ is C^r. Unfortunately the eigenvalues of a symmetric matrix A cannot in general be ordered so as to be C^r in A, even if $n = 2$.

THEOREM 5.7. *Let $0 < \alpha < 1$, $r = 0, 1, 2, \ldots$. Then $h \in C^{r,\alpha}(E)$ if and only if $H \in C^{r,\alpha}(\Gamma_E)$.*

The proof of Theorem 5.7 follows the idea of the proof of Theorem 4.2, and requires a number of preliminary lemmas.

LEMMA 5.8. *There exists a constant C such that if $A, B \in S^{n\times n}$ have eigenvalues $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$ and $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n$ respectively then*

$$|a - b| \leqslant C\|A - B\|, \tag{5.12}$$

where $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$.

Proof. By approximation it is sufficient to prove (5.12) when the a_i and the b_i are distinct. Consider the matrix $A(t) = tA + (1 - t)B$. The discriminant of $A(t)$ is nonzero for $t = 0, 1$, and therefore vanishes for at most finitely many $t \in (0, 1)$; therefore the eigenvalues $v_1(t) \geqslant v_2(t) \geqslant \cdots \geqslant v_n(t)$ are distinct except for at most finitely many $t \in (0, 1)$. But when the $v_i(t)$ are distinct we have by Lemma 5.1 that $|\dot{v}_i(t)| \leqslant C\|A - B\|$. Therefore

$$|a - b| \leqslant \int_0^1 |\dot{v}(t)|\, dt \leqslant \|A - B\|$$

as required. □

LEMMA 5.9. *Let $H \in C^{0,\alpha}(\Gamma_E)$, $0 < \alpha < 1$. Then $h \in C^{0,\alpha}(E)$.*

Proof. If $\|A\|, \|B\| \leqslant M$ then

$$\|h(A) - h(B)\| = \|H(a) - H(b)\| \leqslant C\|a - b\|^\alpha \leqslant C\|A - B\|^\alpha$$

by Lemma 5.8. □

LEMMA 5.10. *There exists a constant K such that if $A = Q\,\mathrm{diag}(a_1, \ldots, a_n)Q^T \in S^{n\times n}$, where $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$, $Q \in \mathbf{O}(n)$, and $B = \mathrm{diag}(b_1, \ldots, b_n) \in S^{n\times n}$ with $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n$ then*

$$\sum_{i\neq j} |Q_{ij}|\,|a_i - a_j| \leqslant K\|A - B\|.$$

Proof. We have that

$$\sum_{i \neq j} |Q_{ij}|\,|a_i - a_j| = C\|Q \operatorname{diag} a - (\operatorname{diag} a)\, Q\|$$

$$= C\|A - \operatorname{diag} a\|$$

$$\leqslant C[\|A - \operatorname{diag} b\| + \|\operatorname{diag} a - \operatorname{diag} b\|]$$

$$\leqslant K\|A - B\|,$$

where we have used Lemma 5.8. □

LEMMA 5.11. *Let $H \in C^{1,\alpha}(\Gamma_E)$, $0 < \alpha < 1$. Then $h \in C^{1,\alpha}(E)$.*

Proof. Let $A, B \in S^{n\times n}$ with $\|A\|, \|B\| \leqslant M$. Without loss of generality we may suppose that $B = \operatorname{diag}(b_1, \ldots, b_n)$ is diagonal with $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n$, and that $A = Q \operatorname{diag}(a_1, \ldots, a_n) Q^T$ with $Q \in \mathbf{SO}(n)$ and $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$. Then by (5.10) and the Hölder continuity of DH,

$$\|Dh(A) - Dh(B)\|$$

$$= \|Q \operatorname{diag}(H_{,1}(a), \ldots, H_{,n}(a))\, Q^T - \operatorname{diag}(H_{,1}(b), \ldots, H_{,n}(b))\|$$

$$\leqslant \|\operatorname{diag}(H_{,1}(a) - H_{,1}(b), \ldots, H_{,n}(a) - H_{,n}(b))\|$$

$$+ \|Q \operatorname{diag}(H_{,1}(a), \ldots, H_{,n}(a))\, Q^T - \operatorname{diag}(H_{,1}(a), \ldots, H_{,n}(a))\|$$

$$\leqslant C[\|a - b\|^\alpha + \|Q \operatorname{diag}(H_{,1}(a), \ldots, H_{,n}(a))$$

$$- \operatorname{diag}(H_{,1}(a), \ldots, H_{,n}(a)) Q\|]$$

$$= C\left[\|a - b\|^\alpha + \sum_{i \neq j} |Q_{ij}|H_{,i}(a) - H_{,j}(a)|\right]$$

$$\leqslant C\left[\|a - b\|^\alpha + \sum_{i \neq j} |Q_{ij}|\,|a_i - a_j|^\alpha\right].$$

By Lemmas 5.8 and 5.12 we deduce that

$$\|Dh(A) - Dh(B)\| \leqslant C\left[\|a - b\|^\alpha + \left(\sum_{i \neq j} |Q_{ij}|\,|a_i - a_j|\right)^\alpha\right]$$

$$\leqslant C\|A - B\|^\alpha.$$

Together with Lemma 5.9 this proves that $h \in C^{1,\alpha}(E)$. □

Proof of Theorem 5.7. If $h \in C^{r,\alpha}(E)$, then clearly $H \in C^{r,\alpha}(\Gamma_E)$. We prove by induction on r that $H \in C^{r,\alpha}(\Gamma_E)$ implies $h \in C^{r,\alpha}(E)$. The result is true for $r = 0, 1$ by Lemmas 5.9 and 5.11. Let $r \geqslant 2$; suppose the result is true for $r - 1$, and let $H \in C^{r,\alpha}(\Gamma_E)$. By Theorem 5.5, $h \in C^2(E)$ and we have that if $A = Q\,\mathrm{diag}(a_1, \ldots, a_n)Q^T$ with the a_i distinct, and if $B \in S^{n\times n}$, then

$$\frac{d^2}{dt^2} h(A + tB)\bigg|_{t=0} = \sum_{r,s,a,b=1}^{n} A_{rsab} B_{rs} B_{ab}\,. \tag{5.13}$$

where

$$A_{rsab} \stackrel{\text{def}}{=} \sum_{i,j=1}^{n} H_{,ii}(a) Q_{ri} Q_{si} Q_{aj} Q_{bj} + 2\sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j} Q_{ri} Q_{sj} Q_{ai} Q_{bj}\,.$$

Note that

$$\sum_{r,s=1}^{n} A_{rsrs} = \sum_{i=1}^{n} H_{,ii}(a) + 2\sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}\,, \tag{5.14}$$

and

$$\sum_{r,s=1}^{n} A_{rssr} = \sum_{i=1}^{n} H_{,ii}(a). \tag{5.15}$$

We compute the second order partial derivatives of h, regarded as a function of the $n(n-1)/2$ variables A_{ij}, $1 \leqslant j \leqslant i \leqslant n$. By (5.13)

$$\sum_{i=1}^{n} \frac{\partial^2 h(A)}{\partial A_{ii}^2} = \sum_{i=1}^{n} A_{iiii}\,, \tag{5.16}$$

and

$$\begin{aligned} \sum_{i>j} \frac{\partial^2 h(A)}{\partial A_{ij}^2} &= \sum_{i>j} (A_{ijij} + A_{ijji} + A_{jiij} + A_{jiji}) \\ &= \sum_{i\neq j} (A_{ijij} + A_{ijji}) \\ &= \sum_{i,j=1}^{n} (A_{ijij} + A_{ijji}) - 2\sum_{i=1}^{n} A_{iiii}\,. \end{aligned} \tag{5.17}$$

Combining (5.14)–(5.17) we deduce that

$$\sum_{i=1}^{n} \frac{\partial^2 h(A)}{\partial A_{ii}^2} + \frac{1}{2}\sum_{i>j} \frac{\partial^2 h(A)}{\partial A_{ij}^2} = \theta(a_1, \ldots, a_n), \tag{5.18}$$

where

$$\theta(a_1, \ldots, a_n) = \sum_{i=1}^{n} H_{,ii}(a) + \sum_{i>j} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}$$

provided the a_i are distinct.

Clearly θ is symmetric. We claim that $\theta \in C^{r-2,\alpha}(\Gamma_E)$. Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \Gamma_E$ and let N be an open ball in $\mathbf{R}^n$ with centre α and with radius $\rho > 0$ sufficiently small so that $\overline{N} \subset \Gamma_E$ and $\inf\{|a_i - a_j| : a = (a_1, \ldots, a_n) \in \overline{N}, i \text{ and } j \text{ such that } \alpha_i \neq \alpha_j\} > 0$. Then for $a = (a_1, \ldots, a_n) \in N$ we have that

$$\theta(a) = \sum_{i=1}^{n} H_{,ii}(a) + \sum_{\substack{i>j \\ \alpha_i \neq \alpha_j}} \frac{H_{,i}(a) - H_{,j}(a)}{a_i - a_j}$$

$$+ \sum_{\substack{i>j \\ \alpha_i = \alpha_j}} \int_0^1 (H_{,ii} - H_{,ij})(a_1, \ldots, a_{j-1}, ta_j + (1-t)a_i, a_{j+1},$$

$$\ldots, a_{i-1}, ta_i + (1-t)a_j, a_{i+1}, \ldots, a_n)\, dt,$$

and since $H \in C^{r,\alpha}(\Gamma_E)$ it follows easily that $\theta \in C^{r-2,\alpha}(N)$. A simple compactness argument now shows that $\theta \in C^{r-2,\alpha}(\Gamma_E)$. By the induction hypothesis $\theta \in C^{r-2,\alpha}(E)$ as a function of A. By the regularity theory for Poisson's equation we deduce from (5.18) that $h \in C^{r,\alpha}(E)$. This completes the induction. □

6. Applications to nonlinear elasticity. Let $M^{n\times n}$ denote the set of real $n \times n$ matrices, and write $M_+^{n\times n} = \{F \in M^{n\times n} : \det F > 0\}$, $S_+^{n\times n} = \{F \in S^{n\times n} : F > 0\}$.

We are concerned with a homogeneous elastic body having *stored-energy function* $W : D \to \mathbf{R}$, where D is an open subset of $M_+^{n\times n}$ invariant under $\mathbf{SO}(n)$ (that is $QF, FQ \in D$ whenever $F \in D$ and $Q \in \mathbf{SO}(n)$). The assumption of homogeneity is made only for simplicity. The function W is defined with respect to a fixed reference configuration in which the body occupies the bounded open subset $\Omega \in \mathbf{R}^n$. The significance of W is that the total energy stored in a deformation $x : \Omega \to \mathbf{R}^n$ is given by

$$E = \int_\Omega W(\nabla x(X))\, dX. \tag{6.1}$$

The above development is standard; the reader unfamiliar with nonlinear elasticity can consult Truesdell & Noll [1965] for a more complete discussion.

We will apply the results of sections 3 and 5 to relate the differentiability properties of W, Φ, Θ, h, H and $\tilde{H}$. We make use of the following technical lemmas.

LEMMA 6.1. *The mapping* $C \mapsto C^{1/2}$ *of* $S_+^{n\times n}$ *to itself is* C^∞.

Proof. Define $f: S^{n\times n} \times S^{n\times n} \to S^{n\times n}$ by $f(U, C) = U^2 - C$. Clearly f is C^∞ and $f(C^{1/2}, C) = 0$ for any $C \in S_+^{n\times n}$. Also

$$D_U f(U, C)(A) = UA + AU, \qquad A \in S^{n\times n}.$$

Suppose that $U \in S_+^{n\times n}$, so that $U = QGQ^T$ for some $Q \in \mathbf{SO}(n)$ and some $G = \operatorname{diag}(d_1, \ldots, d_n)$, where $d_i > 0$ for $1 \leqslant i \leqslant n$. For $G_1 \in S^{n\times n}$ write $H = Q^T G_1 Q$. Then the equation

$$GW + WG = H$$

has a unique solution $W \in S^{n\times n}$ given by $W_{ij} = H_{ij}/(d_i + d_j)$. Therefore the equation $UA + AU = G_1$ has the unique solution $A = QWQ^T$, and hence $D_U(U, C)$ is an isomorphism. By the inverse function theorem there is a unique C^∞ solution $U(C)$ of

$$f(U, C) = 0, \qquad U(C_1) = C_1^{1/2}$$

in a neighbourhood of C_1 in $S_+^{n\times n}$. The uniqueness of $C^{1/2}$ and the continuity of $U(C)$ show that $U(C) = C^{1/2}$. It follows that $C \mapsto C^{1/2}$ is C^∞ in $S_+^{n\times n}$. □

Remark 6.2. A similar proof is given by Gurtin [1981] and attributed by him to W. Noll.

LEMMA 6.3. *The mapping* $U(F) = \sqrt{F^T F}$ *from* $M_+^{n\times n}$ *to* $S_+^{n\times n}$ *is* C^∞. *If* $F = \operatorname{diag}(a_1, \ldots, a_n)$ *with all* $a_i > 0$ *and if* $G \in M^{n\times n}$ *then the first and second derivatives with respect to* t *of* $U(t) \stackrel{\text{def}}{=} U(F + tG)$ *at* $t = 0$ *are given by*

$$\dot{U}(0)_{ij} = \frac{a_i G_{ij} + a_j G_{ji}}{a_i + a_j} \tag{6.12}$$

and

$$\ddot{U}(0)_{ij} = \frac{2}{a_i + a_j} \sum_{k=1}^{n} \left[G_{ki} G_{kj} - \dot{U}(0)_{ik} \dot{U}(0)_{jk} \right] \tag{6.13}$$

respectively.

Proof. The mapping U is a composition of $F \mapsto F^T F$ and $\sqrt{\ }$, and is therefore C^∞ by Lemma 6.1. Expanding $U(t)$ by Taylor's theorem about $t = 0$ we have that

$$\left(\operatorname{diag} a + t\dot{U}(0) + \frac{t^2}{2}\ddot{U}(0) + o(t^2)\right)^2$$
$$= (\operatorname{diag} a)^2 + t\left(G^T \operatorname{diag} a + (\operatorname{diag} a)G\right) + t^2 G^T G. \tag{6.14}$$

Equating coefficients in (6.14) gives

$$\dot{U}(0)\operatorname{diag} a + (\operatorname{diag} a)\dot{U}(0) = G^T \operatorname{diag} a + (\operatorname{diag} a)G$$

and

$$\tfrac{1}{2}\left[\ddot{U}(0)\operatorname{diag} a + (\operatorname{diag} a)\ddot{U}(0)\right] + \dot{U}(0)^2 = G^T G,$$

and (6.12), (6.13) follow. □

THEOREM 6.4. *Let $W: D \to \mathbb{R}$ be isotropic, and let Φ be given by* (6.7).

(i) *Let $r = 0, 1, 2$ or ∞. Then $W \in C^r(D)$ if and only if $\Phi \in C^r(\Gamma_E)$.*

(ii) *Let $0 < \alpha < 1$, $r = 0, 1, 2, \ldots$. Then $W \in C^{r,\alpha}(D)$ if and only if $\Phi \in C^{r,\alpha}(\Gamma_E)$.*

(iii) *Let $F = \operatorname{diag} v \in D$, where $v = (v_1, \ldots, v_n)$ with all $v_i > 0$, and let $G \in M^{n\times n}$.*

Then if $\Phi \in C^1(\Gamma_E)$

$$D_F W(F)G = \sum_{i=1}^{n} \Phi_{,i}(v) G_{ii}, \tag{6.15}$$

and if $\Phi \in C^2(\Gamma_E)$ then

$$D_F^2 W(F)(G,G) = \sum_{i,j=1}^{n} \Phi_{,ij}(v) G_{ii} G_{jj} + \sum_{i\neq j} \frac{v_i \Phi_{,i}(v) - v_j \Phi_{,j}(v)}{v_i^2 - v_j^2} (G_{ij})^2$$
$$+ \sum_{i\neq j} \frac{v_j \Phi_{,i}(v) - v_i \Phi_{,j}(v)}{v_i^2 - v_j^2} G_{ij} G_{ji}. \tag{6.16}$$

Proof. By Lemma 6.3, $W \in C^r(E)$ (resp. $C^{r,\alpha}(E)$) if and only if $W \in C^r(D)$ (resp. $C^{r,\alpha}(D)$). Thus (i) and (ii) follow from Theorems 5.5 and 5.7. (The case $r = \infty$ in (i) is a consequence of (ii).)

Let $\Phi \in C^1(\Gamma_E)$. Then by (5.10) and (6.12),

$$D_F W(F)G = \frac{d}{dt} W(F + tG)\Big|_{t=0}$$
$$= \sum_{i=1}^{n} \Phi_{,i}(v) \dot{U}_{ii}(0)$$
$$= \sum_{i=1}^{n} \Phi_{,i}(v) G_{ii}.$$

Let $\Phi \in C^2(\Gamma_E)$. Then by (5.11) and (6.13),

$$D_F^2 W(F)(G,G) = \frac{d^2}{dt^2} W(F+tG)\Big|_{t=0}$$

$$= \sum_{i=1}^{n} \Phi_{,i}(v)\ddot{U}_{ii}(O) + D_U^2 W(\operatorname{diag} v)\big(\dot{U}(O), \dot{U}(O)\big)$$

$$= \sum_{i \neq j} \frac{\Phi_{,i}(v)}{v_i}\left[(G_{ji})^2 - \left(\frac{v_i G_{ij} + v_j G_{ji}}{v_i + v_j}\right)^2\right]$$

$$+ \sum_{i,j} \Phi_{,ij}(v) G_{ii} G_{jj} + \sum_{i \neq j}\left(\frac{\Phi_{,i}(v) - \Phi_{,j}(v)}{v_i - v_j}\right)\left(\frac{v_i G_{ij} + v_j G_{ji}}{v_i + v_j}\right)^2,$$

and an easy calculation gives (6.16). □

The formula (6.16) has useful applications to the theory of constitutive inequalities.

THEOREM 6.5. *If $W \in C^2(D)$ is isotropic then W satisfies*

$$D^2 W(F)(G,G) = \sum_{i,j,k,l} \frac{\partial^2 W(F)}{\partial F_{ij}\,\partial F_{kl}} G_{ij} G_{kl} > 0 \tag{6.17}$$

for all $F \in D$ and nonzero $G \in M^{n\times n}$ if and only if Φ given by (6.7) satisfies

$$\sum_{i,j=1}^{n} \Phi_{,ij}(v)\lambda_i\lambda_j > 0 \quad \textit{for all} \quad v \in \Gamma_E \quad \textit{and nonzero} \quad \lambda \in \mathbb{R}^n, \tag{6.18}$$

$$\frac{\Phi_{,i}(v) - \Phi_{,j}(v)}{v_i - v_j} > 0 \quad \textit{for all } i \neq j \textit{ and all } v = (v_1, \ldots, v_n) \in \Gamma_E \textit{ with } v_i \neq v_j, \tag{6.19}$$

and

$$\Phi_{,i}(v) + \Phi_{,j}(v) > 0 \quad \textit{for all} \quad i \neq j \quad \textit{and all} \quad v \in \Gamma_E. \tag{6.20}$$

Remark 6.6. If Γ_E is convex then (6.19) follows from (6.18), since if $i < j$ and $v \in \Gamma_E$ with $v_i \neq v_j$ then (6.18) implies that

$$(\Phi_{,i}(v) - \Phi_{,j}(v))(v_i - v_j) = \frac{1}{2}\sum_{k=1}^{n} (\Phi_{,k}(v) - \Phi_{,k}(\bar{v}))(v_k - \bar{v}_k) > 0,$$

where $\bar{v} \overset{\text{def}}{=} (v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$.

Remark 6.7. Suppose that $D = M_+^{n\times n}$, so that $\Gamma_E = \{a = (a_1, \ldots, a_n) \in \mathbb{R}^n :$

$a_i > 0$ for all $i\}$. Then if (6.18) and (6.20) hold it does not follow that W can be extended to a convex function on $M^{n\times n}$ = convex hull of $M_+^{n\times n}$. For example, let $n = 2$ and

$$\Phi(v_1, v_2) = v_1^2 + v_2^2 + \alpha v_1 v_2 .$$

Then (6.18) and (6.20) hold if and only if $|\alpha| < 2$. However, if $0 > \alpha > -2$ then Φ_1 and Φ_2 can be negative, so that by a result of Hill [1970] (see also Ball [1977]) W has no convex extension.

Proof of Theorem 6.5. By (6.17),

$$D^2W(F)(G,G) = \sum_{i,j=1}^{n} \Phi_{,ij}(v) G_{ii} G_{jj} + \frac{1}{2} \sum_{i\neq j} \left[(\alpha_{ij} + \beta_{ij})(G_{ij})^2 + (\alpha_{ij} - \beta_{ij}) G_{ij} G_{ji} \right], \quad (6.21)$$

where if $v_i \neq v_j$

$$\alpha_{ij} = \frac{\Phi_{,i}(v) - \Phi_{,j}(v)}{v_i - v_j}, \qquad \beta_{ij} = \frac{\Phi_{,i}(v) + \Phi_{,j}(v)}{v_i + v_j}.$$

The first sum in (6.21) is positive definite in the G_{ii} if and only if (6.18) holds. The second sum in (6.21) is positive definite in the G_{ij} $(i \neq j)$ if and only if $\alpha_{ij} > 0$ and $\beta_{ij} > 0$. When $v_i = v_j$ then $\alpha_{ij} = \Phi_{,ii}(v) - \Phi_{,ij}(v)$, which is positive if (6.18) holds as then $\Phi_{,ii}(v) = \Phi_{,jj}(v) > 0$ and $\Phi_{,ii}(v)\Phi_{,jj}(v) > \Phi_{,ij}^2(v)$. The result follows. □

A stored-energy function $W \in C^2(D)$ is said to be *strongly elliptic* if

$$\frac{d^2}{dt^2} W(F + ta \otimes b)\Big|_{t=0} = \sum_{i,j,k,l} \frac{\partial^2 W(F)}{\partial F_{ij}\, \partial F_{kl}} a_i b_j a_k b_l > 0$$

whenever $F \in D$ and $a, b \in \mathbb{R}^n$ are nonzero. Two well-known consequences of strong ellipticity of an isotropic W follow immediately from (6.16). These are the *strengthened tension-extension inequalities*

$$\Phi_{,ii}(v) > 0, \qquad i = 1, \ldots, n, \quad (6.22)$$

and the *Baker–Ericksen inequalities*

$$\frac{v_i \Phi_{,i}(v) - v_j \Phi_{,j}(v)}{v_i - v_j} > 0 \qquad \text{if} \quad v_i \neq v_j . \quad (6.23)$$

The inequalities (6.23), which were first derived from strong ellipticity by Hayes [1969], are in fact consequences of the weaker condition of strict rank 1 convexity (see Ball [1982]).

The necessary and sufficient conditions for strong ellipticity in the case $n = 2$ due to Knowles & Sternberg [1977] also follow easily from (6.16).

For related work on weakly closed sets and rank 1 convexity see Aubert & Tahraoui [1982].

Finally we turn to the representations (6.8)–(6.11). Let $D' = \{FF^T : F \in D\} \subset S_+^{n\times n}$. Then $\Gamma_{D'} = \{(v_1^2, \ldots, v_n^2) : v = (v_1, \ldots, v_n) \in \Gamma_E\}$.

THEOREM 6.8. *Let $W: D \to \mathbb{R}$ be isotropic, and let h, H be given by* (6.9), (6.10).

(i) *Let $r = 0, 1, 2$ or ∞. Then $W \in C^r(D)$ if and only if $h \in C^r(D')$ and if and only if $H \in C^r(\Gamma_{D'})$.*

(ii) *Let $0 < \alpha < 1$, $r = 0, 1, 2, \ldots$. Then $W \in C^{r,\alpha}(D)$ if and only if $h \in C^{r,\alpha}(D')$ and if and only if $H \in C^{r,\alpha}(\Gamma_{D'})$.*

Proof. This follows immediately from Theorems 5.5, 5.7, and 6.4 and the fact that the map $(v_1, \ldots, v_n) \mapsto (v_1^2, \ldots, v_n^2)$ from Γ_E to $\Gamma_{D'}$ is a smooth diffeomorphism. □

In contrast to Φ, h and H the functions Θ and $\tilde{H}$ given by (6.8) and (6.11) are in general less differentiable than W.

THEOREM 6.9. *Let $W: D \to \mathbb{R}$ be isotropic, and let Θ and $\tilde{H}$ be given by* (6.9) *and* (6.11) *respectively. Let $r = 0, 1, 2, \ldots$. If $W \in C^{nr}(\bar{D})$ and Γ_E is convex then $\Theta \in C^r(\bar{\Omega})$ and $\tilde{H} \in C^r(\bar{\Omega}')$, where $\Omega \overset{\text{def}}{=} \Omega_{\Gamma_E}$, $\Omega' \overset{\text{def}}{=} \Omega_{\Gamma_{D'}}$.*

Remark 6.10. Theorem 6.9 is optimal. Let I be an open interval with $\bar{I} \subset (0, \infty)$, let $D = \{F \in M_+^{n\times n} :$ each principal stretch $v_i \in I\}$ and suppose that

$$\Phi(v_1, \ldots, v_n) = \sum_{i=1}^{n} \phi(v_i),$$

where $\phi: I \to \mathbb{R}$. Then $\Theta \in C^{r+1}(\bar{\Omega})$ (equivalently, $\tilde{H} \in C^{r+1}(\bar{\Omega}')$) if and only if $\phi \in C^{n(r+1)}(\bar{I})$ by Theorem 3.8. But then $W \in C^{n(r+1)-1}(D)$ by Theorem 5.7 and Lemma 6.3.

Proof of Theorem 6.9. If $W \in C^{nr}(\bar{D})$ then $\Phi \in C^{nr}(\bar{\Gamma}_E)$, and hence $H \in C^{nr}(\bar{\Gamma}_{D'})$. The result follows from Theorem 3.2. □

Acknowledgement. I would like to thank Gérard Barbançon and Jack Carr for many useful discussions about this work.

REFERENCES

1. V. I. ARNOLD, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York-Heidelberg-Berlin (1983).
2. A. AUBERT AND R. TAHRAOUI, *Conditions necessaires de faible fermeture et de* 1*-rang convexité en dimension* 3, (1982), preprint.
3. J. M. BALL, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Rat. Mech.

Anal. **63** (1977), 337–403. *Discontinuous equilibrium solutions and cavitation in nonlinear elasticity*, Phil. Trans. Roy. Soc. London A **306** (1982), 557–611.

4. G. Barbançon, *A propos du théorème de Newton pour les fonctions de classe C^m et d'une generalization de la notion de multiplicateur rugueux*, Ann. Fac. Sci. Phnom Penh (1969). *Théorème de Newton pour les fonctions de classe C^r*, Ann. Scient. Ec. Norm. Sup. 4ème Série **5** (1972), 435–458.
5. G. Barbançon and M. Rais, *Sur le théorème de Hilbert différentiable pour les groupes linéaires finis* (d'après E. Noether), (1982), preprint.
6. J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York (1960).
7. H. Federer, *Geometric Measure Theory*, Springer-Verlag, Berlin (1969).
8. L. E. Fraenkel, *On the embedding of $C^1(\bar{\Omega})$ in $C^{0,\alpha}(\bar{\Omega})$*, J. London Math. Soc. (2), **26** (1982), 290–298.
9. F. R. Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, New York (1960).
10. D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin (1977).
11. G. Glaeser, *Fonctions composées différentiables*, Annals of Mathematics **77** (1963), 193–209.
12. M. E. Gurtin, *An Introduction to Continuum Mechanics*, Academic Press, New York (1981).
13. M. Hayes, *Static implications of the strong-ellipticity condition*, Arch. Rat. Mech. Anal. **33** (1969), 181–191.
14. R. Hill, *Constitutive inequalities for isotropic elastic solids under finite strain*, Proc. Roy. Soc. London A **314**, (1970), 457–472.
15. T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin (1966).
16. J. K. Knowles and E. Sternberg, *On the failure of ellipticity of the equations for finite elastic plane strain*, Arch. Rat. Mech. Anal. **63** (1977), 321–326.
17. F. Rellich, *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach, New York (1969).
18. J. Serrin, *The derivation of stress-deformation relations for a Stokesian fluid*, J. Math. Mech. **8**, 459–469.
19. C. Truesdell and W. Noll, *The nonlinear field theories of mechanics*, Handbuch der Physik Vol. III/3, ed. S. Flügge, Springer, Berlin (1965).
20. H. D. Ursell, *Inequalities between sums of powers*, Proc. London Math. Soc. **9** (1959), 432–450.
21. H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. **36** (1934a), 63–89. *Functions differentiable on the boundaries of regions*, Ann. of Math. **55** (1934b). *Differentiable even functions*, Duke Math. J. **10** (1943), 159–160.

Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, Scotland