In these notes we prove two results that were stated in lectures and whose proofs are not examinable. The first one is the differentiability of the inverse in the inverse function theorem. For completeness we give the whole proof including the parts given in lectures (which of course remain examinable).

Theorem (Inverse Function Theorem)**.** *Let $U \subset \mathbb{R}^n$ be an open set, let $f\colon U \to \mathbb{R}^n$ be a continuously differentiable function, and let $a \in U$. If $f'(a)$ is invertible, then there exist open neighbourhoods V of a with $V \subset U$ and W of $f(a)$ such that $f\restriction_V\colon V \to W$ is a bijection with a continuously differentiable inverse $g\colon W \to V$. Moreover, $g'(y) = \big[f'(g(y))\big]^{-1}$ for all $y \in W$.*

Proof. STEP 1: Without loss of generality, we may assume that $a = f(a) = 0$ and $f'(a) = I$, the identity map. Indeed, let $T = f'(a)$ and consider $h\colon U - a \to \mathbb{R}^n$ given by $h(x) = T^{-1}(f(x+a) - f(a))$. Then h is continuously differentiable with $h'(x) = T^{-1} \circ f'(x+a)$, $h(0) = 0$ and $h'(0) = T^{-1} \circ f'(a) = I$. If we know the result for h, then we can deduce it for f, since $f(x) = T(h(x-a)) + f(a)$.

We will now assume that $a = f(a) = 0$ and $f'(a) = I$. Since f' is continuous, we can fix $r > 0$ such that $D_r(0) \subset U$ and $\|f'(x) - I\| \leqslant \frac{1}{2}$ for all $x \in D_r(0)$.

STEP 2: For all $x, y \in D_r(0)$ we have $\|f(x) - f(y)\| \geqslant \frac{1}{2}\|x - y\|$. In particular, f is injective on $D_r(0)$. To see this, consider $h(x) = x - f(x)$ for $x \in D_r(0)$. Then h is differentiable with $h'(x) = I - f'(x)$ and $\|h'(x)\| \leqslant \frac{1}{2}$. By the Mean Value Inequality, we have $\|h(x) - h(y)\| \leqslant \frac{1}{2}\|x - y\|$ for $x, y \in D_r(0)$. Hence by the triangle-inequality, we have

$$\frac{1}{2}\|x-y\| \geqslant \|h(x) - h(y)\| = \|x - y - (f(x) - f(y))\| \geqslant \|x - y\| - \|f(x) - f(y)\|$$

which implies the claim.

STEP 3: For any s with $0 < s < r/2$ we have $D_s(0) \subset f\big(B_{2s}(0)\big) \subset f\big(D_r(0)\big)$. To prove this, we fix $y \in D_s(0)$ and define $h(x) = x - f(x) + y$ for $x \in D_r(0)$. Then h is differentiable with $h'(x) = I - f'(x)$. As before, the Mean Value Inequality shows that $\|h(x) - h(z)\| \leqslant \frac{1}{2}\|x - z\|$ for all $x, z \in D_r(0)$, and thus h is a contraction mapping. It follows that for $x \in B_{2s}(0)$, we have

$$\|h(x)\| = \|h(x) - h(0) + y\| \leqslant \|h(x) - h(0)\| + \|y\| \leqslant \frac{1}{2}\|x\| + \|y\| + \leqslant 2s\ .$$

Hence h is a contraction mapping $B_{2s}(0) \to B_{2s}(0)$, and $B_{2s}(0)$ is a non-empty complete metric space (being a closed subset of the complete space $\mathbb{R}^n$). By the Contraction Mapping Theorem, h has a fixed point x, which satisfies $f(x) = y$.

STEP 4: Set $W = D_s(0)$ and $V = f^{-1}(W) \cap D_r(0)$. Then V and W are open neighbourhoods of 0, $f\restriction_V$ is a bijection $V \to W$ whose inverse $g\colon W \to V$ is continuously differentiable with $g'(y) = \big[f'(g(y))\big]^{-1}$ for all $y \in W$.

Open balls are open and f is continuous, and so V and W are open and both contain 0 since $f(0) = 0$. By Step 2, $f\restriction_V$ is injective, and by Step 3, it is also surjective. We first note that the inverse $g\colon W \to V$ of $f\restriction_V$ is continuous. Indeed, given $w, z \in W$, set $x = g(w)$, $y = g(z)$ and use Step 2 to obtain

$$\|g(w) - g(z)\| = \|x - y\| \leqslant 2\|f(x) - f(y)\| = 2\|w - z\|\ . \tag{1}$$

Thus, g is 2-Lipschitz and, in particular, continuous. For differentiability, fix $b \in W$. Let $a = g(b)$ and $T = f'(a)$. Then

$$f(a+h) = f(a) + T(h) + \varepsilon(h)\|h\| \tag{2}$$

where $\varepsilon(0) = 0$ and ε is continuous at 0. Now fix $\delta > 0$ such that $D_\delta(b) \subset W$, and for $k \in D_\delta(0)$ set $h = h(k) = g(b+k) - g(b)$. It follows from (2) that

$$k = f(a+h) - f(a) = T(h) + \varepsilon(h)\|h\| \ ,$$

and hence $h = T^{-1}(k) - T^{-1}(\varepsilon(h))\|h\|$. Thus,

$$g(b+k) = g(b) + h = g(b) + T^{-1}(k) - T^{-1}(\varepsilon(h))\|h\| \ .$$

The function $T^{-1} \circ \varepsilon \circ h$ is continuous at 0 and has value 0 there. Moreover, it follows from (1) that

$$\|h\| = \|g(b+k) - g(b)\| \leqslant 2\|k\| \ ,$$

and hence the term $T^{-1}(\varepsilon(h))\|h\|$ is $o(\|k\|)$. This shows that g is differentiable at b with $g'(b) = T^{-1} = \left[f'(g(b))\right]^{-1}$. Since $f' \circ g$ is continuous, and since $A \mapsto A^{-1} \colon \mathcal{G}_n \to \mathcal{M}_n$ is continuous (the entries of $A^{-1} = \frac{1}{\det A} \operatorname{adj} A$ are rational functions in the entries of A), it follows that $g' \colon W \to L(\mathbb{R}^n, \mathbb{R}^n)$ is continuous. □

The second result is a version of Taylor's theorem in higher dimension that was needed in describing stationary points in terms of the second derivative.

Lemma (Second-order Taylor's theorem with Peano remainder)**.** *Let U be an open subset of $\mathbb{R}^m$, let $f \colon U \to \mathbb{R}^n$ be a function and let $a \in U$. If f is twice differentiable at a, then*

$$f(a+h) = f(a) + f'(a)(h) + \frac{1}{2} f''(a)(h,h) + o(\|h\|^2) \ .$$

Proof. It is enough to prove this for each component of f, and thus we can assume that $n = 1$. By definition of the second derivative, there is a function ε such that

$$f'(a+h)(k) = f'(a)(k) + f''(a)(h,k) + \varepsilon(h)(k)\|h\| \tag{3}$$

where $\varepsilon(h) \in L(\mathbb{R}^m, \mathbb{R}^n)$ and $\varepsilon(h) \to 0$ as $h \to 0$. Let us define

$$\eta(h) = f(a+h) - f(a) - f'(a)(h) - \frac{1}{2} f''(a)(h,h)$$

in some neighbourhood of 0, where this is defined. We will be done if we show that $\eta(h) = o(\|h\|^2)$. We fix h and define $\varphi \colon [0,1] \to \mathbb{R}$ by

$$\varphi(t) = f(a+th) - f(a) - tf'(a)(h) - \frac{t^2}{2} f''(a)(h,h) \ .$$

Then φ is continuous and is differentiable on $(0,1)$ by the Chain Rule. By the Mean Value Theorem, we have

$$\eta(h) = \varphi(1) - \varphi(0) = \varphi'(\theta) = f'(a+\theta h)(h) - f'(a)(h) - \theta f''(a)(h,h)$$

for some $\theta = \theta(h) \in (0,1)$. Using (3), we obtain

$$\begin{aligned} \eta(h) &= \big(f'(a)(h) + \theta f''(a)(h,h) + \varepsilon(\theta h)(h)\|\theta h\|\big) - f'(a)(h) - \theta f''(a)(h,h) \\ &= \varepsilon(\theta h)(h)\|\theta h\| \end{aligned}$$

It follows that $|\eta(h)| \leqslant \|\varepsilon(\theta h)\| \cdot \|h\|^2 \theta$. Since $\varepsilon(\theta h) \to 0$ as $h \to 0$, it follows that $\eta(h) = o(\|h\|^2)$ as required. □

One can define higher order derivatives by induction: we say f is k-times differentiable at a if f is $(k-1)$-times differentiable on an open neighbourhood of a and the $(k-1)^{\text{th}}$ derivative $f^{(k-1)}$ is differentiable at a. The k^{th} derivative $f^{(k)}(a)$ of f at a is then the derivative of $f^{(k-1)}$ at a, and it is a k-linear map $\underbrace{\mathbb{R}^m \times \cdots \times \mathbb{R}^m}_{k \text{ times}} \to \mathbb{R}^n$ which for each $h_2, \ldots, h_k \in \mathbb{R}^m$ satisfies

$$f^{(k-1)}(a+h_1)(h_2,\ldots,h_k) = \\ f^{(k-1)}(a)(h_2,\ldots,h_k) + f^{(k)}(a)(h_1,h_2,\ldots,h_k) + o(\|h_1\|) \ .$$

The obvious analogue of the above lemma, a k^{th}-order Taylor's theorem with the Peano remainder, is valid with a similar proof. Other versions of Taylor's theorem with other remainders are also valid; they are obvious generalizations of the one-dimensional results.