CDS 202 Winter 2009 Solution Set 3

Problem 1 (MTA 2.3-3) (PS 2008)

(i) Using the comparison test and the fact that $\|A^n\| \leq \|A\|^n$, it is clear that $\sum_{n=0}^{\infty} \frac{A^n}{n!}$ converges absolutely, and so converges. Since

$$\left\| \sum_{n=0}^{N} \frac{A^n}{n!} \right\| \leq \sum_{n=0}^{N} \frac{\|A\|^n}{n!} \leq e^{\|A\|}$$

it follows that $\|e^A\| \leq e^{\|A\|}$.

(ii) Since A and B commute, we don't have to worry about the ordering of these operators in the series expansion of e^{A+B}, and so this just reduces to the usual proof that $e^{A+B} = e^A e^B$. It follows that

$$e^A e^{-A} = e^{A+(-A)} = e^0 = I,$$

and so $e^{-A} = (e^A)^{-1}$.

(iii) Expand e^{A+H} out into monomials in A, H, and rearrange by grouping all terms containing n H's together (there is no problem with rearranging the terms since the series is absolutely convergent). Then

$$e^{A+H} = \sum_{n=0}^{\infty} \psi_n(A) \cdot H^n$$

where $\psi_n(A) \in L_s^n(E, E)$ is defined by

$$\psi_n(A) \cdot H^n = \sum_{p=n}^{\infty} \frac{\text{sum of all ordered monomials containing } n \text{ } H\text{'s and } p-n \text{ } A\text{'s}}{p!},$$

(Note: it is enough to define $\psi_n(A) \in L_s^n(E, E)$ on the diagonal $H^n = (H, H, \ldots, H)$ by the results of Supplement 2.2B).

So for example,

$$\psi_1(A) \cdot H = H + \frac{HA + AH}{2!} + \frac{HAA + AHA + AAH}{3!} + \ldots$$

To show that $\psi_n(A)$ is actually bounded, just note that

$$\begin{aligned} \|\psi_n(A) \cdot H^n\| &\leq \sum_{p=n}^{\infty} \frac{\binom{p}{n} \|H\|^n \|A\|^{p-n}}{p!} \\ &= \frac{\|H\|^n}{n!} \sum_{p=n}^{\infty} \frac{\|A\|^{p-n}}{(p-n)!} \\ &= \frac{\|H\|^n}{n!} e^{\|A\|}. \end{aligned}$$

This shows $\psi_n(A)$ is bounded when considered as an element of $S^n(E, E)$, and hence as an element of $L^n_s(E, E)$ (cf Supplement 2.2B). Using the above bound, we get that

$$\sum_{n=0}^{\infty} \|\psi_n(A) \cdot H^n\| \leq \sum_{n=0}^{\infty} \frac{\|H\|^n}{n!} e^{\|A\|} = e^{\|A\|} e^{\|H\|},$$

so the series for e^{A+H} converges absolutely, hence converges, and $e^{(\cdot)}$ is analytic at any A, with derivatives

$$D^n \exp(A) \cdot H^n = n!\ \psi_n(A) \cdot H^n.$$

(iv) From

$$e^H = H + \frac{H^2}{2!} + \frac{H^3}{3!} + \dots$$

it is clear that $D\exp(0) = I$, which is invertible. The inverse function theorem then guarantees the existence of a unique inverse in a neighborhood of the origin.

(v) The series is absolutely convergent, hence convergent, for $\|I - A\| < 1$. To show that it equals log:

Differentiate $\exp(\log B) = B$ wrt B to get

$$D\exp(\log B) \circ D\log(B) = \mathrm{Id}_E \Rightarrow D\log(I) = [D\exp(0)]^{-1} = \mathrm{Id}_E.$$

Differentiate again

$$D^2 \exp(\log B)(D\log(B) \cdot H_1, D\log(B) \cdot H_2) + D\exp(\log B)(D^2 \log(B) \cdot (H_1, H_2)) = 0$$

and evaluate at $B = I,\ H_1 = H_2 = H$ to get

$$D^2 \log(I) \cdot (H, H) = -D^2 \exp(0) \cdot (H, H) = -H^2,$$

and so on. The Taylor series for log about I is

$$\begin{aligned}\log(I + H) &= \log(I) + D\log(I) \cdot H + \frac{1}{2!} D^2 \log(I) \cdot (H, H) + \dots \\ &= 0 + H - \frac{H^2}{2} + \dots\end{aligned}$$

which is the same as the series given in the question.

(vi) Let $X = \log A,\ Y = \log B$. From the series expression for log it's clear that

$$AB = BA \Rightarrow XY = YX.$$

From part (ii), we then have that

$$AB = e^X e^Y = e^{X+Y} = e^{\log A + \log B}.$$

Taking the log of both sides (assuming $\log(AB)$ exists) we get that

$$\log AB = \log A + \log B.$$

Problem 2 (MTA 2.3-4) (PS 2008)

Let $f : U \subset \mathbf{E} \to \mathbf{F}$ be of class C^r, $r \geq 1$, $u_0 \in U$ open, and suppose that $\mathbf{D}f(u_0)$ is a linear isomorphism. Define $g : U \times \mathbf{F} \to \mathbf{F}$ by $g(u, v) = f(u) - v$. Then g is easily seen to be C^r with nonsingular partial derivative $\mathbf{D}_1 g(u_0, v_0) = \mathbf{D}f(u_0) : \mathbf{E} \to \mathbf{F}$ for any $v_0 \in \mathbf{F}$. By the implicit function theorem, there exist open nbhds V_0 of v_0, W_0 of $g(u_0, v_0)$, and a unique C^r function $h : V_0 \times W_0 \to U$ with $h(v_0, g(u_0, v_0)) = u_0$ such that

$$g(h(v, w), v) = w \qquad \forall v \in V_0,\ w \in W_0.$$

In particular, take $v_0 = f(u_0)$. Then W_0 is a nbhd of $0 = g(u_0, f(u_0))$, and so

$$\tilde{h} : V_0 \to U, \quad \tilde{h}(v) = h(v, 0),$$

is well defined, and C^r (since it is the restriction of a C^r function). So we have that

$$g(\tilde{h}(v), v) = 0 \quad \Rightarrow \quad v = f(\tilde{h}(v)), \qquad \forall v \in V_0.$$

The above equation implies that

$$\tilde{h} : V_0 \to \tilde{h}(V_0), \qquad f|_{\tilde{h}(V_0)} : \tilde{h}(V_0) \to V_0,$$

are both 1-1 and onto, and inverses of one another. Hence f is a C^r diffeomorphism on the nbhd $\tilde{h}(V_0)$ of u_0 . Moreover, the chain rule applied to $f(\tilde{h}(v)) = v$ implies that

$$\mathbf{D}\tilde{h}(v) = \left[\mathbf{D}f(\tilde{h}(v))\right]^{-1},$$

for v in the nbhd V_0 of $v_0 = f(u_0)$.

Problem 3 (MTA 2.5-12)

Let $F(a_{n-1}, \cdots, a_0, \lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$ be the characteristic polynomial of an operator on $\mathbb{R}^n$. This map is smooth because it is just a polynomial in $a_{n-1}, \cdots, a_0, \lambda$. Now suppose that we fix $a_{n-1}, \cdots, a_0$ and let λ_0 (which depends on $a_{n-1}, \cdots, a_0$) be a simple eigenvalue, i.e.,

$$F(a_{n-1}, \cdots, a_0, \lambda) = (\lambda - \lambda_0)G(a_{n-1}, \cdots, a_0, \lambda),$$

where $G(a_{n-1}, \cdots, a_0, \lambda_0) \neq 0$. Then

$$\frac{\partial F}{\partial \lambda}(\lambda_0) = G(a_{n-1}, \cdots, a_0, \lambda_0) \neq 0.$$

Hence, using the implicit function theorem, we conclude that in a neighborhood of λ_0, there is a unique smooth map $g(a_{n-1}, \cdots, a_0)$ such that

$$F(a_{n-1}, \cdots, a_0, g(a_{n-1}, \cdots, a_0)) = 0.$$

This means that the eigenvalue is a smooth function of $a_{n-1}, \cdots, a_0$. Since each a_i is a polynomial of elements of operator matrix, each is a smooth function of the operator. From this we see that the eigenvalue is a smooth function of the operator. So, we conclude that simple eigenvalues of operators on $\mathbb{R}^n$ are smooth functions of the operator.

Problem 4 (MTA 3.5-1 (i), show that $\mathbf{O(n)}$ is a manifold of dimension $\mathbf{n(n-1)/2}$) (PS 2009)

(i) Consider the determinant map $\det : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$. Then $SL(n, \mathbb{R}) = \det^{-1}\{1\}$. So if we can show that 1 is a regular value of det, it will follow from the Submersion Theorem that $SL(n, \mathbb{R})$ is a closed submanifold of $L(\mathbb{R}^n, \mathbb{R}^n)$ of dimension $\dim \ker(T_A \det) = \dim L(\mathbb{R}^n, \mathbb{R}^n) - \dim \mathbb{R} = n^2 - 1$.

To find $T_A \det$, where $A \in SL(n, \mathbb{R})$, note that

$$\begin{aligned}\det(A + H) - \det(A) &= \det(A)(\det(I + A^{-1}H) - 1) \\ &= \det(A)(\mathrm{trace}(A^{-1}H) + o(H)) \\ &= \det(A)\mathrm{trace}(A^{-1}H) + o(H)\end{aligned}$$

for any $H \in L(\mathbb{R}^n, \mathbb{R}^n)$ so

$$T_A \det(H) = \det(A)\mathrm{trace}(A^{-1}H) = \mathrm{trace}(A^{-1}H).$$

This map is clearly surjective onto $\mathbb{R}$ (e.g. at any $A \in SL(n, \mathbb{R})$, take $H = \frac{c}{n}A$ to get $T_A \det(H) = c$). Therefore 1 is a regular value of det.

For the complex case, the analysis is the same except that the (real) dimension is

$$\dim_{\mathbb{R}} \ker(T \det) = \dim_{\mathbb{R}} L(\mathbb{C}^n, \mathbb{C}^n) - \dim_{\mathbb{R}} \mathbb{C} = 4n^2 - 2.$$

(ii) Define the map $f : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathrm{Symm}(n)$ by $f(A) = A^T A$, where $\mathrm{Symm}(n)$ denotes the (vector) space of symmetric $n \times n$ matrices. Then $O(n) = f^{-1}(I)$. If we can show that I is a regular point of f, then again using the Submersion Theorem it will follow that $O(n)$ is a closed submanifold of $L(\mathbb{R}^n, \mathbb{R}^n)$ of dimension

$$\dim\ker Tf = \dim L(\mathbb{R}^n, \mathbb{R}^n) - \dim \mathrm{Symm}(n) = n^2 - n(n+1)/2 = n(n-1)/2.$$

To find $T_A f$

$$\begin{aligned} f(A+H) - f(A) &= A^T A + A^T H + H^T A + H^T H - A^T A \\ &= A^T H + H^T A + o(H) \end{aligned}$$

so $T_A f(H) = A^T H + H^T A$.

$T_A f$ is surjective at any $A \in O(n)$ since e.g. if $S \in \mathrm{Symm}(n)$, then $T_A f(\frac{1}{2} AS) = S$. So I is a regular value of f.

Problem 5 (Nawaf Bou-Rabee)

Solution (a) Suppose f and g are smooth. df and dg are injective since f and g are immersions. Let $f, g : M_{f,g} \to N_{f,g}$, consider a coordinate chart of $M_f \times M_g$ $(\psi_f \times \psi_g, V_f \times V_g)$ and $(\psi_f \times \psi_g, V_f \times V_g)$ of $N_f \times N_g$. Then,

$$\begin{aligned} d(f \times g)_p &= d((\psi_f^{-1} \circ \hat{f} \circ \phi_f) \times (\psi_g^{-1} \circ \hat{g} \circ \phi_g))_p \\ &= d(\psi_f^{-1} \circ \hat{f} \circ \phi_f)_p \times d(\psi_g^{-1} \circ \hat{g} \circ \phi_g)_p \\ &= (d(\psi_f^{-1}) \circ d(\hat{f}) \circ d(\phi_f))_p \times (d(\psi_g^{-1}) \circ d(\hat{g}) \circ d(\phi_g))_p \\ &= df_p \times dg_p \end{aligned}$$

The steps follow by independence of g from f under the cross product (which permits us to distribute the derivative) and the chain rule which we can use since f and g are smooth. Thus, $dw = d(f \times g) = df \times dg$ is injective since $dw = df \times dg$ is simply a block diagonal matrix with df and dg along the diagonals.

To show this we use a basic result from linear algebra: a linear transformation is injective if and only if its nullspace is trivial. Since df and dg are the block-diagonal terms of dw, the kernel of dw is trivial if and only if the nullspaces of df and dg are trivial,i.e., the zero element in the cross product space is the only term that df and dg map to zero (since they are injective), and hence, the zero element is the only term that dw maps to zero. Thus, dw is injective.

Solution (b) Since f and g are immersions, by the chain rule, the following composition is a product of linear injections:

$$\begin{aligned} d(f \circ g)_p &= d(\psi_f^{-1} \circ \hat{f} \circ \phi_f) \circ (\psi_g^{-1} \circ \hat{g} \circ \phi_g)_p \\ &= (d\psi_f^{-1} \circ d\hat{f} \circ d\phi_f)_{g(p)} \circ (d\psi_g^{-1} \circ d\hat{g} \circ d\phi_g)_p \\ &= df_{g(p)} \circ dg_p \end{aligned}$$

which we will show is a linear injection. Since the composition is defined only when the range of g equals the domain of f, the dimensions of the range of dg_p and domain of $df_{g(p)}$ are the same. Thus, the fact that these are linear injections implies a map of the form: $df_{g(p)} \circ dg_p : \mathbb{R}^k \mapsto \mathbb{R}^l \mapsto \mathbb{R}^m$ where l is the dimension of the domain of $df_{g(p)}$ and the range of dg_p and $k \leq l \leq m$ since each step in the map is injective. Since the range of dg_p is a subset of $\mathbb{R}^l$, the image of this vector subspace is a restriction of $df_{g(p)}$ on a vector subspace. The restriction of a linear injection on a vector subspace is still a linear injection since distinct elements in the image of the subspace will correspond to distinct elements within the subspace because the subspace is subset of the domain of the injection. Thus, the composition is a linear injection.

Solution (c) The restriction of f to a submanifold of its domain is simply the composition of f with the inclusion map γ defined in Problem 2 which we showed is an immersion. Thus, $f|_U$ is a composition of f with γ which is an immersion since by part (b) the composition of immersions is an immersion. Equivalently, if df is injective then its restriction to any vector subspace (defined by the restriction to a submanifold) is also injective.

Solution (d) If the dimensions of M and N are the same, df_p injective implies df_p is onto. Since a matrix is invertible if and only if it is a bijective map, df_p is invertible. By the inverse function theorem, f is a local diffeomorphism at p.

Problem 6 (MTA 3.5-11) (PS 2009)

For any $m \in f^{-1}(n)$, $T_m f$ is surjective and $\dim M = \dim N < \infty$ together imply that $T_m f$ is bijective. By the Inverse Function Theorem f is a local diffeomorphism about m. So for each $m \in f^{-1}(n)$ there is an open nbhd O_m of m such that $f|_{O_m} : O_m \to f(O_m)$ is a diffeomorphism. Note that m is the only element of $f^{-1}(n)$ contained in O_m, since f must be injective when restricted to O_m. The collection $\{O_m\}_{m \in f^{-1}(n)}$ is an open cover of $f^{-1}(n)$. Since f is continuous, $f^{-1}(n)$ is a closed subset of the compact space M, and hence is itself compact. Therefore there exists a finite subcover $\{O_{m_1}, O_{m_2}, \dots, O_{m_k}\}$ of $f^{-1}(n)$. Since each O_{m_i} contains precisely one element of $f^{-1}(n)$, it follows that $f^{-1}(n)$ is finite (it equals $\{m_1, m_2, \dots, m_k\}$).

Since M is Hausdorff, by induction we can construct k mutually disjoint open nbhds of the points $m_1, m_2, \ldots, m_k$. Intersecting each such nbhd with the corresponding O_{m_i} gives k open nbhds $V_1, V_2, \ldots, V_k$ of $m_1, m_2, \ldots, m_k$ such that $f|_{V_i} : V_i \to f(V_i)$ is a diffeomorphism. Define the following sets

$$\begin{aligned} V &= \bigcap_{i=1}^{k} f(V_i) - f(M - \bigcup_{i=1}^{k} V_i) \\ U_i &= V_i \cap f^{-1}(V) \qquad i = 1, 2, \ldots, k. \end{aligned}$$

We claim that these are the required sets from the question. First note that V is open: $M - \cup_{i=1}^{k} V_i$ is closed in M, hence compact. f is continuous, so $f(M - \cup_{i=1}^{k} V_i)$ is a compact subset of the Hausdorff space N, hence it is closed. Then V can be written

$$V = \bigcap_{i=1}^{k} f(V_i) \bigcap \left(N - f(M - \bigcup_{i=1}^{k} V_i) \right).$$

This is a finite intersection of open sets, hence open.

It follows that the U_i are all open. We know that f is a local diffeomorphism on V_i, hence $f|_{U_i} : U_i \to V$ is a diffeomorphism, since $U_i \subset V_i$.

It only remains to prove that $f^{-1}(V) = \cup_{i=1}^{k} U_i$. Clearly $\cup_{i=1}^{k} U_i \subset f^{-1}(V)$. Suppose there exists a point $x \in f^{-1}(V) - \cup_{i=1}^{k} U_i$. Then

$$\begin{aligned} & f(x) \in V, \ x \notin \cup_{i=1}^{k} U_i \\ \Rightarrow & f(x) \notin f(M - \cup_{i=1}^{k} V_i), \ x \notin \cup_{i=1}^{k} U_i \\ \Rightarrow & x \in \cup_{i=1}^{k} V_i, \ x \notin \cup_{i=1}^{k} U_i \\ \Rightarrow & x \in V_i - U_i \quad \text{some } i \text{, since the } V_i \text{ are all disjoint.} \end{aligned}$$

Since F maps V_i diffeomorphically onto $f(V_i)$ and $U_i \subset V_i$ diffeomorphically onto V, it follows that $f(x) \notin V$, which contradicts the above.

Now suppose the M is connected, and f is a submersion. For each $k \in \mathbb{N}$, define the set

$$A_k = \{n \in N : f^{-1}(n) \text{ has precisely } k \text{ points}\}.$$

Clearly all of the A_k are disjoint, and $f(M) = \bigcup_{k \in \mathbb{N}} A_k$. From the previous result, for each point in A_k we can find an open nbhd of the point contained in A_k. Hence A_k is open (in N and therefore in $f(M)$). Also A_k is closed in $f(M)$ since $A_k = f(M) - \bigcup_{j \in \mathbb{N}, j \neq k} A_j$. Since M is connected and f is continuous, $f(M)$ is connected. So the only possibilities are that $A_k = f(M)$ or $A_k = \emptyset$. So if one point of $f(M)$ has an k-point inverse image, then every point does.