10 APPENDIX: Technical results

10.1 The inverse function theorem

Lemma 10.1 *(Contraction mapping principle) Let M be a complete metric space and suppose $T : M \to M$ is a map such that*

$$d(Tx, Ty) \leq kd(x, y)$$

where $k < 1$. Then T has a unique fixed point.

Proof: Choose any point x_0, then

$$\begin{aligned} d(T^m x_0, T^n x_0) &\leq k^m d(x_0, T^{n-m} x_0) \qquad \text{for} \quad n \geq m \\ &\leq k^m (d(x_0, Tx_0) + d(Tx_0, T^2 x_0) + \ldots + d(T^{n-m-1} x_0, T^{n-m} x_0)) \\ &\leq k^m (1 + k + \ldots + k^{n-m-1}) d(x_0, Tx_0) \\ &\leq \frac{k^m}{1-k} d(x_0, Tx_0) \end{aligned}$$

This is a Cauchy sequence, so completeness of M implies that it converges to x. Thus $x = \lim T^n x_0$ and so by continuity of T,

$$Tx = \lim T^{n+1} x_0 = x$$

For uniqueness, if $Tx = x$ and $Ty = y$, then

$$d(x, y) = d(Tx, Ty) \leq kd(x, y)$$

and so $k < 1$ implies $d(x, y) = 0$. □

Theorem 10.2 *(Inverse function theorem) Let $U \subseteq \mathbf{R}^n$ be an open set and $f : U \to \mathbf{R}^n$ a C^∞ function such that Df_a is invertible at $a \in U$. Then there exist neighbourhoods V, W of a and $f(a)$ respectively such that $f(V) = W$ and f has a C^∞ inverse on W.*

Proof: By an affine transformation $x \mapsto Ax + b$ we can assume that $a = 0$ and $Df_a = I$. Now consider $g(x) = x - f(x)$. By construction $Dg_0 = 0$ so by continuity there exists $r > 0$ such that if $\|x\| < 2r$,

$$\|Dg_x\| < \frac{1}{2}$$

It follows from the mean value theorem that

$$\|g(x)\| \leq \frac{1}{2}\|x\|$$

and so g maps the closed ball $\bar{B}(0,r)$ to $\bar{B}(0,r/2)$. Now consider

$$g_y(x) = y + x - f(x)$$

(The choice of g_y is made so that a fixed point $g_y(x) = x$ solves $f(x) = y$).

If now $\|y\| \leq r/2$ and $\|x\| \leq r$, then

$$\|g_y(x)\| \leq \frac{1}{2}r + \|g(x)\| \leq \frac{1}{2}r + \frac{1}{2}r = r$$

so g_y maps the complete metric space $M = \bar{B}(0,r)$ to itself. Moreover

$$\|g_y(x_1) - g_y(x_2)\| = \|g(x_1) - g(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|$$

if $x_1, x_2 \in \bar{B}(0,r)$, and so g_y is a contraction mapping. Applying Lemma 1 we have a unique fixed point and hence an inverse $\varphi = f^{-1}$.

We need to show first that φ is continuous and secondly that it has derivatives of all orders. From the definition of g and the mean value theorem,

$$\begin{aligned}\|x_1 - x_2\| &\leq \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\| \\ &\leq \|f(x_1) - f(x_2)\| + \frac{1}{2}\|x_1 - x_2\|\end{aligned}$$

so

$$\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|$$

which is *continuity* for φ. It follows also from this inequality that if $y_1 = f(x_1)$ and $y_2 = f(x_2)$ where $y_1, y_2 \in B(0,r/2)$ then $x_1, x_2 \in \bar{B}(0,r)$, and so

$$\begin{aligned}\|\varphi(y_1) - \varphi(y_2) - (Df_{x_2})^{-1}(y_1 - y_2)\| &= \|x_1 - x_2 - (Df_{x_2})^{-1}(f(x_1) - f(x_2))\| \\ &\leq \|(Df_{x_2})^{-1}\|\|Df_{x_2}(x_1 - x_2) - f(x_1) + f(x_2)\| \\ &\leq A\|x_1 - x_2\|R\end{aligned}$$

where A is a bound on $\|(Df_{x_2})^{-1}\|$ and the function $\|x_1 - x_2\|R$ is the remainder term in the definition of differentiability of f. But $\|x_1 - x_2\| \leq 2\|y_1 - y_2\|$ so as $y_1 \to y_2$, $x_1 \to x_2$ and hence $R \to 0$, so φ is differentiable and moreover its derivative is $(Df)^{-1}$.

Now we know the derivative of φ:

$$D\varphi = (Df)^{-1}$$

so we see that it is continuous and has as many derivatives as f itself, so φ is C^∞. □

10.2 Existence of solutions of ordinary differential equations

Lemma 10.3 *Let M be a complete metric space and $T : M \to M$ a map. If T^n is a contraction mapping, then T has a unique fixed point.*

Proof: By the contraction mapping principle, T^n has a unique fixed point x. We also have

$$T^n(Tx) = T^{n+1}x = T(T^n x) = Tx$$

so Tx is also a fixed point of T^n. By uniqueness $Tx = x$. □

Theorem 10.4 *Let $f(t, x)$ be a continuous function on $|t - t_0| \leq a$, $\|x - x_0\| \leq b$ and suppose f satisfies a Lipschitz condition*

$$\|f(t, x_1) - f(t, x_2)\| \leq \|x_1 - x_2\|.$$

If $M = \sup |f(t, x)|$ and $h = \min(a, b/M)$, then the differential equation

$$\frac{dx}{dt} = f(t, x), \qquad x(t_0) = x_0$$

has a unique solution for $|t - t_0| \leq h$.

Proof: Let

$$(Tx)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))ds$$

Then Tx is differentiable since f and x are continuous and if $Tx = x$, x satisfies the differential equation (differentiate the definition). We use the metric space

$$X = \{x \in C([t_0 - h, t_0 + h], \mathbf{R}^n) : \|x(t) - x_0\| \leq Mh\}$$

with the uniform metric

$$d(x_1, x_2) = \sup_{|t - t_0| \leq h} \|x_1(t) - x_2(t)\|$$

which makes it complete. If $x \in M$, then $Tx \in M$ and we claim

$$|T^k x_1(t) - T^k x_2(t)\| \leq \frac{c^k}{k!}|t - t_0|^k d(x_1, x_2)$$

For $k = 0$ this is clear, and in general we use induction to establish:

$$\begin{aligned}
\|T^k x_1(t) - T^k x_2(t)\| &\leq \int_{t_0}^{t} \|f(s, T^{k-1}x_1(s) - f(s, T^{k-1}x_2(s)\|ds \\
&\leq c\int_{t_0}^{t} \|T^{k-1}x_1(s) - T^{k-1}x_2(s)\|ds \\
&\leq (c^k/(k-1)!) \int_{t_0}^{t} |s - t_0|^{k-1} ds \; d(x_1, x_2) \\
&\leq (c^k/k!)|t - t_0|^k d(x_1, x_2)
\end{aligned}$$

So T^n is a contraction mapping for large enough N, and the result follows. □

Theorem 10.5 *The solution above depends continuously on the initial data x_0.*

Proof: Take $h_1 \leq h$ and $\delta > 0$ such that $Mh + \delta \leq b$, and let

$$Y = \{y \in C([t_0 - h_1, t_0 + h_1] \times \bar{B}(x_0, \delta); \mathbf{R}^n : \|y(t, x) - x\| \leq Mh, y(t_0, x) = x\}$$

which is a complete metric space as before. Now set

$$(Ty)(t, x) = x + \int_{t_0}^{t} f(s, y(s, x))ds$$

Since $Mh_1 + \delta \leq b$, T maps Y to Y and just as before T^n is a contraction mapping with a unique fixed point which satisfies

$$\frac{\partial y}{\partial t} = f(t, y), \qquad y(t_0, x) = x$$

Since y is continuous in t and x this is what we need. □

If $f(t, x)$ is smooth then we need more work to prove that the solution to the equation is smooth and smoothly dependent on parameters.

10.3 Smooth dependence

Lemma 10.6 *Let $A(t, x), B(t, x)$ be continuous matrix-valued functions and take $M \geq \sup_{t,x} \|B\|$. The solutions of the linear differential equations*

$$\begin{aligned}
\frac{d\xi(t, x)}{dt} &= A(t, x)\xi(t, x), \qquad \xi(t_0, x) = a(x) \\
\frac{d\eta(t, x)}{dt} &= B(t, x)\eta(t, x), \qquad \eta(t_0, x) = b(x)
\end{aligned}$$

satisfy

$$\sup_x \|\xi(t,x) - \eta(t,x)\| \le C\|A - B\| \frac{e^{M|t-t_0|} - 1}{M} + \|a - b\| e^{M|t-t_0|}$$

where C is a constant depending only on A and a.

Proof: By the existence theorem we know how to find solutions as limits of ξ_n, η_n where

$$\begin{aligned} \xi_k &= a + \int_{t_0}^{t} A\xi_{k-1} ds \\ \eta_k &= b + \int_{t_0}^{t} B\eta_{k-1} ds \end{aligned}$$

Let $g_k(t) = \sup_x \|\xi_k(t,x) - \eta_k(t,x)\|$ and $C = \sup_{k,x,t} \|\xi_k\|$. Then

$$g_n(t) \le \|a - b\| + C\|A - B\||t - t_0| + M \int_{t_0}^{t} g_{n-1}(s) ds$$

Now define f_n by $f_0(t) = \|a - b\|$ and then inductively by

$$f_n(t) = \|a - b\| + C\|A - B\||t - t_0| + M \int_{t_0}^{t} f_{n-1}(s) ds$$

Comparing these two we see that $f_n \ge g_n$. This is a contraction mapping, so that $f_n \to f$ with

$$f(t) = \|a - b\| + C\|A - B\||t - t_0| + M \int_{t_0}^{t} f(s) ds$$

and solving the corresponding differential equation we get

$$f(t) = \|a - b\| e^{M|t-t_0|} + C\|A - B\| \frac{e^{M|t-t_0|} - 1}{M}$$

As $g_n(t) \le f_n(t)$,

$$\sup_x \|\xi_n(t,x) - \eta_n(t,x)\| \le f_n(t)$$

and the theorem follows by letting $n \to \infty$. □

Theorem 10.7 *If f is C^k and*

$$\frac{d}{dt}\alpha(t,x) = f(t, \alpha(t,x)), \qquad \alpha(0,x) = x$$

then α is also C^k.

Proof: The hardest bit is $k = 1$. Assume f is C^1 so that $\partial f/\partial t$ and $\partial f/\partial x_i$ exist and are continuous. We must show that α is C^1 in all variables. If that were true, then the matrix valued function λ where $(\lambda_i = \partial\alpha/\partial x_i)$ would be the solution of the differential equation

$$\frac{d\lambda}{dt} = D_x f(t, \alpha)\lambda \tag{25}$$

so we shall solve this equation by the existence theorem and prove that the solution is the derivative of α. Let $F(s) = f(t, a + s(b - a))$. Then

$$\frac{dF}{ds} = D_x f(t, a + s(b - a))(b - a)$$

so

$$f(t, b) - f(t, a) = \int_0^1 D_x f(t, a + s(b - a))(b - a)ds$$

But then

$$\begin{aligned}\frac{d}{dt}(\alpha(t, x + y) - \alpha(t, x)) &= f(t, \alpha(t, x + y)) - f(t, \alpha(t, x)) \\ &= \int_0^1 D_x f(t, \alpha(t, x) + s(\alpha(t, x + y) - \alpha(t, x)))(\alpha(t, x + y) - \alpha(t, x))ds\end{aligned}$$

Let $A(t, x) = D_x f(t, \alpha(t, x))$ and $\xi(t, x) = \lambda(t, x)y$ and

$$B_y(t, x) = \int_0^1 D_x f(t, \alpha(t, x)+s(\alpha(t, x+y)-\alpha(t, x)))ds, \quad \eta_y(t, x) = \alpha(t, x+y)-\alpha(t, x)$$

Apply the previous lemma and we get

$$\sup_{|t|\leq\epsilon} \|\lambda(t, x)y - (\alpha(t, x + y) - \alpha(x))\| = o(\|y|)$$

and so $D_x\alpha = \lambda$, which is continuous in (t, x). Since also $d\alpha/dt = f(t, \alpha)$ this means that α is C^1 in all variables.

To continue, suppose inductively that the theorem is true for $k - 1$, and f is C^k. Then $A(t, x) = D_x f(t, \alpha(t, x))$ is C^{k-1} but since

$$\frac{d\lambda}{dt} = A\lambda$$

we have λ is C^{k-1}. Now $D_x\alpha = \lambda$ so the x_i-derivatives of α are C^{k-1}. But also $d\alpha/dt = f(t, \alpha)$ is C^{k-1} too, so α is C^k. $\square$

10.4 Partitions of unity on general manifolds

Definition 39 *A partition of unity on M is a collection $\{\varphi_i\}_{i\in I}$ of smooth functions such that*

- $\varphi_i \geq 0$
- $\{\operatorname{supp}\varphi_i : i \in I\}$ *is locally finite*
- $\sum_i \varphi_i = 1$

Here *locally finite* means that for each $x \in M$ there is a neighbourhood U which intersects only finitely many supports $\operatorname{supp}\varphi_i$.

Theorem 10.8 *Given any open covering $\{V_\alpha\}$ of M there exists a partition of unity $\{\varphi_i\}$ on M such that* $\operatorname{supp}\varphi_i \subset V_{\alpha(i)}$ *for some $\alpha(i)$.*

Proof: (by exhaustion – !)

1. M is locally compact since each $x \in M$ has a neighbourhood homeomorphic to, say, the open unit ball in $\mathbf{R}^n$. So take U homeomorphic to a smaller ball, then $\bar{U}$ is compact. Since M is Hausdorff, $\bar{U}$ is closed (compact implies closed in Hausdorff spaces).

2. M has a countable basis of open sets $\{U_j\}_{j\in\mathbf{N}}$, so $x \in U_j \subset U$ and $\bar{U}_j \subset \bar{U}$ is compact so M has a countable basis of open sets with $\bar{U}_j$ compact.

3. Put $G_1 = U_1$. Then

$$\bar{G}_1 \subset \bigcup_{j=1}^{\infty} U_j$$

so by compactness there is $k > 1$ such that

$$\bar{G}_1 \subset \bigcup_{j=1}^{k} U_j = G_2$$

Now take the closure of G_2 and do the same. We get compact sets $\bar{G}_j$ with

$$\bar{G}_j \subset G_{j+1} \qquad M = \bigcup_{j=1}^{\infty} U_j$$

4. By construction we have

$$\bar{G}_j \backslash G_{j-1} \subset G_{j+1} \backslash \bar{G}_{j-2}$$

and the set on the left is compact and the one on the right open. Now take the given open covering $\{V_\alpha\}$. The sets $V_\alpha \cap (G_{j+1} \backslash \bar{G}_{j-2})$ cover $\bar{G}_j \backslash G_{j-1}$. This latter set is compact so take a finite subcovering, and then proceed replacing j with $j+1$. This process gives a countable locally finite *refinement* of $\{V_\alpha\}$, i.e. each $V_\alpha \cap (G_{j+1} \backslash \bar{G}_{j-2})$ is an open subset of V_α. It is locally finite because

$$G_{j+1} \backslash \bar{G}_{j-2} \cap G_{j+4} \backslash \bar{G}_{j+1} = \emptyset$$

5. For each $x \in M$ let j be the largest natural number such that $x \in M \backslash \bar{G}_j$. Then $x \in V_\alpha \cap (G_{j+2} \backslash \bar{G}_{j-1})$. Take a coordinate system within this open set and a bump function f which is identically 1 in a neighbourhood W_x of x.

6. The W_x cover $\bar{G}_{j+1} \backslash G_j$ and so as x ranges over the points of $G_{j+2} \backslash \bar{G}_{j-1}$ we get an open covering and so by compactness can extract a finite subcovering. Do this for each j and we get a countable collection of smooth functions ψ_i such that $\psi_i \geq 0$ and, since the set of supports is locally finite,

$$\psi = \sum \psi_i$$

is well-defined as a smooth function on M. Moreover

$$\operatorname{supp} \psi_i \subset V_\alpha \cap (G_m \backslash \bar{G}_{m-3}) \subset V_\alpha$$

so each support is contained in a V_α. Finally define

$$\varphi_i = \frac{\psi_i}{\psi}$$

then this is the required partition of unity. □

10.5 Sard's theorem (special case)

Theorem 10.9 *Let M and N be differentiable manifolds of the same dimension n and suppose $F : M \to N$ is a smooth map. Then the set of critical values of F has measure zero in N. In particular, every smooth map F has at least one regular value.*

Proof: Since a countable union of null sets (=sets of measure zero) is null, and M and N have a countable basis of open sets, it suffices to consider the local case of $F : U \to \mathbf{R}^n$. Moreover since U is a countable union of compact cubes we need only prove that the image of the set of critical points in the compact cube $K = \{x \in \mathbf{R}^n : |x_i| \leq 1\}$ is of measure zero.

Now suppose $a \in K$ is a critical point, so that the image of DF_a is contained in a proper subspace of $\mathbf{R}^n$, and so is annihilated by a linear form f. Let $H \subset \mathbf{R}^n$ be the hyperplane $f(x - F(a)) = 0$. Then

$$d(F(x), H) \leq \|F(x) - (F(a) + DF_a(x - a))\| \tag{26}$$

On the other hand since F is C^∞, from Taylor's theorem we have a constant C such that

$$\|F(x) - F(y) - DF_y(x - y)\| \leq C\|x - y\|^2$$

for all $x, y \in K$, since K is compact. Substituting in (26) this yields

$$d(F(x), H) \leq C\|x - a\|^2$$

If $\|x - a\| \leq \eta$, then $d(F(x), H) \leq C\eta^2$. Let $M = \sup\{\|DF_x\| : x \in K\}$, then by the mean value theorem

$$\|F(x) - F(a)\| \leq M\|x - a\|$$

for $x, a \in K$ and so $d(F(x), F(a)) \leq M\eta$. Thus $F(x)$ lies in the intersection of a slab of thickness $2C\eta^2$ around H and a ball of radius $M\eta$ centred on $F(a)$. Putting the ball in a cube of side $2M\eta$, the volume of this intersection is less than

$$2C\eta^2(2M\eta)^{n-1} = 2^n CM^{n-1}\eta^{n+1}$$

Now subdivide the cube into N^n cubes of side $1/N$, and repeat the argument for each cube. Since now $\|x - y\| \leq \sqrt{n}/N$, critical points in this cube lie in a volume less than

$$2^n CM^{n-1}\left(\frac{\sqrt{n}}{N}\right)^{n+1}$$

Since there are at most N^n such volumes, the total is less than

$$\left(2^n M^{n-1} C n^{(n+1)/2}\right) N^{-1}$$

which tends to zero as $N \to \infty$.

Thus the set of critical values is of measure zero. □