Martin Schmidt
Ross Ogilvie

Introduction to Partial Differential Equations
Exercise sheet 1

6th September, 2022

This exercise sheet is revision and does not count towards the exercise points. But please feel free to attempt them (before or after the tutorial) and submit them for correction any way.

1. **Chain rule in multiple variables.** Recall the chain rule for functions of multivariable variables (Satz 10.4(iii) in Schmidt's Analysis II script): Let $f : U \subset X \to Y$ be differentiable at $x_0 \in U$ and $g : V \subset Y \to Z$ be differentiable at $f(x_0) \in f[U] \subset V$. Then $g \circ f$ is differentiable at x_0 and

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

(a) Why does this chain rule above use function composition, when the chain rule for functions of a single variable uses multiplication? i.e.

$$\frac{d}{dx}(x^2+1)^3 = 3(x^2+1)^2 \cdot 2x. = 6x(x^2+1)^2.$$

(b) Suppose that $u : \mathbb{R}^n \to \mathbb{R}$ and $x : \mathbb{R} \to \mathbb{R}^n$. Express the chain rule with partial derivatives to show that

$$\frac{d}{dt}u(x(t)) = \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}\frac{dx_i}{dt}.$$

(c) Write the above formula in terms of gradients and dot products.

(d) Consider the function $u(x, y) = x^2 + 2y$ and the polar coordinates $x = r\cos\theta, y = r\sin\theta$. Compute the radial and angular derivatives of u.

(e) Consider a scalar function $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ of $2n+1$ variables and a function $u : \mathbb{R}^n \to \mathbb{R}$. Write an expression for the derivative of $F(\nabla u(x), u(x), x)$ with respect to x_1.

Solution.

(a) In Analysis II we considered the derivative at each point as a linear map. When you write a linear map as a matrix, then composition of linear maps corresponds to matrix multiplication. In the case of single variable functions, the matrices have only a single entry and so the matrix multiplication is just multiplying the two derivative expressions.

(b) Applying the chain rule, with the derivatives written as matrices:

$$(u \circ x)'(t) = u'(x(t)) \circ x'(t) = \begin{bmatrix} \frac{\partial u}{\partial x_1}(x(t)) & \frac{\partial u}{\partial x_2}(x(t)) & \dots & \frac{\partial u}{\partial x_n}(x(t)) \end{bmatrix} \begin{bmatrix} \frac{dx_1}{dt}(t) \\ \frac{dx_2}{dt}(t) \\ \vdots \\ \frac{dx_2}{dt}(t) \end{bmatrix}$$

$$= \sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x(t))\frac{dx_i}{dt}(t)$$

(c) The function x is a single-variable vector-valued function. Therefore $x' : \mathbb{R} \to \mathbb{R}^n$ is vector-valued too. On the other hand, u is a function of multiple variables but gives a scalar. The gradient of a scalar function has the same inputs, but is vector-valued:

$$\nabla u : \mathbb{R}^n \to \mathbb{R}^n, \quad \begin{bmatrix} \frac{\partial u}{\partial x_1}(x) & \frac{\partial u}{\partial x_2}(x) & \dots & \frac{\partial u}{\partial x_n}(x) \end{bmatrix},$$

which is also vector valued. We seen then that $\frac{du}{dt} = \nabla u \cdot x'$.

(d) We can use either of the formulas from (b) or (c). We get

$$\begin{aligned}\partial_r u &= (2x)(\cos\theta) + (2)(\sin\theta) = 2r\cos^2\theta + 2\sin\theta \\ \partial_\theta u &= (2x)(-r\sin\theta) + (2)(r\cos\theta) = -2r^2\sin\theta\cos\theta + 2r\cos\theta.\end{aligned}$$

(e) Let us write the variables of F as $F(p, u, x)$, where p and x are vectors. We can see here that u is both the name of an input of F and a function in its own right. This ambiguity is standard, and usually avoids introducing many extra variables. We do introduce a separate letter for the p variable however, rather than trying to write derivatives of F with respect to components of ∇u. If we consider $x_2, \dots, x_n$ as constants, which is the idea of a partial derivative then we can write the composition as $F \circ G$ for $G(x_1) = (\nabla u(x), u(x), x) \in \mathbb{R}^{2n+1}$. The advantage now is that this is the form in part (b) and (c) and we can use that version of the chain rule instead of the full matrix form. We compute

$$\begin{aligned}\frac{dF}{dx_1} &= \sum_{i=1}^{n} \frac{\partial F}{\partial p_i}\frac{\partial(\nabla u)_i}{\partial x_1} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x_1} + \sum_{i=1}^{n}\frac{\partial F}{\partial x_i}\frac{\partial x_i}{\partial x_1} \\ &= \sum_{i=1}^{n}\frac{\partial F}{\partial p_i}\frac{\partial}{\partial x_1}\left(\frac{\partial u}{\partial x_i}\right) + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x_1} + \frac{\partial F}{\partial x_1} \\ &= \sum_{i=1}^{n}\frac{\partial F}{\partial p_i}\frac{\partial^2 u}{\partial x_1 \partial x_i} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x_1} + \frac{\partial F}{\partial x_1}.\end{aligned}$$

2. Contour Diagrams.

Consider the function $f : \{(t, x) \in \mathbb{R}^2 \mid t \geq 0\} \to \mathbb{R}$ defined by $f(t, x) = \arctan(t - x^2)$.

(a) Draw a *contour diagram* for this function. A contour is another word for a level set $f^{-1}[\{c\}]$.

(b) What is the maximum and minimum of this function?

(c) What is the behaviour of this function for large values of t?

Solution.

(a) The level set $f^{-1}[\{c\}]$ of this function is the curve $t - x^2 = \tan c$. This is a parabola.

(b) For a fixed value of t this function has a maximum at $x = 0$ of $\arctan t$ since arctan is an increasing function. It has no minimum, but tends to $-\pi/2$ as $x \to \pm\infty$.

(c) The limit of this function as $t \to \infty$ is not so well-defined. For a fixed value of x, it tends to $\pi/2$ (the pointwise limit). If we move along the level set as we take $t \to \infty$ then clearly the value of f is not changing. It does not converge in L^1.

3. Multiindices and the Generalised Leibniz rule. In this question we introduce multiindex notation. A *multiindex* of n variables is a vector $\gamma \in \mathbb{N}_0^n$.

(a) Let $x = (x_1, x_2, x_3)$ be coordinates on $\mathbb{R}^3$. Write out the full expression for the derivative $\partial^{(0,2,1)}$.

(b) Why do we need to assume that partial derivatives commute for multiindex notation to be useful?

(c) Which multiindices satisfy $|\gamma| \leq 2$ and which satisfy $\gamma \leq (0, 2, 1)$?

Solution.

(a) $\partial^{(0,2,1)} = \partial_1^0 \partial_2^2 \partial_3^1 = \partial_2^2 \partial_3 = \frac{\partial^2}{(\partial x_2)^2} \frac{\partial}{\partial x_3}$. The lower number is the coordinate and the upper number is the order of the derivative. The zeroth order derivative is just the function itself.

(b) The multiindex notation applies the partial derivatives in a certain order. It does not have the capacity to express the same derivatives applied in a different order, for example the difference between $\partial_1 \partial_2$ and $\partial_2 \partial_1$. So, to be useful, this order must not make a difference, i.e. the partial derivatives must commute.

(c) The multiindices with $|\gamma| \leq 2$ are

$$\begin{aligned}
&\text{order } 0 : (0, 0, 0),\\
&\text{order } 1 : (1, 0, 0), (0, 1, 0), (0, 0, 1),\\
&\text{order } 2 : (2, 0, 0), (1, 1, 0), (1, 0, 1), (0, 2, 0), (0, 1, 1), (0, 0, 2).
\end{aligned}$$

As well has having a 'level' structure coming from the order of the multiindex, there is an ordering between some multiindices: $\gamma \geq \delta$ if and only if $\gamma - \delta \in \mathbb{N}_0^+$. For example $(0, 2, 1) \geq (0, 1, 1)$ because $(0 - 0, 2 - 1, 1 - 1) \in \mathbb{N}_0^+$. On the other hand $(1, 0, 1) \ngeq (0, 1, 0)$ because $(1, -1, 1) \notin \mathbb{N}_0^+$. Note also that $(0, 1, 0) \ngeq (1, 0, 1)$. We say in this case that the two multiindices are *incomparable*. The multiindices with $\gamma \leq (0, 2, 1)$ are

$$\begin{aligned}
&\text{order } 0 : (0, 0, 0),\\
&\text{order } 1 : (0, 1, 0), (0, 0, 1),\\
&\text{order } 2 : (0, 2, 0), (0, 1, 1),\\
&\text{order } 3 : (0, 2, 1).
\end{aligned}$$