Course MA2321: Michaelmas Term 2015.

Worked Solutions to Assignment II.

1. *Let l, m and n be positive integers, let let D and E be open sets in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, let $\mathbf{p}$ and $\mathbf{q}$ be limit points of the sets D and E respectively, where in $\mathbf{p} \in \mathbb{R}^n$ and $\mathbf{q} \in \mathbb{R}^m$, let $\mathbf{r} \in \mathbb{R}^l$, let $\varphi: D \to \mathbb{R}^m$ be a function from D to $\mathbb{R}^m$ with the property that $\varphi(D) \subset E$, let $\psi: E \to \mathbb{R}^l$ be a function from E to $\mathbb{R}^l$. Suppose that the following three conditions are satisfied:*

   (i) $\lim\limits_{\mathbf{x}\to\mathbf{p}} \varphi(\mathbf{x}) = \mathbf{q}$;

   (ii) $\lim\limits_{\mathbf{y}\to\mathbf{q}} \psi(\mathbf{y}) = \mathbf{r}$;

   (iii) there exists some positive real number δ_0 such that $\varphi(\mathbf{x}) \neq \mathbf{q}$ for all $\mathbf{x} \in D$ satisfying $|\mathbf{x} - \mathbf{p}| < \delta_0$.

   Prove that $\lim\limits_{\mathbf{x}\to\mathbf{p}} \psi(\varphi(\mathbf{x})) = \mathbf{r}$.

   First Solution. Let some positive real number ε be given. Then there exists some positive real number η such that $|\psi(\mathbf{y}) - \mathbf{r}| < \varepsilon$ whenever $\mathbf{y} \in E$ satisfies $0 < |\mathbf{y} - \mathbf{q}| < \eta$. There then exists some positive real number δ_1 such that $|\varphi(\mathbf{x}) - \mathbf{q}| < \eta$ whenever $\mathbf{x} \in D$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$. Also there exists some positive real number δ_0 such that $\varphi(\mathbf{x}) \neq \mathbf{q}$ whenever $\mathbf{x} \in D$ satisfies $|\mathbf{x} - \mathbf{p}| < \delta_0$. Let δ be the minimum of δ_0 and δ_1. Then $\delta > 0$, and $0 < |\varphi(\mathbf{x}) - \mathbf{q}| < \eta$ whenever $\mathbf{x} \in D$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta$. But this then ensures that $|\psi(\varphi(\mathbf{x})) - \mathbf{r}| < \varepsilon$ whenever $\mathbf{x} \in D$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta$. The result follows.

   Second Solution. It follows from the definitions of limits that, given any postive real number ε_1, there exists some positive real number δ_1 such that $|\varphi(\mathbf{x}) - \mathbf{q}| < \varepsilon_1$ whenever $\mathbf{x} \in D$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$. Also, given any postive real number ε, there exists some positive real number δ_2 such that $|\psi(\mathbf{y}) - \mathbf{r}| < \varepsilon$ whenever $\mathbf{y} \in D$ satisfies $0 < |\mathbf{y} - \mathbf{q}| < \delta_2$. Let $\varepsilon_1 = \delta_2$. Then $|\psi(\varphi(\mathbf{x})) - \mathbf{r}| < \varepsilon$ whenever $\mathbf{x} \in D$ satisfies both $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$ and. $\varphi(\mathbf{x}) \neq \mathbf{q}$. Also there exists some positive real number δ_0 such that $\varphi(\mathbf{x}) \neq \mathbf{q}$ whenever $\mathbf{x} \in D$ satisfies $|\mathbf{x} - \mathbf{p}| < \delta_0$. Let δ be the minimum of δ_0 and δ_1. Then $\delta > 0$, and if $\mathbf{x} \in D$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta$ then both $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$ and also $\varphi(\mathbf{x}) \neq \mathbf{q}$, and therefore $|\psi(\varphi(\mathbf{x})) - \mathbf{r}| < \varepsilon$. The result follows.

Further Comments

A. There were those who used an incorrect definition of limits, supposing for example that $\lim_{\mathbf{y}\to\mathbf{q}} \psi(\mathbf{y}) = \mathbf{r}$ if and only if, given any positive real number ε, there exists some positive real number δ_2 such that $|\psi(\mathbf{y}) - \mathbf{r}| < \varepsilon$ whenever $\mathbf{y} \in D$ satisfies $|\mathbf{y} - \mathbf{q}| < \delta_2$. Of course this is not the definition of limits of functions given in the lecture notes, and Question 1 on first assignment for MA2321 for Michaelmas Term 2015 should in particular have alerted those in the class to the fact that the definition of limits as $\mathbf{y} \to \mathbf{q}$ only controls the behaviour of $\psi(\mathbf{y})$ for values of $\mathbf{y}$ that lie sufficiently close to $\mathbf{q}$ but are not equal to $\mathbf{q}$.

B. Note that the fact that $\lim_{\mathbf{x}\to\mathbf{p}} \varphi(\mathbf{x}) = \mathbf{q}$ by itself is consistent with a scenario in which there exists an infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots$ of points of D converging to $\mathbf{p}$ with the property that $\varphi(\mathbf{x}_j) = \mathbf{q}$ for all positive integers j. It would then be the case that $\psi(\varphi(\mathbf{x}_j)) = \psi(\mathbf{q})$ for all positive integers j, and there is nothing in the requrement that $\lim_{\mathbf{y}\to\mathbf{q}} \psi(\mathbf{y}) = \mathbf{r}$ to control the value of $\psi(\mathbf{q})$. Thus it might well be the case that $\psi(\mathbf{q}) \neq \mathbf{r}$. And in this scenario it would not be the case that $\lim_{\mathbf{x}\to\mathbf{p}} \psi(\varphi(\mathbf{x})) = \mathbf{r}$. This is in fact the precise scenario that was exhibited in Question 1 of the first assignment for MA2321 for Michaelmas Term 2015. Thus all three conditions (i), (ii) and (iii) are required in order to ensure that $\lim_{\mathbf{x}\to\mathbf{p}} \psi(\varphi(\mathbf{x})) = \mathbf{r}$, and a solution to the problem should make it clear how all three conditions contribute to the result.

C. Given that conditions (i) and (iii) are both required for the result to hold, the manner in which these conditions are applied simultaneously should be made clear in the solution, for example by explicitly determining $\delta > 0$ as the minimum of positive real numbers δ_1 and δ_0, where δ_1 comes from an application of (i) and δ_0 comes from an application of (iii). This should make clear and explicit an understanding that the condition of the form $0 < |\varphi(\mathbf{x}) - \mathbf{q}| < \eta$ is required to ensure that $|\psi(\varphi(\mathbf{x})) - \mathbf{r}| < \varepsilon$, and that this condition is ensured by ensuring, through the application of distinct conditions (i) and (iii), that both $|\varphi(\mathbf{x}) - \mathbf{q}| < \eta$ and $\varphi(\mathbf{x}) \neq \mathbf{q}$.

2. *Let $\varphi: V \to \mathbb{R}^m$ be a function mapping some open subset V of $\mathbb{R}^n$ into $\mathbb{R}^m$, let $h: V \to \mathbb{R}$ be a real-valued function on V, and let $\mathbf{p}$ be a limit*

point of V. Suppose that $\lim_{\mathbf{x}\to\mathbf{p}} \varphi(\mathbf{x}) = 0$. Suppose also that there exist positive constants C and δ_0 such that $|h(\mathbf{x})| \leq C$ for all $\mathbf{x} \in V$ satisfying $|\mathbf{x}-\mathbf{p}| < \delta_0$. Using only the definition of limits (i.e., the "epsilon-delta" definition"), without citing to any other lemmas, propositions, theorems etc. concerning limits, prove that $\lim_{\mathbf{x}\to\mathbf{p}} (h(\mathbf{x})\varphi(\mathbf{x})) = 0$.

Solution. It follows from the definition of limits that, given any positive number ε_1, there exists some positive real number δ_1 such that $|\varphi(\mathbf{x}| < \varepsilon_1$ whenever $\mathbf{x} \in V$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$. Let some positive real number ε be given, and let $\varepsilon_1 = \varepsilon/C$. Then there exists some positive real number δ_1 such that $|\varphi(\mathbf{x}| < \varepsilon/C$ whenever $\mathbf{x} \in V$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta_1$. Also there exists some positive real number δ_0 such that $|h(x)| \leq C$ whenever whenever $\mathbf{x} \in V$ satisfies $|\mathbf{x} - \mathbf{p}| < \delta_0$. Let δ be the minimum of δ_0 and δ_1. Then $\delta > 0$, and if $\mathbf{x} \in V$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta$ then $|h(\mathbf{x})| \leq C$ and $|\varphi(\mathbf{x}| < \varepsilon/C$, and therefore

$$|h(\mathbf{x})\varphi(\mathbf{x})| < C \times \frac{\varepsilon}{C} = \varepsilon.$$

Further Comments

A. Note that the value of the positive number ε is given. You have no control over it. All you know about it is that it is a positive real number. You then have to determine some positive real number δ such that $|h(\mathbf{x})\varphi(\mathbf{x})| < \varepsilon$ whenever $\mathbf{x} \in V$ satisfies $0 < |\mathbf{x} - \mathbf{p}| < \delta$. This is achieved by showing the existence of some positive number δ_0 such that $|h(\mathbf{x}| \leq C$ whenever $0 < |\mathbf{x} - \mathbf{p}| < \delta_0$, using the definition of limits to show the existence of some positive number δ_1 such that $|\varphi(\mathbf{x})| < \varepsilon_1$ whenever $0 < |\mathbf{x}-\mathbf{p}| < \delta_1$, where ε_1 has been determined so that $C\varepsilon_1 \leq \varepsilon$. You can then take δ to be the minimum of δ_0 and δ_1. But you should note that the value of ε determines the value of ε_1, and not vice versa.

B. Questions 1 and 2 on this assignment were formulated as a result of writing up (in Section 8 of the MA2321 notes, not examinable at the Annual and Supplemental Examinations 2016) a proof of the Inverse Function Theorem. The Inverse Function Theorem was examined at the Annual Examination in module MA2321 in 2015, 2014, 2013, 2012, 2011 and 2010. The intention had been to conclude MA2321 with a proof of this theorem in Michaelmas Term 2015, and a sketch of the proof was provided in the final day of lectures. However the proof of the full theorem (including the smoothness of the inverse function), written out in full, with all details fully worked-out to the extent that they are in the current MA2321 notes, would have proved too long for an examination question. In writing up the proof of the Inverse Function Theorem, a section of the proof, verifying that the inverse function is differentiable, was extracted from the main proof, and presented as a separate lemma, numbered Lemma 8.10 in the extended notes. (This portion of the proof of the Inverse Function Theorem corresponds to about eight lines of the proof taught in previous years.) Nevertheless, although the proof of Lemma 8.10 might seem to be more detailed, there are a couple of principles used in it without further explanation that correspond to question 1 (about two-thirds of the way through the proof of Lemma 8.10) and question 2 (in the final stages of the proof of Lemma 8.10). The proof of Lemma 8.10 is written on the assumption that those engaging with it would be sufficiently experienced with ε–δ arguments that they could be expected to fill in any details for themselves. But the results required were extracted and included as the first two questions of this assignment.

3. *This question concerns real-valued functions* $f\colon D \to \mathbb{R}$, *where either* $D = \mathbb{R}^n$ *or else* $D = \mathbb{R}^n \setminus \{\mathbf{0}\}$, *and where the function satisfies an identity, determined by a real constant* s, *that requires that*

$$f(t\mathbf{x}) = t^s f(\mathbf{x})$$

for all $\mathbf{x} \in D$ *and for all positive real numbers* t. *Functions satisfying this condition as said to be* homogeneous of degree s.

We use the notation $\partial_j f$ *to denote the partial derivative of* f *with respect to the* j*th coordinate on* $\mathbb{R}^n$, *so that* $(\partial_j f)(\mathbf{p})$ *denotes the value of partial derivative* $\dfrac{\partial f}{\partial x_j}$ *at the point* $\mathbf{p}$ *for* $\mathbf{p} \in D$.

Moreover, given $\mathbf{v} \in \mathbb{R}^n$, *where* $(v_1, v_2, \ldots v_n)$, *the value* $(Df)_{\mathbf{x}}\mathbf{v}$ *resulting from the action of the (total) derivative* $(Df)_{\mathbf{x}}$ *on the vector* $\mathbf{v}$ *is expressed in terms of the partial derivatives of* f *according to the following formula:*

$$(Df)_{\mathbf{x}}\mathbf{v} = \sum_{j=1}^{n} v_j (\partial_j f)(\mathbf{x}).$$

Throughout this question, let $f\colon D \to \mathbb{R}$ be a differentiable real-valued function defined on D, where either $D = \mathbb{R}^n$ or $D = \mathbb{R}^n \setminus \{\mathbf{0}\}$ that satisfies the identity $f(t\mathbf{x}) = t^s f(\mathbf{x})$ for all $\mathbf{x} \in D$ and for all positive real numbers t.

(a) *Suppose that the function f is differentiable. For each positive real number t, let $\lambda_t\colon \mathbb{R}^n \to \mathbb{R}^n$ be the function defined such that $\lambda_t(\mathbf{x}) = t\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. The Chain Rule guarantees that*

$$D(f \circ \lambda_t)_{\mathbf{x}} = (Df)_{\lambda_t(\mathbf{x})}\,(D\lambda_t)_{\mathbf{x}}.$$

The function $\lambda_t\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, and the derivative of any linear transformation is that linear transformation itself (see Lemma 7.4). Explain briefly why

$$(\partial_j f)(t\mathbf{x}) = t^{s-1}(\partial_j f)(\mathbf{x})$$

for $j = 1, 2, \ldots, n$, for all $\mathbf{x} \in D$ and for all positive real numbers t.

First Solution. Let $\mathbf{v} \in \mathbb{R}^n$. Then

$$(D\lambda_t)_{\mathbf{x}}\mathbf{v} = \lambda_t(\mathbf{v}) = t\mathbf{v},$$

and therefore

$$\begin{aligned} D(f \circ \lambda_t)_{\mathbf{x}}\mathbf{v} &= (Df)_{\lambda_t(\mathbf{x})}(D\lambda_t)_{\mathbf{x}}\mathbf{v} = (Df)_{\lambda_t(\mathbf{x})}(t\mathbf{v}) \\ &= t(Df)_{\lambda_t(\mathbf{x})}\mathbf{v} = t\sum_{j=1}^{n} v_j(\partial_j f)_{\lambda_t(\mathbf{x})} \end{aligned}$$

for all $\mathbf{v} \in \mathbb{R}^n$. But $f(\lambda_t(\mathbf{x})) = t^s\mathbf{x}$ for all $\mathbf{x} \in D$ and $t > 0$. It follows that

$$D(f \circ \lambda_t)_{\mathbf{x}}\mathbf{v} = t^s(Df)_{\mathbf{x}}\mathbf{v} = t^s\sum_{j=1}^{n} v_j(\partial_j f)_{\mathbf{x}}$$

for all $\mathbf{v} \in \mathbb{R}^n$ where $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. It follows that $t(\partial_j f)_{\lambda_t(\mathbf{x})} = t^s(\partial_j f)_{\mathbf{x}}$, and thus $(\partial_j f)_{\lambda_t(\mathbf{x})} = t^{s-1}(\partial_j f)_{\mathbf{x}}$ for $j = 1, 2, \ldots, n$, as required.

Second Solution. We apply a more traditional version of the Chain Rule in partial derivatives (as it would be understood by mathematicians of the 18th and early 19th centuries). Let the positive real number t be regarded as a fixed constant, and let real variables $u_1, u_2, \ldots, u_n$ be defined so that $u_j = tx_j$ for $j = 1, 2, \ldots, n$, so that $u_1, u_2, \ldots, u_n$ are dependent variables that depend on the independent variables $x_1, x_2, \ldots, x_n$. Then

$$\frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial f(u_1, u_2, \ldots, u_n)}{\partial u_k} \frac{\partial u_k}{\partial x_j}.$$

But

$$\frac{\partial u_k}{\partial x_j} = \begin{cases} t & \text{if } k = j; \\ 0 & \text{if } k \neq j. \end{cases}$$

It follows that

$$\begin{aligned} \frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_j} &= t\frac{\partial f(u_1, u_2, \ldots, u_n)}{\partial u_j} = t(\partial_j f)(\mathbf{u}) \\ &= t(\partial_j f)(t\mathbf{x}) \end{aligned}$$

for $j = 1, 2, \ldots, n$. But

$$\begin{aligned} \frac{\partial f(tx_1, tx_2, \ldots, tx_n)}{\partial x_j} &= \frac{\partial (t^s f(x_1, x_2, \ldots, x_n))}{\partial x_j} \\ &= t^s \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_j} \\ &= t^s (\partial_j f)(\mathbf{x}). \end{aligned}$$

Thus $(\partial_j f)(t\mathbf{x}) = t^{s-1}(\partial_j f)(\mathbf{x})$, as required.

(b) *Again suppose that the function f is differentiable. Let $P = \{t \in \mathbb{R} : t > 0\}$, and let $\mu_{\mathbf{x}} \colon P \to \mathbb{R}^n$ be defined such that $\mu_{\mathbf{x}}(t) = t\mathbf{x}$ for all $t \in P$ and $\mathbf{x} \in D$. The Chain Rule (Theorem 7.9) guarantees that*

$$D(f \circ \mu_{\mathbf{x}})_t = (Df)_{\mu_{\mathbf{x}}(t)} \, (D\mu_{\mathbf{x}})_t,$$

where

$$D(f \circ \mu_{\mathbf{x}})_t = \frac{d}{dt} \left((f \circ \mu_{\mathbf{x}})(t) \right) = \frac{d}{dt} \left(f(t\mathbf{x}) \right).$$

Use this result to prove that

$$\sum_{j=1}^{n} x_j (\partial_j f)(\mathbf{x}) = s f(\mathbf{x}),$$

or, in traditional notation,

$$\sum_{j=1}^{n} x_j \frac{\partial f}{\partial x_j} = sf.$$

First Solution. For this part, $\mathbf{x}$ is to be considered fixed, and we examine the dependence of $f(t\mathbf{x})$ as t varies. The linear transformation $\mu_{\mathbf{x}}: \mathbb{R} \to \mathbb{R}^n$ maps $\mathbb{R}$ into $\mathbb{R}^n$, and the same is true of its derivative: for any $t > 0$, the linear transformation $(D\mu_{\mathbf{x}})_t$ maps a real number v to an n-dimensional vector $(D\mu_{\mathbf{x}})_t v$. Moreover $\mu_{\mathbf{x}}$ is a linear transformation, and therefore its derivative at t is equal to $\mu_{\mathbf{x}}$ itself. Thus

$$(D\mu_{\mathbf{x}})_t v = \mu_{\mathbf{x}}(v) = v\mathbf{x}$$

for all real numbers $\mathbf{v}$. It follows that

$$\begin{aligned} D(f \circ \mu_{\mathbf{x}})_t v &= (Df)_{\mu_{\mathbf{x}}(t)} \, (D\mu_{\mathbf{x}})_t v = (Df)_{\mu_{\mathbf{x}}(t)}(v\mathbf{x}) \\ &= v(Df)_{\mu_{\mathbf{x}}(t)}\mathbf{x} = v \sum_{j=1}^{n} x_j (\partial_j f)(t\mathbf{x}). \end{aligned}$$

But

$$D(f \circ \mu_{\mathbf{x}})_t v = \frac{df(t\mathbf{x})}{dt} \, v = \frac{d}{dt} \left(t^s f(\mathbf{x}) \right) \, v = st^{s-1} f(\mathbf{x}) v.$$

Combining the above identities, we find that

$$\sum_{j=1}^{n} x_j (\partial_j f)(t\mathbf{x}) = st^{s-1} f(\mathbf{x}).$$

Thus if we set $t = 1$, we obtain the identity

$$\sum_{j=1}^{n} x_j (\partial_j f)(\mathbf{x}) = sf(\mathbf{x}).$$

Second Solution. Again we apply a traditional version of the Chain Rule in partial derivatives, regarding the components $x_1, x_2, \ldots, x_n$ of the vector $\mathbf{x}$ as fixed quantities. Let real variables $u_1, u_2, \ldots, u_n$ be defined so that $u_j = tx_j$ for $j = 1, 2, \ldots, n$. We treat $u_1, u_2, \ldots, u_n$ as dependent variables that depend on the independent variable t. Then

$$\begin{aligned} \frac{df(tx_1, tx_2, \ldots, tx_n)}{dt} &= \sum_{k=1}^{n} \frac{\partial f(u_1, u_2, \ldots, u_n)}{\partial u_k} \frac{du_k}{dt} \\ &= \sum_{k=1}^{n} \frac{\partial f(u_1, u_2, \ldots, u_n)}{\partial u_k} x_j \\ &= \sum_{k=1}^{n} x_j (\partial_j f)(t\mathbf{x}). \end{aligned}$$

But

$$\begin{aligned} \frac{df(tx_1, tx_2, \ldots, tx_n)}{dt} &= \frac{\partial (t^s f(x_1, x_2, \ldots, x_n))}{\partial x_j} \\ &= \frac{d}{dt}(t^s) f(x_1, x_2, \ldots, x_n) \\ &= st^{s-1} f(x_1, x_2, \ldots, x_n), \end{aligned}$$

and therefore

$$\sum_{k=1}^{n} x_j (\partial_j f)(t\mathbf{x}) = st^{s-1} f(x_1, x_2, \ldots, x_n).$$

On setting $t = 1$, we find that

$$\sum_{k=1}^{n} x_j (\partial_j f)(\mathbf{x}) = sf(x_1, x_2, \ldots, x_n),$$

as required.

Comment on Notation. The mathematics of the 20th and 21st centuries typically employs an extraordinary wide range of notations. This is particularly the case in areas such as differential geometry that are fundamental to areas of theoretical physics like general relativity and string theory. Those seriously engaging with mathematics and theoretical physics therefore need to be flexible and adaptable in adjusting to new notations. The Chain Rule for differentiating compositions of functions of several variables may be stated in 19th century notation (as it is to be found in calculus textbooks). But to progress further, one must engage with the more abstract account of differentiability in which the derivative of a smooth function $\varphi: V \to \mathbb{R}^m$ at a point $\mathbf{x}$ of an open set V in $\mathbb{R}^n$ is a linear transformation $(D\varphi)_{\mathbf{x}}: \mathbb{R}^n \to \mathbb{R}^m$. This linear transformation provides the best linear approximation to the behaviour of the function φ around the point $\mathbf{x}$. The Chain Rule in this context then states that the derivative of a composition of differentiable functions is the composition of the derivatives of those functions. Where those derivatives are represented, with respect to appropriate bases, by matrices, then the derivative of the composition is evaluated by multiplying together the matrices representing those derivatives with respect to the relevant bases. This leads on to the notion of the derivative of a smooth mapping between smooth manifolds as being a linear transformation between tangent spaces to those smooth manifolds. In addition to the MA2321 notes, the basic theory is presented in Chapter 2 of the textbook *Analysis on Manifolds* by James R. Munkres recommended for the MA2322 Calculus on Manifolds course. (Note that Munkres adopts the notation $(D\varphi)(\mathbf{x})$, not $(D\varphi)_{\mathbf{x}}$, for the derivative of φ at a point $\mathbf{x}$ of the domain of the function φ.) The final chapter (Chapter 9) of *Analysis on Manifolds* proceeds further in the direction one must travel to arrive at differential geometry and its applications to areas of theoretical physics such as general relativity and string theory. Ideally the approaches adopted in both the first and second solutions to parts (a) and (b) should be understood, as should the correspondences between them.

4. *Throughout this question, let s be a real number, and let $f: \mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}$ be a continuous (but not necessarily differentiable) function that also satisfies the identity $f(t\mathbf{x}) = t^s f(\mathbf{x})$ for all $\mathbf{x} \in D$ and for all positive real numbers t, where s is a real constant. Let S^{n-1} denote the unit sphere in $\mathbb{R}^n$, defined so that*

$$S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| = 1\}.$$

The set S^{n-1} is a closed bounded subset of Euclidean space. The Extreme Value Theorem (Theorem 6.21) therefore guarantees the existence of points $\mathbf{u}$ and $\mathbf{v}$ on the unit sphere S^{n-1} such that $f(\mathbf{u}) \leq f(\mathbf{x}) \leq f(\mathbf{v})$ for all $\mathbf{x} \in S^{n-1}$.

(a) *Let $f: \mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}$ satisfy the conditions stated above. In the case where $s > 0$ prove that $\lim_{\mathbf{x} \to \mathbf{0}} f(\mathbf{x}) = 0$.*

Also, in the case where $s < 0$, prove that if the function f is non-zero at any point of $\mathbb{R}^n \setminus \{\mathbf{0}\}$ then a value $f(0,0)$ for f at the origin $(0,0)$ cannot be determined that will make the function continuous at $(0,0)$.

Solution. We consider first the case $s > 0$. In this case the function $f\colon \mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}$ is continuous, but not necessarily differentiable, on $\mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}$ and satisfies the condition $f(t\mathbf{x}) = t^s f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and for all positive real numbers t.
The restriction of the function f to the unit sphere S^{n-1} is continuous, and therefore there exist $\mathbf{u}, \mathbf{v} \in S^{n-1}$ such that $f(\mathbf{u}) \leq f(\mathbf{x}) \leq f(\mathbf{v})$ for all $\mathbf{x} \in S^{n-1}$. Let K be a positive real number for which $|f(\mathbf{u})| \leq K$ and $|f(\mathbf{v})| \leq K$. Then $|f(\mathbf{x})| \leq K$ whenever $|\mathbf{x}| = 1$. It then follows from the homogeneity of the function f that $|f(\mathbf{x})| \leq K|\mathbf{x}|^s$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Given a positive real number ε, there exists a positive real number δ such that $K\delta^s \leq \varepsilon$. (Indeed we can take $\delta = \sqrt[s]{\varepsilon/K}$.) If $0 \leq |\mathbf{x}| < \delta$ then $|f(\mathbf{x})| < K\delta^s \leq \varepsilon$. Thus $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = 0$.

The case when $s < 0$ is more straightforward. Suppose that $f(\mathbf{z}) > 0$ for some $\mathbf{z} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Then $f(2^{-j}\mathbf{z}) \to +\infty$ as $j \to +\infty$. Alternatively if $f(\mathbf{z}) < 0$ then $f(2^{-j}\mathbf{z}) \to -\infty$ as $j \to +\infty$. It follows that if $f(\mathbf{z}) \neq 0$ for at least one $\mathbf{z} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ then $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x})$ cannot exist.

Comment. Many attempted to prove the result in the case where $s > 0$ as follows. It is clear that

$$\lim_{t\to 0} f(t\mathbf{x}) = \lim_{t\to 0} t^s f(\mathbf{x}) = 0$$

for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. It was claimed that

$$\text{“}\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = \lim_{t\to 0} f(t\mathbf{x})\text{”}.$$

It was then claimed that, because the limit on the right hand side of this purported identity exists for all non-zero $\mathbf{x}$, the limit on the left hand side must also exist. This argument is fallacious. If the limit on the left hand side of the identity exists, then the limit on the right hand side exists and equals that on the left hand side. But the converse result is in general false. The existence of the limit on the right hand side for all non-zero $\mathbf{x} \in \mathbb{R}^n$ does not ensure the existence of the limit on the left hand side. One counterexample is provided by Question 2 on the first assignment for MA2321 in Michaelmas Term 2015, where the function in question was defined such that

$$f(x,y) = \begin{cases} \dfrac{2x^2y^3}{x^4+y^6} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

The restriction to this function to any line passing through the origin is differentiable and thus continuous. (Indeed, more generally, the restriction of this function to any line in $\mathbb{R}^2$ is continuous on that line.) But the function takes the value 1 at all non-zero points of the curve $x^2 = y^3$ and takes the value -1 at all non-zero points of the curve $x^2 = -y^3$. The curves $x^2 = y^3$ and $x^2 = -y^3$ each pass through the origin. It follows that $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist, though $\lim_{t\to 0} f(ta, tb)$ exists for all $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. A similar example is found in the MA2321 lecture notes for Michaelmas Term 2015 preceding Proposition 7.10. The function in question is $g\colon \mathbb{R}^2 \to \mathbb{R}$, where

$$g(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^4} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

It should also be noted that, had this fallacious argument been valid, it would have provided a way of proving the Extreme Value Theorem for continuous functions on the unit sphere S^{n-1} without using the Bolzano-Weierstrass or any of the deep results related to "compactness". But some sort of tool dependent on the Bolzano-Weierstrass Theorem, the Heine-Borel Theorem or some other manifestation of "compactness" would be required to obtain a result equivalent to the Extreme Value Theorem for continuous real-valued functions on unit spheres in Euclidean spaces.

(b) *The function f is C^k if all partial derivatives of the function of order less than or equal to k are defined and continuous throughout the domain. Prove that if a function f defined over the whole of $\mathbb{R}^n$ satisfies the conditions stated at the beginning of this question, and if it is C^k throughout $\mathbb{R}^n$ (and in particular at and around the origin) then either s is a non-negative integer or else $s \geq k$. (It follows from this that if f is smooth then s must necessarily be a non-negative integer.)*

Solution. It follows from repeated application of 3(a) that the partial derivatives of the partial derivatives of f of order k must be homogeneous of degree $s-k$ on the complement of the origin. If it were the case that some kth order partial derivative of f were non-zero at some point of $\mathbb{R}^n$ away from the origin, where $k > s$ then it would follow from part (a) of this question that the limit of that kth order partial derivative would not exist at the origin. It follows that if f is C^k and $k > s$ then the kth order partial derivatives of f must be zero throughout $\mathbb{R}^n$. This can only happen if the partial derivatives of order $k-1$ are constant throughout $\mathbb{R}^n$. Now a non-zero function that is constant must be homogeneous of degree zero. Thus if $m+1$ is the smallest non-negative integer for which all partial derivatives of f of order $m+1$ are zero, then some mth order partial derivative of f must be a non-zero constant function, and must therefore be homogeneous of degree zero. But any mth order partial derivative of f is homogeneous of order $s-m$. It follows that $s = m$. We deduce from this that if f is C^k for some integer k satisfying $k > s$ then s must be an integer, and the partial derivatives of f of order s must be constant functions. We deduce that if the function f is C^k and is homogeneous of degree s then either $s \geq k$ or else s is a non-negative integer.

(In the case when s is a non-negative integer, some functions homogeneous of degree s will be smooth, but others will not be. Question 6 on this assignment provides results relevant to the case $s = 2$ and $n = 2$.)

5. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a continuous function on $\mathbb{R}^n$ that is homogeneous of degree 1 (see Question 3). Suppose that f is differentiable at the origin* $(0,0)$. *Prove that f is a linear transformation.*

Solution. The function f is continuous at zero, because it is differentiable there. The continuity and homogeneity of f ensure that $f(\mathbf{0}) = 0$. It follows that

$$\begin{aligned}(Df)_{\mathbf{0}}(\mathbf{v}) &= \lim_{t\to 0}\left(\frac{1}{t}(f(\mathbf{0}+t\mathbf{v}) - f(\mathbf{0}))\right) = \lim_{t\to 0}\left(\frac{1}{t}f(t\mathbf{v})\right) \\ &= \lim_{t\to 0}\left(\frac{1}{t}(tf(\mathbf{v}))\right) = \lim_{t\to 0} f(\mathbf{v}) = f(\mathbf{v}).\end{aligned}$$

Thus $f = (Df)_{\mathbf{0}}$.

Now the derivative $(Df)_{\mathbf{0}}$ of f at the origin is a linear transformation. It follows that f must itself be a linear transformation, as required.

6. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a real-valued function on $\mathbb{R}^2$ that is homogeneous of degree two. Suppose that f is continuously differentiable on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$, so that the first order partial derivatives exist at and are continuous around all points of $\mathbb{R}^2$ with the possible exception of the origin* $(0,0)$. *We denote the partial derivatives of f with respect to the Cartesian*

coordinates x and y by f_x and f_y respectively. We consider the possible existence and values of second order partial derivatives at the origin.

(a) *Prove that* $\left.\dfrac{\partial^2 f}{\partial x^2}\right|_{(0,0)}$ *exists if and only if* $f(1,0) = f(-1,0)$ *in which case*

$$\left.\frac{\partial^2 f}{\partial x^2}\right|_{(0,0)} = 2f(1,0).$$

Solution. It follows from 3(b) that $xf_x(x,y)+yf_y(x,y) = 2f(x,y)$ for all x and y. It follows that

$$tf_x(t,0) = 2f(t,0) = 2t^2 f(1,0),$$

and thus $f_x(t,0) = 2tf(1,0)$ for all $t > 0$. Similarly

$$-tf_x(-t,0) = 2f(-t,0) = 2t^2 f(-1,0),$$

and thus $f_x(-t,0) = -2tf(-1,0)$ for all $t > 0$. Also continuity of f_x ensures that $f_x(0,0) = 0$. It follows that

$$\begin{aligned} \lim_{h\to 0^+} \frac{f_x(h,0) - f_x(0,0)}{h} &= \lim_{t\to 0^+} \frac{f_x(t,0)}{t} \\ &= f_x(1,0) = 2f(1,0) \end{aligned}$$

and

$$\begin{aligned} \lim_{h\to 0^-} \frac{f_x(h,0) - f_x(0,0)}{h} &= -\lim_{t\to 0^+} \frac{f_x(-t,0)}{t} \\ &= -f_x(-1,0) = 2f(-1,0). \end{aligned}$$

It follows that the one-sided limits

$$\lim_{h\to 0^+} \frac{f_x(h,0) - f_x(0,0)}{h} \quad \text{and} \quad \lim_{h\to 0^-} \frac{f_x(h,0) - f_x(0,0)}{h}$$

exist. The second order partial derivative $\left.\dfrac{\partial^2 f}{\partial x^2}\right|_{(0,0}$ exists if and only if the above one-sided limits are equal to one another, in which case $f(-1,0) = f(1,0)$ and $\left.\dfrac{\partial^2 f}{\partial x^2}\right|_{(0,0} = 2f(1,0)$.

(b) *Prove that* $\left.\dfrac{\partial^2 f}{\partial x\,\partial y}\right|_{(0,0)}$ *exists if and only if* $f_y(1,0) = -f_y(-1,0)$ *in which case*

$$\left.\frac{\partial^2 f}{\partial x\,\partial y}\right|_{(0,0)} = f_y(1,0).$$

Solution. It follows from 3(a) that f_y is homogeneous of degree one, and therefore and thus $f_y(t,0) = tf_y(1,0)$ for all $t > 0$. Similarly and thus $f_y(-t,0) = tf_y(-1,0)$ for all $t > 0$. Also continuity of f_y ensures that $f_y(0,0) = 0$. It follows that

$$\begin{aligned} \lim_{h\to 0^+} \frac{f_y(h,0) - f_y(0,0)}{h} &= \lim_{t\to 0^+} \frac{f_y(t,0)}{t} \\ &= f_y(1,0) \end{aligned}$$

and

$$\begin{aligned} \lim_{h\to 0^-} \frac{f_y(h,0) - f_y(0,0)}{h} &= -\lim_{t\to 0^+} \frac{f_y(-t,0)}{t} \\ &= -f_y(-1,0). \end{aligned}$$

It follows that the one-sided limits

$$\lim_{h\to 0^+} \frac{f_y(h,0) - f_y(0,0)}{h} \quad \text{and} \quad \lim_{h\to 0^-} \frac{f_y(h,0) - f_y(0,0)}{h}$$

exist. The second order partial derivative $\left.\dfrac{\partial^2 f}{\partial x\,\partial y}\right|_{(0,0}$ exists if and only if the above one-sided limits are equal to one another, in which case $f_y(-1,0) = -f_y(1,0)$ and $\left.\dfrac{\partial^2 f}{\partial x^2}\right|_{(0,0} = f_y(1,0)$.

(c) *Prove that if* $f\colon \mathbb{R}^2 \to \mathbb{R}$ *is a function that is twice continuously differentiable throughout* $\mathbb{R}^2$ *(including at* $(0,0)$*), and if the function* f *is homogeneous of degree two, then there exist constants* E*,* F *and* G *such that*

$$f(x,y) = Ex^2 + 2Fxy + Gy^2.$$

Solution. If the function f is twice continuously differentiable then its first derivatives must be continuously differentiable and homogeneous of degree 1. It then follows from Question 5 above that f_x and f_y must be linear transformations. Therefore there must exist constants p, q, r and s such that

$$f_x(x,y) = px + qy, \quad f_y(x,y) = rx + sy.$$

But it follows from 3(b) that

$$\begin{aligned} f(x,y) &= \tfrac{1}{2}(xf_x(x,y) + yf_y(x,y)) \\ &= \tfrac{1}{2}(px^2 + (q+r)xy + sy^2) \end{aligned}$$

The result follows.

Further Comments.

A. Questions 3 to 6 arose from reflection on the example preceding Theorem 7.13 in the lecture notes for module MA2321 in Michaelmas Term 2015. That example considered the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined such that

$$f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

The example in the notes involves explicit expressions for the first order partial derivatives f_x and f_y of the function f that are moderately complicated.

The function f above is homogeneous of degree 2, and its first order partial derivatives are therefore homogeneous of degree 1. The homogeneity of the function f ensures that the function and its first order partial derivatives are determined by the values of the function on the unit circle $x^2 + y^2 = 1$. It is therefore possible use information about the function and its derivatives on the unit circle to deduce results concerning the behaviour of the function at the origin, and the methods could be generalized so as to apply more generally in situations where a function is determined by its values on the unit circle so as to be homogeneous of degree two throughout the plane $\mathbb{R}^2$.

Question 3 establishes the homogeneity of partial derivatives of a homogeneous function, and also Euler's Law concerning the relation between a differentiable homogeneous function and its first order partial derivatives. Question 4 establishes that a homogeneous function can only be smooth at the origin if its degree of homogeneneity is a non-negative integer. (Even then the requirement that the degree of homogeneity is a non-negative integer would not be enough to ensure smoothness.) Of course a function on the unit circle can be extended to $\mathbb{R}^2$ so as to be continuous homogeneous of degree s, for any desired non-negative real number s. Question 5 considers the particular case of differentiable homogeneous functions of degree one, showing that such a function has to be a linear transformation. (This result is used in 6(a).) This leads on to Question 6 itself.

Consider the amount of freedom there is in constructing a function f that is homogeneous of degree two on $\mathbb{R}^2$ and is twice differentiable at all points of $\mathbb{R}^2$ with the possible exception of the origin $(0,0)$. Such a function is determined by its values on the unit circle, and we can write

$$f(r\cos\theta, f\sin\theta) = r^2 g(\theta)$$

for all positive real numbers r and for all real numbers θ, where $g\colon\mathbb{R}\to\mathbb{R}$ is a twice-differentiable function satisfying $g(\theta+2\pi) = g(\theta)$ for all real numbers θ. Moreover, given any continuous real-valued function h defined over some open interval containing the closed interval $[0,2\pi]$, we can determine a real-valued function H on $[0,2\pi]$ that satisfies $H''(\theta) = h(\theta)$ by integrating twice. We can then determine constants p, q and r so as to ensure that if

$$g(\theta) = H(\theta) + p + q\theta + r\theta^2$$

then $g(0) = g(2\pi)$, $g'(0) = g'(2\pi)$ and $g''(0) = g''(2\pi)$. The function g can then be extended to the whole of $\mathbb{R}$ so as to satisfy $g(\theta+2\pi) = g(\theta)$ for all real numbers r. The resulting function will be twice-differentiable on $\mathbb{R}$ and will therefore give rise to a real-valued function f on $\mathbb{R}^2$ that is homogeneous of degree 2 and is twice-differentiable on the complement of the origin. The function h used to construct g need not be differentiable anywhere, and there is no linkage relating its behaviour around one point θ_1 of $[0,2\pi]$ to its behaviour around any other point θ_2 of $[0,2\pi]$.

Now Question 6 shows that $\dfrac{\partial^2 f}{\partial x^2}$ is only defined at the origin when $f(1,0) = f(-1,0)$. For a "general" function f constructed as described above, there is no more reason to suppose that $f(1,0) = f(-1,0)$ than there is to suppose that, if temperature were represented by a twice-differentiable function on the surface of the earth, the temperature at any one point would equal the temperature at the antipodal point on the other side of the earth. Similarly $\dfrac{\partial^2 f}{\partial x\,\partial y}$ is only defined at the origin when $f_y(1,0) = -f_y(-1,0)$. Interchanging x and y, and combining results, we see that in order that the identity

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}$$

be satisfied at the origin, where f is homogenous of degree 2 and is twice-differentiable away from the origin, the identities

$$f_y(1,0) = -f_y(-1,0) = f_x(0,1) = -f_x(0,-1).$$

must be satisfied. If the function f is constructed from an arbitrary continuous real-valued function h on $[0, 2\pi]$ in the manner described above, it would definitely be the exception rather than the rule for such identities to link the behaviour of the function f around the points $(1,0)$, $(0,1)$, $(-1,0)$ and $(0,-1)$ of the unit circle.

B. Question 6(c) should drive home the point that the only real-valued functions on $\mathbb{R}^2$ that are homogeneous of degree two and are also twice-continuously-differentiable throughout $\mathbb{R}^2$ are polynomial functions of degree two. Amongst homogeneous functions of degree two that are twice-continuously-differentiable away from the origin, the polynomial functions surely represent the exception rather than the rule.

C. Theorem 7.13 in the lecture notes for module MA2321 in Michaelmas Term 2015 shows that a real-valued function $f: V \to \mathbb{R}$ defined over an open set V in $\mathbb{R}^2$, then

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial^2 f}{\partial y\,\partial x}$$

provided that

$$\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}, \quad \frac{\partial^2 f}{\partial x\,\partial y} \quad \text{and} \quad \frac{\partial^2 f}{\partial y\,\partial x}$$

exist and are continuous throughout V. Corollary 7.14 generalizes this result to real-valued functions defined over open sets in n-dimensional Euclidean space. Those who have taken module

MA2322 (Calculus on Manifolds) should be aware that the exterior derivative operator d operating on smooth differential forms over a smooth manifold satisfies $d(d\omega) = 0$. This result depends on the result of Corollary 7.14, and if the result of Corollary 7.14 were not valid, then an enormous amount of differential geometry related to notions of "curvature" would just disappear, along with the theory of general relativity and much else. But contininuity of second order partial order derivatives is necessary in order to ensure equality of mixed second order partial derivatives.

D. Next we note that, given a function that fails to satisfy a differentiability condition at a single point, we can construct another function that fails to be differentiable at a dense set of points. For simplicity we move to one dimension. Consider the function $f: \mathbb{R} \to \mathbb{R}$ where

$$f(x) = \frac{1}{1 + |x|}$$

for all real numbers x. This function is not differentiable at the zero, those it is differentiable for all non-zero values of x. It is also "Lipschitz continuous" as it satisfies the condition $|f(u) - f(v)| \leq |u - v|$ for all real numbers u and v. Let $q_1, q_2, q_3, \ldots$ be an infinite sequence in which every rational number occurs exactly once. (The set of rational numbers is "countable", and thus such an infinite sequence exists.) Let

$$g(x) = \sum_{j=1}^{\infty} 2^{-j} f(x - q_j).$$

Then the function $g: \mathbb{R} \to \mathbb{R}$ is a continuous function that is not differentiable for any rational value of its argument. (The continuity of g follows from the standard result that the sum of a uniformly convergent sequence of continuous functions is guaranteed to be continuous.) Other more sophisticated examples of "mathematical pathology" include Weierstrass's examples of real-valued functions that are continuous everywhere but differentiable nowhere, and the "Peano space-filling curve" that is a continuous surjective function from the closed interval $[0, 1]$ to the closed unit square $[0, 1]^2$. For further information, see the Wikipedia article on "Pathological (mathematics)" at the following URL:

`https://en.wikipedia.org/wiki/Pathological_%28mathematics%29`

It is a mistake to think that the functions represented by algebraic expressions, combined with trigonometric, exponential and logarithms functions, such as one finds in Calculus textbooks have behaviour that is typical of continuous, differentiable or smooth functions in general.