MA442 MIDTERM EXAM 1

RYAN WANS

Problem 1 (Direct proof of the Inversion Theorem)**.** My solutions to parts (a)-(e) follow.

(a) We have $f \in C^1(\Omega) \implies Df \in C^0(\Omega)$, and thus for each $\epsilon > 0$ there exists some $r > 0$ such that $x \in B_r(x_0) \implies \|Df(x_0) - Df(x)\|_{pp} \leq \epsilon$. Since $Df(x_0) \in \mathrm{Hom}(\mathbb{R}^p, \mathbb{R}^p) \implies \Gamma \in \mathrm{Hom}(\mathbb{R}^p, \mathbb{R}^p)$, then the result $\|\Gamma x\|_q \leq \|\Gamma\|_{pq} \|x\|_p$ allows us to conclude that

$$\|\Gamma \circ (Df(x_0) - Df(x))\|_{pp} \leq \|\Gamma\|_{pp} \|Df(x_0) - Df(x)\|_{pp} \leq \|\Gamma\|_{pp} \epsilon,$$

and thus by the linearity of Γ and inverse property $\Gamma(Df(x_0)) = \mathrm{Id}$,

$$\|\mathrm{Id} - \Gamma \circ Df(x_0)\|_{pp} \leq \|\Gamma\|_{pp} \epsilon = \frac{1}{2} \tag{1}$$

when we choose $\epsilon = (2\|\Gamma\|_{pp})^{-1}$. Since any linear operator between normed spaces has bounded norm, we know that $\|\Gamma\|_{pp} < \infty$.

(b) $DF_y(x) = D(f(x) - y) = Df(x) - Dy = Df(x)$ by property of the derivative and $y \in \mathbb{R}^p \implies Dy = 0$. Thus Df exists over Ω implies that DF_y exists over Ω. We have that $\|F_y(x_0)\|_p = \|y - y_0\|_p < s = \frac{r}{2\|\Gamma\|_{pp}}$ since $y \in B_s(y_0)$. Thus we have that,

$$\|\Gamma \circ F_y(x_0)\|_p \leq \|\Gamma\|_{pp} \|F_y(x_0)\|_p < \|\Gamma\|_{pp} \frac{r}{2\|\Gamma\|_{pp}} = \frac{r}{2}. \tag{2}$$

Finally, we upper bound

$$\|\mathrm{Id} - \Gamma \circ DF_y(x)\|_{pp} = \|\mathrm{Id} - \Gamma \circ Df(x)\|_{pp} \leq \frac{1}{2} \tag{3}$$

which follows from $Dy = 0$ and equation 1 for $x \in B_r(x_0)$.

(c) By convexity of open balls, $x_1, x_2 \in B_r(x_0) \implies$ line segment $[x_1, x_2] \subset B_r(x_0)$. From a proposition given in class (related to Lemma 41.3 in Bartle), from the Mean Value Theorem, we have that

$$\begin{aligned}\|G_y(x_1) - G_y(x_2)\|_p &\leq \sup_{c \in [x_1, x_2]} \|DG_y(c)\|_{pp} \|x_1 - x_2\|_p \\ &\leq \sup_{c \in B_r(x_0)} \|DG_y(c)\|_{pp} \|x_1 - x_2\|_p.\end{aligned}$$

Furthermore, we have that

$$DG_y(c) = D[c - \Gamma \circ F_y(c)] = \mathrm{Id} - D[\Gamma \circ F_y(c)] = \mathrm{Id} - \Gamma \circ DF_y(x)$$

by linearity of the derivative, application of chain rule, and that fact that $D\Gamma = \Gamma$. Since we restrict ourselves to $x \in B_r(x_0)$, we can apply the result in equation 3 and use $DF_y = Df$ to get that

$$\sup_{c \in B_r(x_0)} \|DG_y(c)\|_{pp} = \frac{1}{2},$$

and thus $\|G_y(x_1) - G_y(x_2)\|_p \leq \frac{1}{2} \|x_1 - x_2\|_p$.

(d) We perform induction over n. The base case of $k = 0$ follows from $\phi_1(y) - \phi_0(y) = G_y(x_0) - x_0 = \Gamma \circ F_y(x_0)$ and thus $\|\phi_1(y) - \phi_0(y)\|_p = \|\Gamma \circ F_y(x_0)\|_p < r/2$ by the result in equation 2. We now hypothesize that for $k = n - 1$, the inequality

$$\|\phi_n(y) - \phi_{n-1}(y)\|_p \leq \frac{1}{2^{n-1}} \|\phi_1(y) - \phi_0(y)\|_p$$

holds true. Then when $k = n$, we have that,

$$\begin{aligned}
\|\phi_{n+1}(y) - \phi_n(y)\|_p &= \|G_y(\phi_n(y)) - G_y(\phi_{n-1}(y))\|_p \\
&\le \frac{1}{2}\|\phi_n(y) - \phi_{n-1}(y)\|_p \qquad (*) \\
&\le \frac{1}{2^n}\|\phi_1(y) - \phi_0(y)\|_p \le \frac{1}{2^n}\cdot\frac{r}{2} = \frac{r}{2^{n+1}}
\end{aligned}$$

by the result in part (c), the induction hypothesis, and the inequality from the base case of $k = 0$. We are able to apply part (c) since $y \in B_s(y_0)$ by assumption. Since, for $k \in \mathbb{N}$, we have

$$\begin{aligned}
\|\phi_{k+1}(y) - \phi_0(y)\|_p &= \|\phi_{k+1}(y) - \phi_k(y) + \cdots + \phi_1(y) - \phi_0(y)\|_p \\
&\le \|\phi_{k+1}(y) - \phi_k(y)\|_p + \cdots + \|\phi_1(y) - \phi_0(y)\|_p \\
&< \sum_{n=0}^{k}\frac{r}{2^{n+1}} = \frac{r}{2}\sum_{n=0}^{k}\frac{1}{2^n} = r - \frac{r}{2^{k+1}} < r
\end{aligned}$$

and thus for all $k \in \mathbb{N}$, we have $\|\phi_{k+1}(y) - \phi_0(y)\|_p = \|\phi_{k+1}(y) - x_0\|_p < r$ which implies that $\phi_{k+1}(y) \in B_r(x_0)$.

(e) For continuity of each $\phi_k(y)$ over $B_s(y_0)$, we again perform again induction over k. For $k = 0$, we have that $\phi_0(y) = x_0$ is constant and therefor continuous over $y \in B_s(y_0)$. We now hypothesize that $\phi_{k-1}(y)$ is continuous over $y \in B_s(y_0)$. Then we have that

$$\phi_k(y) = G_y(\phi_{k-1}(y)) = \phi_{k-1}(y) - \Gamma \circ (f(\phi_{k-1}(y)) - y).$$

Since $\phi_{k-1}\colon B_s(y_0) \to B_r(x_0)$ and f is continuous over $B_r(x_0) \subset \Omega$, the composition $f \circ \phi_{k-1}$ of continuous functions is continuous over $B_s(y_0)$. The affine function $f(\phi_{k-1}(y)) - y$, compostion of continous functions $\Gamma \circ (f(\phi_{k-1}(y)) - y)$, and thus the entire expression of ϕ_k are continuous over $B_s(y_0)$. Since (ϕ_k) is Cauchy as shown in $(*)$, we know there exists some $\phi = \lim_k(\phi_k)$. Then for any $y \in B_s(y_0)$, we have

$$\begin{aligned}
\|\phi(y) - \phi_0(y)\|_p &= \|\phi(y) - \phi_n(y) + \phi_n(y) - \phi_0(y)\|_p \\
&\le \|\phi(y) - \phi_n(y)\|_p + \|\phi_n(y) - \phi_0(y)\|_p \\
&< \left(\sum_{k=n}^{\infty}\|\phi_{k+1}(y) - \phi_k(y)\|_p\right) + \left(r - \frac{r}{2^n}\right) \\
&\le \left(\sum_{k=n}^{\infty}\frac{r}{2^{k+1}}\right) + \left(r - \frac{r}{2^n}\right) = \frac{r}{2^{n+1}} + r - \frac{r}{2^n} < r.
\end{aligned}$$

by triangle inequality, the same method as in (d) with the result from (d), another result from (d), and finally the fact that $r(2^{n+1})^{-1} < r(2^n)^{-1}$. Thus, $\phi(y) \in B_r(x_0)$ and for any $\epsilon > 0$ and any $y \in B_s(y_0)$, if we choose $N \in \mathbb{N}$ such that $\epsilon < \frac{r}{2^{N+1}}$, then we have that $n \ge N \implies \|\phi(y) - \phi_n(y)\|_p < \epsilon$, and thus that $(\phi_k) \rightrightarrows \phi$. Finally, $\|\phi(y) - G_y(\phi_{n-1}(y))\|_p \le r(2^{n+1})^{-1}$ implies that $\lim_n G_y(\phi_{n-1}(y)) = \phi(y)$ and therefore $\phi(y) = G_y(\phi(y))$ which is continuous over $B_s(y_0)$ by the continuity of each ϕ_k over $B_s(y_0)$, continuity of G_y, and the Uniform Convergence Theorem. By analysis of G_y itself, we get that

$$G_y(\phi(y)) = \phi(y) - \Gamma \circ F_y(\phi(y)) = \phi(y)$$

implies that $\Gamma \circ (f(\phi(y)) - y) = 0$ which, by bijectivity of Γ, means that $f(\phi(y)) - y = 0 \iff f(\phi(y)) = y$ for $y \in B_s(y_0)$. In conclusion, we have that $f \circ \phi = \mathrm{Id}$ over $B_s(y_0)$ and thus ϕ is the right inverse of f over $B_s(y_0)$.

Problem 2 (Tangent vectors and Lagrange's Theorem)**.** My solutions to parts (a) and (b) follow.

(a) ($\Rightarrow$) We have that $\tau \in T_{x_0}(S) \implies$ there exists some $\gamma(t)\colon (-\delta,\delta) \to S$ such that $\gamma(0) = x_0$ and $D\gamma(0) = \tau$. Since $\gamma(t) \in S$ for $t \in (-\delta,\delta)$ implies that $g(\gamma(\tau)) = 0$ over this domain, we get that $D(g(\gamma(t))) = Dg(\gamma(t)) \circ D\gamma(t) = 0$ by the chain rule and thus when $t = 0$, we have $Dg(x_0) \circ \tau = 0$.

($\Leftarrow$) By regularity of $x_0 \in S$, we know that $\operatorname{rank} Dg(x_0) = k$ for $Dg(x_0)\colon \mathbb{R}^p \to \mathbb{R}^k$ with $p \geq k$. Let $[p] = \{1, 2, \ldots, p\}$ and $I = \{i_1, \ldots, i_k\} \subset [p]$ be the set of indicies of columns of $Dg(x_0)$ that contain pivots. Then if we define $K = D_{x_{i_1},\ldots,x_{i_k}} g(x_0)$, then $\operatorname{rank} K = k \implies \det K \neq 0 \implies K \in \mathrm{GL}_k(\mathbb{R})$ is bijective. For $\Omega \subset \mathbb{R}^p = \mathbb{R}^{p-k} \times \mathbb{R}^k$ open and $x_0 = (a,b) \in \Omega$, then $g \in C^1(\Omega)$ allows us to apply the Implicit Function Theorem. Then there exists an open neighborhood $W \subset \mathbb{R}^{p-k}$ with $a \in W$ and a unique mapping $\phi\colon W \to \mathbb{R}^k$ in $C^1(W)$ such that $b = \phi(a)$ and $g(x, \phi(x)) = 0$ for $x \in W$. Moreover, there exists an open neighborhood $U \subset \Omega$ with $x_0 \in U$ such that

$$g(x,y) = 0 \text{ for } (x,y) \in U \iff y = \phi(x) \text{ for } x \in W.$$

Let $J = [p] \setminus I$ and define $L = D_{x_{j_1},\ldots,x_{j_{p-k}}} g(x_0)$ for each $j_i \in J$ and WLOG assume that

$$Dg(x_0) = \begin{bmatrix} L & K \end{bmatrix}.$$

Moreover, let $\tau = (\tau_a, \tau_b) \in \Omega$. We know from the Implicit Function Theorem results that $D\phi(x_0) = -K^{-1}L$. Furthermore, $\tau \in \ker Dg(x_0) \implies Dg(x_0)\tau = L\tau_a + K\tau_b = 0$ and thus we can express $\tau_b = -K^{-1}L\tau_a$. Then for the parametric curve $s\colon (-\delta/||\tau_a||, \delta/||\tau_a||) \to W$ where we define $s(t) = a + t\tau_a$, we can further define

$$\gamma(t) = (s(t), \phi(s(t))) = (a + t\tau_a, \phi(a + t\tau_a))$$

for $\gamma\colon (-\delta,\delta) \to W \times \mathbb{R}^k$. Since W is open and $a \in W^{\circ}$, there exists some $\delta > 0$ such that $B_\delta(a) \subset W$, and this is how we choose the domain of such γ. Moreover, given such a $\delta > 0$, if we restrict the domain of $s(t)$ to $(-\delta/||\tau_a||, \delta/||\tau_a||)$, then $s(t) \in B_\delta(a) \subset W$. Of course any $t \in (-\delta,\delta)$ has the property that $g(\gamma(t)) = 0$ by result of the Implicit Function Theorem since $s(t) \in W$ and thus $\gamma(t) \in S$. Trivially, $\gamma(0) = (s(0), \phi(s(0))) = (a, \phi(a)) = (a,b) = x_0$. To prove the second property of γ, we observe that

$$D\gamma(0) = \begin{bmatrix} Ds(0) & D\phi(s(0)) \end{bmatrix} = \begin{bmatrix} \tau_a & -K^{-1}L \circ \tau_a \end{bmatrix} = \begin{bmatrix} \tau_a & \tau_b \end{bmatrix} = \tau.$$

Therefore, by definition given in the problem, $\tau \in T_{x_0}(S)$.
From the fact that $\tau \in T_{x_0}(S) \iff Dg(x_0)(\tau) = 0$ we know that $T_{x_0}(S) = \ker Dg(x_0)$. Thus by the Rank-Nullity Theorem, we have

$$\operatorname{rank} Dg(x_0) + \dim\ker D_g(x_0) = p$$

and thus $\dim T_{x_0}(S) = \dim\ker D_g(x_0) = p - \operatorname{rank} Dg(x_0) = p - k$. The subspace axioms are trivially checked since $Dg(x_0)(0) = 0$ implies $0 \in T_{x_0}(s)$, and $u, v \in T_{x_0}(S) \implies (\alpha u + \beta v) \in T_{x_0}(S)$ for $\alpha, \beta \in \mathbb{R}$, by linearity of the derivative.

(b) Let $f\colon \Omega \to \mathbb{R}$ be class $C^1(\Omega)$ and suppose f has a relative extrema at $x_0 \in S$ which is also regular. Then for any $\tau \in T_{x_0}$, we know that there exists some $\gamma\colon (-\delta,\delta) \to S$ of class C^1 such that $\gamma(0) = x_0$ and $D\gamma(0) = \tau$. The define some $h : (-\delta,\delta) \to \mathbb{R}$ such that $h(t) = f(\gamma(t))$ which is inherently also of class C^1. Then x_0 relative extrema of f on S implies that h has a local extrema at $t = 0$. Since Ω open means that $x_0 \in \Omega \implies x_0 \in \Omega^{\circ}$, and Dh exists in $B_\delta(0)$, then by the 1st Derivative Test Theorem, we have that $Dh(0) = 0$. By chain rule,

$$Dh(t) = D(f \circ \gamma(t)) = Df(\gamma(t)) \circ D\gamma(t)$$

and when $t = 0$ we have

$$Dh(0) = Df(\gamma(0)) \circ D\gamma(0) = Df(x_0) \circ \tau = 0.$$

Because every $\tau \in T_{x_0}(S)$ has such γ and h functions, then $Df(x_0)$ vanishes over $\tau \in T_{x_0}(S)$.

Problem 3 (Lebesgue's Criterion for Riemann Integrability)**.** My solutions to parts (a)-(c) follow.

(a) We begin with the fact that D_α has content zero. By definition, for each $\epsilon > 0$, there exists a collection of cells $\{I_1, \dots, I_r\}$ such that $D_\alpha \subset \bigcup_{k=1}^r I_k$ and $\sum_{k=1}^r c(I_k) \leq \epsilon$. We choose such $\epsilon = \alpha$. WLOG let each $\{I_1, \dots, I_r\}$ be open, which was both proven in class and stated in Bartle pg. 413. Then we have that $D_\alpha \subset \bigcup_{k=1}^r I_k \subset \bigcup_{k=1}^r \overline{I_k}$ and moreover that $x \in D_\alpha \implies x \in {I_k}^\circ$ for some $1 \leq k \leq r$. We now construct a partition Q of I such that WLOG the subcells within Q are given by $\{\overline{I_1}, \dots, \overline{I_r}, I_{r+1}, \dots, I_N\}$ where each subcells is obviously closed. It remains obvious that $D_\alpha \subset \bigcup_{k=1}^r I_k$ as shown just above, and thus we have property (i). For property (ii), we use the fact that each I_k has content to obtain

$$\sum_{k=1}^r c(\overline{I_k}) = \sum_{k=1}^r (c({I_k}^\circ) + c(\partial I_k)) = \sum_{k=1}^r c({I_k}^\circ) \leq \alpha$$

from our construction via the content-zero definition and $c(\partial I_k) = 0$. For property (iii), note that $\overline{J_{k>r}} \cap J_{j \leq r} = \emptyset$ and therefore each $\overline{J_k}$ for $k > r$ is completely disjoint from D_α. Then from the definition of D_α, if such J_k is disjoint, then $\operatorname{osc}_{J_K} f < \alpha$ for each $r < k \leq N$.

We now have a partition P_α that satisfies (i)-(iii). Then we have that

$$\sum_{k=1}^N (M_k - m_k) c(J_k) = \sum_{k=1}^N \operatorname*{osc}_{J_k}(f) c(J_k) = \sum_{k=1}^r \operatorname*{osc}_{J_k}(f) c(J_k) + \sum_{k=r+1}^N \operatorname*{osc}_{J_k}(f) c(J_k)$$

The first of the two sums is directly related to property (ii), and since $\operatorname{osc}_{J_k} f \leq 2\|f\|_I$, for *any* J_k, then $\sum_{k=1}^r \operatorname{osc}_{J_k}(f) c(J_k) \leq 2\|f\|_I \sum_{k=1}^r c(J_k) \leq 2\|f\|_I \alpha$. By property (iii), the cells disjoint from D_α have $\operatorname{osc}_{J_k} f < \alpha$, and thus $\sum_{k=r+1}^N \operatorname{osc}_{J_k}(f) c(J_k) \leq \alpha \sum_{k=r+1}^N c(J_k) \leq \alpha c(I)$. Finally, we've arrived at

$$\sum_{k=1}^N \operatorname*{osc}_{J_k}(f) c(J_k) \leq 2\alpha \|f\|_I + \alpha c(I).$$

(b) Suppose D_α has content zero for each $\alpha > 0$. By a claim given in lecture, we have that for any D_α with $\alpha > 0$ there exists a partition P_α of I such that there exists some subcells $J_{k_1}, \dots, J_{k_r}$ in P_α such that

$$D_\alpha \subset \left(\bigcup_{i=1}^r J_{k_i} \right)^\circ \quad \text{and} \quad \sum_{i=1}^r c(J_{k_i}) < \alpha.$$

WLOG permute the ordering of subcells J of P_α such that if $\{J_1, \dots, J_N\}$ are the N subcells that form the partition P_α of I, then $J_{k_1} = J_1, \dots, J_{k_r} = J_r$. Accordingly, we have that $J_i \cap D_\alpha = \emptyset$ for $r < i \leq N$ and thus $\operatorname{osc}_{J_i} f < \alpha$. We now desire continuity of f over $I^* = I \setminus \left(\bigcup_{i=1}^r J_r\right)^\circ$ which is compact. Since we have that $\operatorname{osc}_{I^*} f < \alpha$, we have that

$$\lim_{\alpha \searrow 0} \sum_{i=1}^r c(J_i) = 0 \quad \text{and} \quad \lim_{\alpha \searrow 0} \operatorname*{osc}_{I^*} f = 0.$$

Therefore, since I is closed, f is bounded, $E = \lim_{\alpha \to 0+} \bigcup_{i=1}^r J_r$ has content zero, and $\lim_{\alpha \to 0+} \operatorname{osc}_{I^*} f = 0 \implies \operatorname{osc}_{I \setminus E} f = 0 \implies f$ is continuous over $I \setminus E$, then by Theorem 43.9, f is integrable over I.

(c) Suppose that $f \colon I \to \mathbb{R}$ is integrable over I. By the Riemann Criterion for Integrability, we have that for any $\epsilon > 0$ there exists a partition P_ϵ of I such that if $\{J_1, \dots, J_n\} \succ P_\epsilon$ then $\sum_{k=1}^n \operatorname{osc}_{J_k}(f) c(J_k) < \epsilon$. We are now going to split our analysis of the intersections of such J_k with D_α into 3 disjoint and exhaustive cases. We have that either (i) D_α intersects the interior of J_k, (ii) D_α never intersects the closure of J_k, and (iii) D_α intersects the boundary of J_k but not the interior.

$$\sum_{k=1}^n \operatorname*{osc}_{J_k}(f) c(J_k) = \sum_{\substack{k \\ {J_k}^\circ \cap D_\alpha \neq \emptyset}} \operatorname*{osc}_{J_k}(f) c(J_k) + \sum_{\substack{k \\ \overline{J_k} \cap D_\alpha = \emptyset}} \operatorname*{osc}_{J_k}(f) c(J_k) + \sum_{\substack{k \\ {J_k}^\circ \cap D_\alpha = \emptyset \\ \partial J_k \cap D_\alpha \neq \emptyset}} \operatorname*{osc}_{J_k}(f) c(J_k) < \alpha\epsilon$$

We've also chosen WLOG to multiply the right side of the inequality by α which makes some of the proceeding analysis a little nicer. We know from the integrability criterion that in order for the left-hand sum to vanish, each of the consitutent sums in the middle must also vanish.
$({J_k}^\circ \cap D_\alpha \neq \emptyset)$ We know that each $\operatorname{osc}_{J_k}(f) \geq \alpha$ in this case since ${J_k}^\circ$ contains points from D_α, which has this oscillation condition. Then we have that

$$\alpha \sum_{\substack{k \\ {J_k}^\circ \cap D_\alpha \neq \emptyset}} c(J_k) \leq \sum_{\substack{k \\ {J_k}^\circ \cap D_\alpha \neq \emptyset}} \operatorname*{osc}_{J_k}(f) c(J_k) < \alpha\epsilon$$
$$\sum_{\substack{k \\ {J_k}^\circ \cap D_\alpha \neq \emptyset}} c(J_k) < \epsilon$$

and thus the content of such J_k converges to zero as $\epsilon \to 0$. Therefore, the "bulk majority", or interior, of D_α has content zero. We will prove the remaining portion of D_α has content zero in the last sum.
$(\overline{J_k} \cap D_\alpha = \emptyset)$ J_k contains zero points from D_α and thus we know that each $\operatorname{osc}_{J_k}(f) < \alpha$. Of course, as $\epsilon = \alpha\epsilon \to 0$, this term of the sum vanished and therefore the same vanishes as well.
$({J_k}^\circ \cap D_\alpha = \emptyset$ and $\partial J_k \cap D_\alpha \neq \emptyset)$ If the boundary of J_k contains a dense subset of D_α, we must consider that it is possible to have $\operatorname{osc}_{\partial J_k}(f) \geq \alpha$. In this case though, we have that

$$\operatorname*{osc}_{J_k}(f) c(\partial J_k) = \operatorname*{osc}_{\partial J_k}(f) c(\partial J_k) = 0$$

which follows from $\operatorname{osc}_{J_k} f = \operatorname{osc}_{\partial J_k} f$ and the fact that J_k has content implies that the ∂J_k has content zero. Therefore this sum term completely vanishes as desired.
We've shown that each disjoint and exhaustive sum term of the sum $\sum_{k=1}^{n} \operatorname{osc}_{J_k}(f) c(J_k)$ converges to zero as $\alpha\epsilon \to 0$. Then, we have shown that since f is integrable over I, it must satisfy the Riemann Criterion for Integrability, which bounds the content of each D_α to zero for each $\alpha > 0$.