Contents

Chapter 1: Topological Manifolds

Topological Manifolds

The study of differential geometry begins with tens of pages of definitions.

Definition 1.1: Topological Manifold

Let M be a topological space. M is a topological manifold of dimension m if it is Hausdorff, second-countable, and locally homeomorphic to $\mathbb{R}^n$.

Definition 1.2: Local homeomorphism

M locally homeomorphic to $\mathbb{R}^n$ if every point $x \in M$ an open set U, equipped with a homeomorphism which sends points in U into an open subset of $\mathbb{R}^n$.

$$\phi : U \to \phi(U)$$

The tuple (U, ϕ) is called a coordinate chart.

Definition 1.3: More on coordinate charts

- A coordinate chart (U, ϕ) is centered at $p \in M$ if $p \in U$ and $\phi(p) = 0 \in \mathbb{R}^n$.
- We call U the coordinate domain, and
- we call ϕ the coordinate map.
- If the choice of (U, ϕ) is unambiguous, then the local coordinates of p are simply the coordinates of $\phi(p)$ in $\mathbb{R}^n$, and
- we sometimes also denote $\phi(U)$ by $\hat{U}$ if it is unambiguous to do so.
- If $\hat{U}$ is an open ball/cube, then U is called a coordinate ball/cube.

The central theme of point-set topology (or even metric topology) is that of passing a topological argument to the basis or to a neighbourhood. Manifolds in particular have a nice basis.

Proposition 1.1: Basis of precompact coordinate balls

Every topological manifold has a countable basis of precompact coordinate balls.

Proposition 1.2: Additional facts about topological manifolds

If M is a topological manifold,

- M is locally compact. (Lee, Proposition 1.12)
- M is paracompact, and every open cover has a refinement that is another countably locally finite open cover whose elements are chosen from an arbitrary (but fixed) basis of M. (Lee, Theorem 1.15)
- M is locally-path connected.
- M is connected iff it is path-connected.
- M is metrizable. (Munkres Chapter 6)

Smooth Manifolds

We wish to perform calculus on manifolds.

Definition 2.1: Smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$

Let $F : \mathbb{R}^n \to \mathbb{R}^m$, replacing $\mathbb{R}^n$ and $\mathbb{R}^m$ with open subsets if necessary. F is smooth if its (scalar-valued) component functions has continuous partial derivatives of all orders. The set of smooth functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ is sometimes denoted by $C^\infty(\mathbb{R}^n, \mathbb{R}^m)$. If $m = 1$, we sometimes write $C^\infty(\mathbb{R}^n)$, similar to a test function on the Schwartz Space.

Definition 2.2: Transition map from ϕ to ψ

Let (U, ϕ) and (V, ψ) be coordinate charts on M. The composite function (whenever $U \cap V \neq \varnothing$)

$$\psi \circ \phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$$

is called the transition map. Notice $\psi \circ \phi^{-1}$ is by definition a homeomorphism.

Definition 2.3: Smoothly compatiable

Two coordinate charts on M, (U, ϕ) and (V, ψ) are called smoothly compatible if either their domains are disjoint, or their transition map is a diffeomorphism on $\mathbb{R}^m$.

Definition 2.4: Smooth atlas

An atlas $\mathcal{A}$ of M is a collection of charts $\{(U_\alpha, \phi_\alpha)\}$ whose collection of coordinate domains $\{U_\alpha\}$ for an open cover of M.

It is called a smooth atlas if any two charts in the atlas are pairwise smoothly compatible.

Definition 2.5: Smooth manifold

A smooth atlas $\mathcal{A}$ on M is maximal if it is not contained (properly) in any other smooth atlas as a subset. In other words, if (U', ϕ') is a chart on M that is smoothly compatible with all elements in $\mathcal{A}$, then $(U', \phi') \in \mathcal{A}$ already.

This smooth atlas is often very large, it includes all translations of charts, dilations, and composition with diffeomorphisms in $\mathbb{R}^m$, restrictions onto open subsets, etc. A maximal smooth atlas is sometimes called a complete atlas, or a smooth manifold structure.

A smooth manifold is the tuple $(M, \mathcal{A})$, where $\mathcal{A}$ is some smooth atlas. It can happen if M is originally a topological manifold with a huge number of charts, some of which are smoothly compatible with others, that $\mathcal{A}$ is a strict subset, and both $\mathcal{A}_1$ and $\mathcal{A}_2$ are maximal smooth atlases on M, but $\mathcal{A}_1 \neq \mathcal{A}_2$. We often omit $\mathcal{A}$ and write M if the smooth atlas is understood or not of importance.

Definition 2.6: Smooth coordinate terminologies

Let $(M, \mathcal{A})$ be a smooth manifold.

- Any coordinate chart $(U, \phi) \in \mathcal{A}$ is called a smooth chart, similar to definition 1.3
- We call U the *smooth coordinate domain* or *smooth coordinate neighbourhood* of any $p \in U$, and
- we call ϕ the *smooth coordinate map*.
- The terms *smooth coordinate ball* and *smooth coordinate cube* are used similarly.
- A set $B \subseteq M$ is a *regular coordinate ball* if its image is a smooth coordinate ball centered at the origin; and the closure of this ball in $\mathbb{R}^m$ is a subset of the image of another smooth coordinate ball, centered at the origin.

Definition 2.7: Standard smooth structure on $\mathbb{R}^n$

The maximal smooth atlas containing $(\mathbb{R}^n, \mathrm{id}_{\mathbb{R}^n})$ is called the *standard smooth structure on* $\mathbb{R}^n$.

Manifolds with boundary are not as important as regular manifolds for now, but they are worth mentioning.

Definition 2.8: Closed n-dimensional upper half-plane $\mathbb{H}^n \subseteq \mathbb{R}^n$

We define the following symbols for the upper half plane.

- $\mathbb{H}^n = \left\{x \in \mathbb{R}^n,\ x^n \geq 0\right\}$,
- $\operatorname{Int} \mathbb{H}^n = \left\{x \in \mathbb{R}^n,\ x^n > 0\right\}$,
- $\partial\mathbb{H}^n = \left\{x \in \mathbb{R}^n,\ x^n = 0\right\}$

Definition 2.9: Manifolds with boundary

A topological space M is called a manifold with boundary if it is Hausdorff, second-countable, and locally homeomorphic to an open subset of $\mathbb{H}^n$ (endowed with the subspace topology from $\mathbb{R}^n$).

A chart (U, ϕ) is an *interior chart* if its coordinate image is disjoint from the 'boundary' of the upper-half plane. This means $\phi(U) \cap \partial\mathbb{H}^n = \varnothing$. Similarly, (V, ψ) is a *boundary chart* if its range contains a point in $\partial\mathbb{H}^n$; so $\psi(V) \cap \partial\mathbb{H}^n \neq \varnothing$.

Similar to definition 1.3 and definition 2.6, we use the terms *coordinate half-ball*, *coordinate half-cube*, *regular coordinate half-ball*.

Let $p \in M$, it is called an *interior point of* M (not to be confused with the topological interior) if it is in the domain of some interior chart, and p is called a *boundary point of* M if there exists a boundary chart that sends p into $\partial\mathbb{H}^n$. The set of interior points and boundary points of M will be denoted by $\operatorname{Int} M$ and ∂M.

Example 2.1: Sphere as a topological manifold

The n-sphere as a topological manifold. Define

$$S^n = \left\{x \in \mathbb{R}^{n+1},\ |x| = 1\right\}$$

We claim that $\{U_i^{\pm}\}_{i=1}^{n+1}$ form an open cover, where

$$U_i^+ = \left\{x \in S^n, x^i > 0\right\} \quad U_i^- = \left\{x \in S^n, x^i < 0\right\}$$

Each $U_i^{\pm}$ is the inverse image of $\pi_i^{-1}((0, +\infty)) \cap S^n$ or $\pi_i^{-1}((0, -\infty)) \cap S^n$, hence open. For every $x \in S^n$, there exists at least some $1 \leq j \leq n+1$ that makes the j-th coordinate of x, $x^j \neq 0$. So

$$S^n = \bigcup_i U_i^{\pm}$$

Denote the unit ball $\left\{x \in \mathbb{R}^n,\ |x| < 1\right\}$ in $\mathbb{R}^n$ by $\mathbb{B}^n$.

Chapter 2: Smooth Maps

Smooth Maps

Definition 1.1: Smooth functions $C^\infty(M, \mathbb{R}^k)$

Let $F : M \to \mathbb{R}^k$ be a vector-valued function on a smooth manifold M. We say F is a smooth function if for every $p \in M$, there exists a smooth chart $p \in (U, \phi)$ such that the *coordinate representation of F at p, with respect to* (U, ϕ) is a smooth function from $\mathbb{R}^m$ to $\mathbb{R}^k$, denoted by $\hat{F}$ (in the sense of Definition 2.1).

$$\hat{F} = F \circ \phi^{-1} : \phi(U) \to \mathbb{R}^k \in C^\infty(\phi(U), \mathbb{R}^k)$$

if $k = 1$, then we denote the space of *test functions* on M by $C^\infty(M) = C^\infty(M, \mathbb{R})$

Definition 1.2: Smooth maps between manifolds $C^\infty(N, M)$

Let $F : N \to M$ be a map between smooth manifolds N and M (note we switched the order). F is a smooth map if at every $p \in M$, there exists

- a chart in the smooth atlas of N (the domain), $p \in (U, \phi)$,
- another chart in the smooth atlas of M (the range), $F(U) \subseteq (V, \psi)$,
- such that, the *coordinate representation of F at p with respect to* (U, ϕ)*, and* (V, ψ) is a smooth function from $\mathbb{R}^n$ to $\mathbb{R}^m$, also denoted by $\hat{F}$.

$$\hat{F} = \psi \circ F \circ \phi^{-1} : \phi(U) \to \psi(V) \in C^\infty(\mathbb{R}^n, \mathbb{R}^m) \tag{1}$$

The following propositions summarizes common operations on smooth maps, a few sources of them.

Proposition 1.1: Smooth maps are continuous

If $F : N \to M$ is a smooth map, then F is continuous with respect to the topologies on N and M.

Proof. Let $p \in N$ be fixed, because F is smooth this induces two smooth charts, one in the domain and another in the range; as in definition 1.2. $F(p)$ is a point in M. From eq. (1), $\hat{F}|\phi(U)$ is a smooth hence continuous function. Since $\phi : U \mapsto \phi(U)$ and $\psi : V \mapsto \psi(V)$ are homeomorphisms,

$$F|U = \underbrace{\psi^{-1}}_{\text{continuous}} \circ \underbrace{\hat{F}|\phi(U)}_{\text{smooth}} \circ \underbrace{\phi}_{\text{continuous}} \quad \text{is continuous on} \quad U$$

Let the point p range through all the points in N, so F is continuous at every p, hence on N. ■

Proposition 1.2: Characterizations of Smooth Maps in $C^\infty(N, M)$

Let N and M be smooth manifolds, and $F : N \to M$. F is a smooth map iff

- For every $p \in N$, there exists smooth charts $p \in (U, \phi)$ and $F(p) \in (V, \psi)$ such that $U \cap F^{-1}(V)$ is an open set in N, and the composite map (the coordinate representation)

$$\psi \circ F \circ \phi^{-1}|(U \cap F^{-1}(V)) : \phi(U \cap F^{-1}(V)) \to \psi(V) \quad \text{is smooth}$$

- F is continuous and there exist smooth atlases $\{(U_\alpha, \phi_\alpha)\} \subseteq \mathcal{A}_N$, and $\{(V_\beta, \psi_\beta)\} \subseteq \mathcal{A}_M$ such that the coordinate representation

$$\psi_\beta \circ F \circ \phi_\alpha^{-1} : \phi(U_\alpha \cap F^{-1}(V_\beta)) \to \psi(V_\beta) \quad \text{is smooth}$$

  whenever it makes sense.

- the restriction of F onto any arbitrary open set U, $F|U : U \mapsto M$ is smooth (in the sense of open submanifold).

Proof. By Proposition 1.1, and the fact that complete atlases are closed under restrictions onto open sets, it is clear that the original definition implies the two. The first definition also clearly implies the original definition, as we can restrict

$$(U, \phi) \mapsto \Big(U \cap F^{-1}(V), \phi|(U \cap F^{-1}(V))\Big)$$

since $U \cap F^{-1}(V)$ is open in the domain manifold.

The second definition implies the original one as well, since the smooth atlases are taken from the maximal atlas, we can pass the argument to any smoothly-compatible chart. Atlases must cover both the domain and the range, and coordinate transitions between smoothly compatible charts are diffeomorphisms. If F is smooth on a subcollection of those charts, meaning

$$\psi_\beta \circ F \circ \phi_\alpha^{-1} \in C^\infty(\phi_\alpha(U_\alpha) \cap F^{-1}(V_\beta), \psi_\beta(V_\beta))$$

it is smooth with respect to every pair of (smooth) charts in the two atlases $\mathcal{A}_N$, $\mathcal{A}_M$, as a composition of smooth maps:

$$\psi \circ \underbrace{\psi^{-1} \circ \psi_\beta}_{\text{smooth}} \circ F \circ \underbrace{\phi_\alpha^{-1} \circ \phi}_{\text{smooth}} \circ \phi^{-1}$$

where we can restrict $\phi_\alpha \mapsto \phi_\alpha|(U_\alpha \cap F^{-1}(V_\beta))$ by continuity of F.
I will prove the third and last equivalence later. ■

Proposition 1.3: Sources of smooth maps

Let N, M, P be smooth manifolds, then

- Every constant map is smooth,
- The identity map $\mathrm{id}_M : M \to M$ is smooth,
- The inclusion map $\iota : W \to M$ is smooth, where W is an open submanifold of M.
- The composition of smooth maps is again a smooth map: if $F \in C^\infty(N, M)$ and $G \in C^\infty(M, P)$, then $(G \circ F) \in C^\infty(N, P)$

Diffeomorphisms

Definition 2.1: Diffeomorphism between Manifolds $\mathcal{D}(N, M)$

Let N and M be smooth manifolds, $F : N \to M$ is a diffeomorphism if it is a smooth bijective map with a smooth inverse. We denote the space of diffeomorphisms from N to M by $\mathcal{D}(N, M)$.

Proposition 2.1: Properties of Manifold Diffeomorphisms

Let N, M and P be smooth manifolds, then

- The composition of diffeomorphisms is again a diffeomorphism, that is, if $F \in \mathcal{D}(N, M)$ and $G \in \mathcal{D}(M, P)$, then $(G \circ F) \in \mathcal{D}(N, P)$.
- The open-manifold restriction of a diffeomorphism onto its image is again a diffeomorphism,
- Every diffeomorphism is a homeomorphism and an open map.

Proof. Trivial. ■

Partitions of Unity

See Folland Chapters 4 and 8. Including Urysohn's Lemma, Tietze's Extension Theorem, the usual construction of C_c^∞ bump functions.

Chapter 3: Tangent Spaces

Algebra of Germs on $C^\infty(N)$

The tangent space is a powerful concept that acts almost like the dual in distribution theory.

Definition 1.1: Algebra of Germs at p: $C_p^\infty(N)$

Let N be a smooth manifold and $p \in N$. We define an equivalence relation on the space of test functions on N, $C^\infty(N)$. If $f, g \in C^\infty(N)$, we write $f \sim g$ if $f = g$ for some open neighbourhood about p. We denote this equivalence class by $C_p^\infty(N)$, and it is clear $C^\infty(N)$ is closed under pointwise multiplication by the product rule, and form an algebra; so $C_p^\infty(N)$ is an algebra too.

Tangent spaces of manifolds

Definition 2.1: Vector space of derivations at p: T_pN

Let $\nu : C_p^\infty(N) \to \mathbb{R}$ be a linear functional on the vector space of germs at p. It is called a derivation at p if it satisfies the product rule, if $f, g \in C_p^\infty(N)$,

$$\nu(fg) = g(p)\nu(f) + f(p)\nu(g)$$

then we say

- ν is a tangent vector at p,
- $\nu \in T_pN$,
- ν is an element of the *tangent space at* p.
- ν is a derivation on N at p.

Proposition 2.1: Properties of derivations at p

Let N be a smooth manifold and $p \in N$.

- If $f \in C_p^\infty$ is constant in some neighbourhood of p, then $\nu(f) = 0$ for every $\nu \in T_pN$,
- If $f(p) = g(p) = 0$, then $\nu(fg) = 0$ for tangent vector ν at p.

Tangent spaces of $\mathbb{R}^n$

Proposition 3.1: Basis of $T_p\mathbb{R}^n$

Let $\mathbb{R}^n$ be equipped with the standard smooth structure as in Definition 2.7. The vector space of derivations at $p \in \mathbb{R}^n$ are spanned by the n partial derivatives at p

$$\left.\frac{\partial}{\partial x^j}\right|_p : f \mapsto \left.\frac{\partial}{\partial x^j} f(x)\right|_p, \quad 1 \leq j \leq n,\ f \in C^\infty(\mathbb{R}^n)$$

Moreover, the n vectors form a basis, and $\dim T_p\mathbb{R}^n = n$.

Definition 3.1: Standard Basis of $T_p\mathbb{R}^n$

The standard basis for the tangent space at $p \in \mathbb{R}^n$ is the n partial derivatives at p.

$$T_p\mathbb{R}^n = \left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right\} \tag{2}$$

Definition 3.2: Coordinate functions x^j on $\mathbb{R}^n$

Let $x^1, \ldots x^n$ be the standard coordinates on $\mathbb{R}^n$. If $1 \leq j \leq n$, x^j is a function (also represented by the same symbol as the standard coordinate) from $\mathbb{R}^n$ to $\mathbb{R}$, which maps each $p = (p^1, \ldots, p^j, \ldots, p^n)$ to its j-th coordinate $p \mapsto p^j$.

This map is linear, hence smooth, has the matrix representation of $\mathcal{M}\{x^j\} = (\delta_{jk})_{1,k} \in \mathbb{R}^{1\times n}$ where δ_{jk} denotes the discrete mass at j.

Furthermore, the coordinate functions behave like the dual basis for the derivations $\partial/\partial x^j|_p$ at p

$$\left.\frac{\partial}{\partial x^j}\right|_p (x^k) = \delta_{jk}$$

If $\nu \in T_p\mathbb{R}^n$, and has the basis representation

$$\nu = \sum_j \nu^j \left.\frac{\partial}{\partial x^j}\right|_p = \nu^j \left.\frac{\partial}{\partial x^j}\right|_p$$

then

$$\nu = \sum_j \nu(x^j) \left.\frac{\partial}{\partial x^j}\right|_p = \nu(x^j) \left.\frac{\partial}{\partial x^j}\right|_p \tag{3}$$

Differential of a smooth map $F \in C^\infty(N, M)$

Definition 4.1: Differential of a smooth map $dF_p : T_pN \to T_{F(p)}M$

Let $F \in C^\infty(N, M)$, $p \in N$ and $\nu \in T_pN$ be a tangent vector at p. The differential of a smooth map is a linear map that sends tangent vectors in T_pN to tangent vectors in $T_{F(p)}M$. If $f \in C^\infty(M)$ is a test function on M, then $f \circ F$ is a test function on N, and

$$dF_p(\nu)(f) = \nu\Big(\underbrace{f \circ F}_{C^\infty(N)}\Big)$$

We state the following without proof, as the proof is tedious. It simply involves unboxing the definition of the differential $dF_p : T_pN \to T_pM$, and projecting onto the coordinates.

Proposition 4.1: Properties of the differential

Let N, M and P be smooth manifolds, if $F \in C^\infty(N, M)$, $G \in C^\infty(M, P)$ then

- dF_p is a linear map between T_p and $T_{F(p)}N$,
- $d(G \circ F)_p = dG_{F(p)} \circ dF_p$
- $d(\mathrm{id}_N)_p = \mathrm{id}_{T_pN}$,
- if $F \in \mathcal{D}(N, M)$, then dF_p is a linear isomorphism between T_pN and $T_{F(p)}M$, and
$$(dF_p)^{-1} = d(F^{-1})_{F(p)}$$

Proposition 4.2: Matrix representation of the differential of $F : N \to M$

Let $F \in C^\infty(N, M)$, and $p \in N$ induces two charts $p \in (U, \phi)$ and $F(p) \in (V\,\psi)$. The matrix representation of the differential at p, $dF_p : T_pN \to T_{F(p)}M$ is nothing but the Jacobian matrix of size $m \times n$ of the coordinate representation at p.

$$\mathcal{M}\{dF_p\} = \begin{bmatrix} \frac{\partial \hat{F}^1}{\partial x^1}\Big|_{\phi(p)} & \frac{\partial \hat{F}^1}{\partial x^2}\Big|_{\phi(p)} & \cdots & \cdots & \frac{\partial \hat{F}^1}{\partial x^n}\Big|_{\phi(p)} \\ \frac{\partial \hat{F}^2}{\partial x^1}\Big|_{\phi(p)} & \frac{\partial \hat{F}^2}{\partial x^2}\Big|_{\phi(p)} & \cdots & \cdots & \frac{\partial \hat{F}^2}{\partial x^n}\Big|_{\phi(p)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \hat{F}^m}{\partial x^1}\Big|_{\phi(p)} & \frac{\partial \hat{F}^m}{\partial x^2}\Big|_{\phi(p)} & \cdots & \cdots & \frac{\partial \hat{F}^m}{\partial x^n}\Big|_{\phi(p)} \end{bmatrix} \tag{4}$$

Alternately, if we write $\hat{p} = \phi(p)$ as the $\mathbb{R}^m$ coordinates at p, then

$$\mathcal{M}\{dF_p\} = \begin{bmatrix} \frac{\partial \hat{F}^1}{\partial x^1}\Big|_{\hat{p}} & \frac{\partial \hat{F}^1}{\partial x^2}\Big|_{\hat{p}} & \cdots & \cdots & \frac{\partial \hat{F}^1}{\partial x^n}\Big|_{\hat{p}} \\ \frac{\partial \hat{F}^2}{\partial x^1}\Big|_{\hat{p}} & \frac{\partial \hat{F}^2}{\partial x^2}\Big|_{\hat{p}} & \cdots & \cdots & \frac{\partial \hat{F}^2}{\partial x^n}\Big|_{\hat{p}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \hat{F}^m}{\partial x^1}\Big|_{\hat{p}} & \frac{\partial \hat{F}^m}{\partial x^2}\Big|_{\hat{p}} & \cdots & \cdots & \frac{\partial \hat{F}^m}{\partial x^n}\Big|_{\hat{p}} \end{bmatrix} \tag{5}$$

Differential of a smooth map $F \in C^\infty(\mathbb{R}^n, \mathbb{R}^m)$

An important application of this is the following. We begin with the $\mathbb{R}^m \to \mathbb{R}^n$ case. We will see that if p and $F(p)$ are represented by another pair of coordinate charts (smoothly compatible with the previous pair), then the rank of dF_p does not change. So the rank of the differential is an invariant of the choice of coordinate chart.

Definition 5.1: Matrix representation of the differential of $F : \mathbb{R}^m \to \mathbb{R}^n$

Let $F \in C^\infty(\mathbb{R}^m, \mathbb{R}^n)$, and $p \in \mathbb{R}^m$ induces two charts $p \in (U, \mathrm{id}_{\mathbb{R}^m})$ and $F(p) \in (V\ \mathrm{id}_{\mathbb{R}^n})$, where $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$. The matrix representation of the differential at p, $dF_p : T_p\mathbb{R}^m \to T_{F(p)}\mathbb{R}^n$ is nothing but the Jacobian matrix of F at p.

$$\mathcal{M}\{dF_p\} = DF(p) = \begin{bmatrix} \frac{\partial F^1}{\partial x^1}\Big|_{p} & \frac{\partial F^1}{\partial x^2}\Big|_{p} & \cdots & \cdots & \frac{\partial F^1}{\partial x^m}\Big|_{p} \\ \frac{\partial F^2}{\partial x^1}\Big|_{p} & \frac{\partial F^2}{\partial x^2}\Big|_{p} & \cdots & \cdots & \frac{\partial F^2}{\partial x^m}\Big|_{p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial F^n}{\partial x^1}\Big|_{p} & \frac{\partial F^n}{\partial x^2}\Big|_{p} & \cdots & \cdots & \frac{\partial F^n}{\partial x^m}\Big|_{p} \end{bmatrix} \tag{6}$$

Change of Coordinates Matrix

We will go through the section on the Change of Coordinates, and how different coordinate charts change the representation of a derivation at $p \in M$, where M is some smooth manifold.

Definition 6.1: Standard basis of T_pN

From proposition 2.1, since ϕ^{-1} is a diffeomorphism, $d\Big(\phi^{-1}|_{\phi(p)}\Big) : T_{\phi(p)}\mathbb{R}^n \to T_pN$ is a linear isomorphism. Hence T_pN is a n-dimensional vector space, and the standard basis vectors of $T_p\mathbb{R}^n$ are denoted by

$$\left\{ \frac{\partial}{\partial x^1}\bigg|_p, \dots, \frac{\partial}{\partial x^n}\bigg|_p \right\} \tag{7}$$

where each basis vector $\frac{\partial}{\partial x^1}\Big|_p \triangleq d\Big(\phi^{-1}|_{\phi(p)}\Big)$ is the push-forward derivation (through ϕ^{-1}) of the j-th standard basis vector in $T_{\phi(p)}N$.

Proposition 6.1: Differential of $\psi \circ \phi^{-1} : M \to\ M$

Let M be a smooth manifold, and fix $p \in M$. If $\nu \in T_pM$ is given with respect to the bases

$$\left\{ \frac{\partial}{\partial x^1}\bigg|_p, \dots, \frac{\partial}{\partial x^m}\bigg|_p \right\} \quad \text{and} \quad \left\{ \frac{\partial}{\partial y^1}\bigg|_p, \dots, \frac{\partial}{\partial y^m}\bigg|_p \right\}$$

Defined by

$$\frac{\partial}{\partial x^j}\bigg|_p \triangleq d\left(\phi^{-1}\Big|_{\phi(p)}\right)\left(\frac{\partial}{\partial x^j}\bigg|_{\phi(p)}\right) \quad \text{and} \quad \frac{\partial}{\partial y^j}\bigg|_p \triangleq d\left(\psi^{-1}\Big|_{\psi(p)}\right)\left(\frac{\partial}{\partial y^j}\bigg|_{\psi(p)}\right)$$

and we write ν in terms of the first basis

$$\nu = \nu^j\ \frac{\partial}{\partial x^j}\bigg|_p = \sum_{j=1}^m \nu^j\ \frac{\partial}{\partial x^j}\bigg|_p$$

and the second basis

$$\nu = \nu^j\ \frac{\partial y^k}{\partial x^j}\bigg|_{\phi(p)}\ \frac{\partial}{\partial y^k}\bigg|_p = \sum_{k=1}^m \sum_{j=1}^m \nu^j\ \frac{\partial y^k}{\partial x^j}\bigg|_{\phi(p)}\ \frac{\partial}{\partial y^k}\bigg|_p$$

If $f \in C^\infty(M)$, then

$$\nu(f) = \nu^j\ \frac{\partial}{\partial x^j}\bigg|_p f = \nu^j\ \frac{\partial y^k}{\partial x^j}\bigg|_{\phi(p)}\ \frac{\partial}{\partial y^k}\bigg|_p f$$

Proof. Recall $\frac{\partial}{\partial x^j}\Big|_p f \triangleq \frac{\partial}{\partial x^j}\Big|_{\phi(p)} f \circ \phi^{-1}$, similarly for $\frac{\partial}{\partial y^j}\Big|_p f$. Deriving f and p and by vector space operations on T_pM, the first basis expansion gives

$$\nu^j\ \frac{\partial}{\partial x^j}\bigg|_p f = \nu^j\ \frac{\partial}{\partial x^j}\bigg|_{\phi(p)} f \circ \phi^{-1} \tag{8}$$

and the second expression reads

$$\nu^j \left.\frac{\partial y^k}{\partial x^j}\right|_{\phi(p)} \left.\frac{\partial}{\partial y^k}\right|_p f = \nu^j \left.\frac{\partial y^k}{\partial x^j}\right|_{\phi(p)} \left.\frac{\partial}{\partial y^k}\right|_{\psi(p)} f \circ \psi^{-1} \tag{9}$$

Since $f \circ \phi^{-1} \in C^\infty(\mathbb{R}^m, \mathbb{R})$, we see the expressions are indeed equal. By the chain rule, if

$$\psi \circ \phi^{-1}(x^1, \dots x^m) = (y^1, \dots y^m)$$

then

$$D(\psi \circ \phi^{-1})(\phi(p)) = \begin{bmatrix} \left.\frac{\partial y^1}{\partial x^1}\right|_{\phi(p)} & \left.\frac{\partial y^1}{\partial x^2}\right|_{\phi(p)} & \cdots & \cdots & \left.\frac{\partial y^1}{\partial x^m}\right|_{\phi(p)} \\ \left.\frac{\partial y^2}{\partial x^1}\right|_{\phi(p)} & \left.\frac{\partial y^2}{\partial x^2}\right|_{\phi(p)} & \cdots & \cdots & \left.\frac{\partial y^2}{\partial x^m}\right|_{\phi(p)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \left.\frac{\partial y^m}{\partial x^1}\right|_{\phi(p)} & \left.\frac{\partial y^m}{\partial x^2}\right|_{\phi(p)} & \cdots & \cdots & \left.\frac{\partial y^m}{\partial x^m}\right|_{\phi(p)} \end{bmatrix}$$

It follows from Proposition 3.6d) that the matrix $D(\psi \circ \phi^{-1})|_{\phi(p)}$ is invertible, as $\psi \circ \phi^{-1}$ is a diffeomorphism. ■

Proposition 6.2: Rank of a dF_p is invariant under coordinate change

Let F be a smooth map between M and N, at every $p \in M$, rank dF_p is an invariant over (smoothly compatible) pairs of charts in M and N.

Proof. Let $p \in (U_1, \phi_1) \cap (U_2, \phi_2)$, and $F(p) \in (V_1, \psi_1) \cap (V_2, \psi_2)$. Where all charts are smoothly compatible if it makes sense to talk about it. Both $\phi_2 \circ \phi_1^{-1}$ and $\psi_2 \circ \psi_1^{-1}$ are diffeomorphisms, and the change of basis matrices $D(\phi_2 \circ \phi_1^{-1})\big|_{\phi_1(p)}$ and $D(\psi_2 \circ \psi_1^{-1})\big|_{\psi_1(F(p))}$ are invertible by Proposition 3.6d) again, so the ranks dF_p with respect to any of the two charts are equal.

$$\underbrace{D(\psi_2 \circ \psi_1^{-1})\Big|_{\psi_1(F(p))}}_{\text{invertible}} \Big(\mathcal{M}\{dF_p\}\Big) \underbrace{D(\phi_2 \circ \phi_1^{-1})\Big|_{\phi_1(p)}}_{\text{invertible}}$$

■

Chapter 4: Submersions, Immersions and Embeddings

Matrices

The following is of utmost importance. It states that that rank of a matrix, square or otherwise, is an 'open condition'.

Example 1.1: Lee Example 1.28 (Matrices of Full Rank)

Let $A \in \mathcal{M}(m \times n, \mathbb{R})$ be the set of $m \times n$ matrices with real entries. A has rank m iff there exists some $m \times m$ sub-matrix of A, denoted by S st S is invertible. We wish to show the set of rank-m matrices is invertible. Indeed, let

$$F : \mathcal{M}(m \times n, \mathbb{R}) \to \mathbb{R},\ \Delta_{m\times m}(A) = \sum_{\substack{S \text{ is a } m\times m \\ \text{sub-matrix of } A}} |\det\{S\}|$$

Since $S \mapsto \det\{S\}$ is continuous in the entries of S, hence continuous in the entries of A, $\Delta_{m\times m}$ is continuous.
So the set $\left\{A \in \mathcal{M}(m \times n,\ \mathbb{R}),\ \operatorname{rank} A = m\right\} = F^{-1}(\mathbb{R} \setminus \{0\})$ is open.

Estimates in vector calculus

Before proving the inverse function theorem, we will need several Lemmas

Proposition 2.1: Rudin Theorem 9.7

If A and B are in $L(\mathbf{X}, \mathbf{Y})$, then

$$\|BA\| \leq \|B\|\|A\|$$

Proof. Let $\|x\| = 1$, and

$$\|B(Ax)\| \leq \|B\|\|Ax\| \leq \|B\|\|A\|\|x\|$$

this holds for every $\|x\| = 1$, hence

$$\|BA\| \leq \|B\|\|A\|$$

■

Proposition 2.2: Rudin Theorem 9.19

Let f map a convex open set $U \subseteq \mathbb{R}^n$ into $\mathbb{R}^m$, if f is differentiable (pointwise) in U, and there exists some M st its derivative its bounded (in the operator norm)

$$\|Df(x)\| \leq M \quad x \in U$$

then, for every pair of elements x_1, x_2 in U,

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

Proof. This proof 'passes the argument' to the scalar-valued version, in short: if x_1 and x_2 are in U. Define

$$c(t) = (1-t)x_1 + tx_2$$

as the convex combination of x_1 and x_2. The takeaway intuition here is that it suffices to check on the line joining the two points', to obtain an estimate for $\|f(x_1) - f(x_2)\|$. Indeed, define

$$g(t) = f(c(t)) \text{ is a curve } g : \mathbb{R} \to \mathbb{R}^m$$

Using Theorem 5.19, of which we will state below

Proposition 2.3: Rudin Theorem 5.19

Let $g : [0,1] \to \mathbb{R}^m$, and g be differentiable on $(0,1)$, then there exists some $x \in (0,1)$ with

$$|f(b) - f(a)| \leq (b-a)|f'(x)|$$

Proof. Read from Rudin Theorem 5.19. ■

Since $Dg(t) = Df(c(t)) \circ Dc(t)$ by the Chain Rule, and $Dc(t) = b - a$ by inspection,

$$\|Dg(t)\| = \|Df(c(t)) \circ Dc(t)\| \leq \|Df\|\|Dc\| = \|Df\|(b-a)$$

This holds for every $t \in [0,1]$. Applying Theorem 5.19 gives

$$\underbrace{\|g(1) - g(0)\|}_{\text{curve endpoints}} \leq M\|b - a\|$$

Replacing $\|g(1) - g(0)\| = \|f(x_1) - f(x_2)\|$ and $\|Df\| \leq M$ we get

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

■

Inverse Function Theorem (Rudin)

Proposition 3.1: Rudin Theorem 9.24

Suppose $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, and $Df(a)$ is invertible for some $a \in \mathbb{R}^n$, and define $b = f(a)$. Then,

(a) there exist open sets U and V in $\mathbb{R}^n$ such that $a \in U$, $b \in V$, and f is one-to-one on U, and $f(U) = V$.

(b) if g is the inverse of f (which exists, by Part a), defined in V by $g(f(x)) = x$ for every $x \in U$ then $g \in C^1(\mathbb{R}^n, \mathbb{R}^n)$

Proof of Part A. We define $Df(a) = A \in \mathbb{R}^{n \times n}$, so A is invertible, and $\|A^{-1}\| \neq 0$, where $\|\cdot\|$ denotes the operator norm. Recall all norms on finite-dimensional vector spaces are equivalent, this will be useful later.
Choose $\lambda > 0$ st

$$\lambda = \left\|A^{-1}\right\|^{-1} 2^{-1} \tag{10}$$

By continuity of $Df(x)$ at the point a, let $\lambda > 0$, this induces a $B(\delta, a)$ with $x \in B(\delta, a)$ means

$$\underbrace{\|Df(x) - Df(a)\|}_{\text{operator norm}} < \lambda \tag{11}$$

as $Df : \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^n)$ takes a point in $\mathbb{R}^n$ and returns a linear map., with $L(\mathbb{R}^n, \mathbb{R}^n)$ endowed with the usual vector space structure. Fix $y \in \mathbb{R}^n$, and define

$$\phi(x) = x + \underbrace{A^{-1}(y - f(x))}_{\text{offset}}$$

this is now a function solely in x, and $\phi(x) = x \iff f(x) = y$ is clear, but such a fixed point is not necessarily unique. We claim that it is unique in $B(\delta, a)$. We will use the contractive mapping principle.

Differentiating $\phi(x)$ reads

$$D\phi(x) = \underbrace{I}_{I = A^{-1}A} - A^{-1}Df(x) = A^{-1}(A - Df(x))$$

Proposition 2.1 tells us the norm of a product is bounded above by the product of the norms. Using eqs. (10) and (11), if $x \in U$ we have

$$\|D\phi(x)\| = \left\|A^{-1}(A - Df(x))\right\| \leq \left\|A^{-1}\right\| \|A - Df(x)\| \leq 2^{-1}$$

The total derivative of ϕ is uniformly bounded in U, applying Proposition 2.2 tells us that ϕ is a contractive mapping

$$\|D\phi(x)\| \leq 2^{-1} \implies \|\phi(x_1) - \phi(x_2)\| \leq 2^{-1}\|x_1 - x_2\|$$

for x_1, x_2 in U.

To show $f|U$ is indeed a bijection, fix $y \in f(U)$ so $y = f(x)$ for some $x \in U$, and there can only be one fixed point stemming from $\phi|U$, with $\phi(z) = z + A^{-1}(y - f(z))$ being the 'fixed point detector'. Write $(f|U)^{-1}(y) = \lim\{(\phi|U)(x_n)\}_n$ and every point in $f(U)$ has a unique inverse.

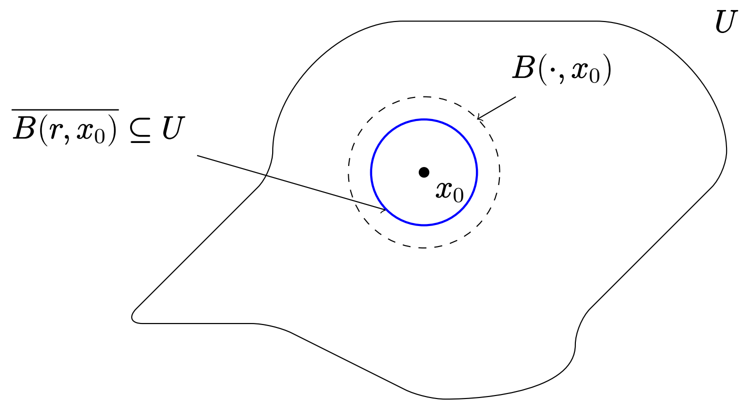

Figure 1: Every point x_0 in an open set U admits an open ball that hides in U

For the last part of the proof, we wish to show $V = f(U)$ is open. Let $y_0 \in V$ and we can 'hone into' the inverse of y_0 using the same construction as earlier. So $f(x_0) = y_0$ for some unique $x_0 \in U$.

If x_0 is in U, it induces an open ball (see fig. 1) st

$$x_0 \in B(r, x_0) \subseteq \overline{B(r, x_0)} \subseteq U, \quad r > 0$$

We claim the open ball $B(\lambda r, y_0) \subseteq V$. Indeed, suppose $y \in \mathbb{R}^n$ with

$$d(y, y_0) < \lambda r$$

If ϕ is the 'fixed-point detector' with respect to y (the point we are trying to prove that is in $f(U)$), in fact: we will prove $y \in f(\overline{B(r, x_0)}) \subseteq f(U)$.

$$\underbrace{\phi(x_0) - x_0}_{\text{removing the offset from } \phi(x_0)} = A^{-1}(y - f(x_0)) = A^{-1}(y - y_0)$$

using the operator norm on $A^{-1}(y - y_0)$ reads

$$\|\phi(x_0) - x_0\| = \left\|A^{-1}(y - y_0)\right\| \leq \left\|A^{-1}\right\| \|y - y_0\| \leq \left\|A^{-1}\right\| \lambda r = r2^{-1}$$

We will drag y into the image of the closed ball as follows: suppose x is another point that lies in the closed ball, ϕ is contractive on $\overline{B} \subseteq U$ regardless of the point y that induces ϕ. But $\overline{B}$ is closed, hence it is complete. So the Cauchy sequence (from the contractive mapping theorem) produces exactly one point in $\overline{B}$. It remains to show that if we start our sequence at some point $x \in \overline{B}$, then $\phi(x) \in \overline{B}$ as well, and a simple induction will produce our contractive sequence.

To this, fix $x \in \overline{B}$, and

$$\begin{aligned}|\phi(x) - x_0| &\leq |\phi(x) - \phi(x_0)| + |\phi(x_0) - x_0| \\ &\leq \overbrace{2^{-1}|x - x_0|}^{\text{contraction on } \overline{B} \subseteq U} + \overbrace{r2^{-1}}^{\text{earlier}} \\ &= r\end{aligned}$$

therefore ϕ contracts to a fixed point $x^* \in \overline{B}$, and $f(x^*) = y$. So $y \in f(\overline{B}) \subseteq f(U)$ as desired. ■

Proof of Part B. The proof is quite long, and we will only focus on the important bits. Rudin uses the technique of approximating smooth functions using first-order terms. He writes

$$\begin{cases} f(x) & = y \\ f(x+h) & = y + k \end{cases} \implies k = f(x+h) - f(x)$$

Furthermore, if $x \in U$, then the derivative $Df(x)$ is invertible, this is from Theorem 9.8, obtains an estimate on the open ball in $GL(n, \mathbb{R})$. Roughly speaking, this open ball 'drags' other matrices into $GL(n, \mathbb{R})$. If A is invertible, and B is a conformable matrix with A, then

$$\underbrace{\|B - A\|}_{\substack{\text{distance} \\ \text{between} \\ A,B}} \left\|A^{-1}\right\| < 1 \implies B \in GL(n, \mathbb{R})$$

If $x \in B(\delta, a)$, then Equation (11) reads

$$\|Df(x) - A\| < \lambda \implies \|Df(x) - A\| \left\|A^{-1}\right\| < 2^{-1} < 1$$

so $Df(x)$ is invertible with inverse T.

And we estimate the deviation $|k|^{-1} \leq \lambda |h|^{-1}$ by using the contraction inequality with y as the basepoint for ϕ. Skipping a few lines ahead (to the confusing part), we see that

$$|h| \leq |h - A^{-1}k| + |A^{-1}k| \leq 2^{-1}|h| + |A^{-1}k|$$

subtracting over, and multiplying across gives a upper bound on $|k|^{-1}$

$$2^{-1}|h| \leq |A^{-1}k| \implies 2^{-1}|h| \leq \left\|A^{-1}\right\| |k| \implies |k|^{-1} \leq \underbrace{\frac{2}{\|A^{-1}\|}}_{\lambda} |h|^{-1}$$

Notice $2\lambda \|A^{-1}\| = 1$, so $2/\|A^{-1}\| = \lambda$. Finally, we 'factor out' $-T$ on the line just before the difference quotient.

$$\begin{aligned}
\overbrace{g(y+k)-g(y)-Tk}^{\text{numerator in difference quotient}} &= h - Tk \\
&= -T\Big(\underbrace{f(x+h)-f(x)}_{=k} - \underbrace{Df(x)h}_{=T^{-1}h}\Big)
\end{aligned}$$

We see that $T = Dg(y)$, indeed:

$$\begin{aligned}
\frac{|g(y+k)-g(y)-Tk|}{|k|} &\leq \frac{\|T\|}{\lambda}\frac{|f(x+h)-f(x)-Df(x)h|}{|h|} \\
&\lesssim \frac{|f(x+h)-f(x)-Df(x)h|}{|h|}
\end{aligned}$$

$$= \underbrace{o(h) = o(k)}_{|h|\lesssim|k|} \to 0$$

Finally, $Df|U : U \to GL(n,\mathbb{R})$ is a continuous mapping. By Theorem 9.8, $(Df|U)^{-1} : U \to GL(n,\mathbb{R})$ is continuous as well. Therefore $g \in C^1(U,U)$, and $f|U$ is a C^1-diffeomorphism. ■

Remark 3.1

The inverse function theorem is extremely powerful. If a f is a C^1 map from and into $\mathbb{R}^n$, and the total differential of f is full rank (hence invertible, as it is square) at some point $a \in \mathbb{R}^n$, the theorem states three things:

- For points x within a small enough neighbourhood a, the total differential $Df(x)$ is invertible,
- On this same neighbourhood (denoted by U), $f(U)$ is a bijection,
- the inverse of f is a C^1 map. This makes $f|U$ a C^1-diffeomorphism

Inverse Function Theorem on Manifolds

Let F be a smooth map between two smooth manifolds M and N, with dimensions m and n respectively.

Definition 4.1: Rank of a map

The rank of F at $p \in M$ is the rank of the linear map:

$$dF_p : T_pM \to T_{F(p)}N$$

Definition 4.2: Constant rank maps

A smooth map $F \in C^\infty(M, N)$ has constant rank if its differential $dF_p : T_pM \to T_{F(p)}N$ has the same rank at every point $p \in M$.

There are three types of constant rank maps that are of interest.

Definition 4.3: Smooth submersion

F is a smooth submersion if dF_p is a surjection onto $T_{F(p)}N$ at p-everywhere. That is, $\operatorname{rank} dF_p = \dim T_{F(p)}N = \dim N$

Definition 4.4: Smooth immersion

F is a smooth immersion if dF_p is an injection onto $T_{F(p)}N$ at p-everywhere. That is, $\operatorname{rank} dF_p = \dim T_pM = \dim M$

Definition 4.5: Smooth embedding

F is a smooth embedding if it is a smooth immersion, and it is a homeomorphism onto its range $F(M) \subseteq N$.

Definition 4.6: Local diffeomorphism

F is a local diffeomorphism if every $p \in M$ in its domain induces a neighbourhood $U \subseteq M$ with $F|U : U \to F(U)$ is a diffeomorphism (in the sense of two open sub-manifolds).

Proposition 4.1: Rank as an open condition

Suppose $F : M \to N$ is a smooth map, and $p \in M$. If dF_p is a surjection (resp. injection), pointwise at p, there exists a neighbourhood U of p where $F|U$ is a smooth submersion (resp. immersion)

Proof. Trivial. See Example 1.1. ∎

Proposition 4.2: Inverse Function Theorem on Manifolds

Let M and N be smooth manifolds, and $F : M \to N$ be a smooth map. Suppose the differential of F is invertible at some point $p \in M$, then there exists connected neighbourhoods U_0 of p, and V_0 of $F(p)$ such that $F|U_0 : U_0 \to V_0$ is a diffeomorphism.

Proof. Trivial. See the regular inverse function theorem Proposition 3.1 on Euclidean space, and pass the argument back to the manifolds using coordinate charts. ■

Proposition 4.3: Rank Theorem for Manifolds

Let $F : M \to N$ be a smooth map with constant rank r, then at every $p \in M$, there exists smooth charts $p \in (U, \phi)$ and $F(U) \subseteq (V, \psi)$, where the coordinate representation of F takes the form

$$\hat{F}(x) = \begin{bmatrix} \mathrm{id}_{r\times r} & 0_{r\times m-r} \\ 0_{n-r\times r} & 0_{n-r\times m-r} \end{bmatrix} x, \qquad \text{or equivalently} \tag{12}$$

$$\hat{F}(x^1, \dots, x^r, x^{r+1}, \dots, x^m) = (x^1, \dots, x^r) \tag{13}$$

Proof. Tedious. However, some techinques are worth remembering:

- Passing the argument to the Euclidean case as usual,
- We are free to shrink the sizes of open cubes and balls, and exploit local compactness,
- Suppose we are given a matrix of size $m \times n$, which has rank r, then we can attach a sub-matrix to make it square and invertible, then rehearse the usual arguments with the Inverse Function Theorems Propositions 3.1 and 4.2 to obtain a neighbourhood small enough that preserves the rank of the square matrix. Then pass the argument back to the smaller sub-matrix.

The last bullet point is worth elaborating, suppose we are given a rectangular matrix, where A is square and invertible. Take $z = (x, y)^T$ with dimensions that make the formulas below make sense.

$$Mz = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \implies \begin{bmatrix} M \\ I \end{bmatrix} (z, y)^T = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix} \begin{bmatrix} z \\ y \end{bmatrix} \quad \text{is square and invertible}$$

by Proposition 4.1, we see that there exists a neighbourhood about the square matrix $\begin{bmatrix} M \\ I \end{bmatrix}$ such that it remains invertible, hence a neighbourhood about M that makes A invertible (as a sub-matrix), so the rank of M is preserved. ■

Corollary 4.1: Rank Theorem for Manifolds - Special Cases

Let $F : M \to N$ be a smooth map with constant rank. If F is a smooth immersion, then Equation (12) takes the form:

$$\hat{F}(x) = \begin{bmatrix} \mathrm{id}_{m\times m} \\ 0_{n-m\times m} \end{bmatrix} x, \qquad \text{or equivalently} \tag{14}$$

$$\hat{F}(x^1, \ldots, x^m) = (x^1, \ldots, x^m, 0 \ldots, 0) \tag{15}$$

If F is a smooth submersion,

$$\hat{F}(x) = \begin{bmatrix} \mathrm{id}_{n\times n} & 0_{n\times m-n} \end{bmatrix} x, \qquad \text{or equivalently} \tag{16}$$

$$\hat{F}(x^1, \ldots, x^n, x^{n+1}, \ldots, x^m) = (x^1, \ldots, x^n) \tag{17}$$

More on immersions and embeddings

Proposition 5.1: Characterization of smooth immersions

F is a smooth immersion iff every point $p \in M$ has a neighbourhood $U \subseteq M$ where $F|U : U \to N$ is a smooth embedding.

Proof. We will prove it for when M and N are smooth manifolds, see Lee for the full proof with boundary. It involves extending the argument by composing F with an inclusion map. From Lemma 3.11 (Lee), if $a \in \partial\mathbb{H}^n$, then the differential of the inclusion map $\iota : \mathbb{H}^n \to \mathbb{R}^n$ is a linear isomorphism between tangent spaces.

$$d\iota_a : T_a\mathbb{H}^n \to T_a\mathbb{R}^n, \quad \underbrace{T_a\mathbb{H}^n \cong T_a\mathbb{R}^n}_{\text{isomorphic}}$$

If for every $p \in M$, there exists a neighbourhood U of p with $F|U : U \to N$ a smooth embedding, then $dF|U_p$ has rank m, so dF_p has rank m, and the differential is injective pointwise everywhere. Conversely, if dF_p is a smooth immersion, the Rank Theorem (Proposition 4.3) tells us there exists connected neighbourhoods of p and $F(p)$, where F has coordinate representation in Equation (14) with respect to an appropriate choice of coordinate charts centered at p, so $\hat{F}(\hat{p}) = 0 \in \mathbb{R}^n$. Let $\hat{p} \in \hat{U}$ and $\hat{F}(\hat{p}) \in \hat{V}$, the proof then devolves into a linear-map problem. $\hat{F}$ given by the expression in Equation (14) is clearly injective. Therefore it is bijective onto its range, its inverse is nothing but the map that removes the extra zeroes at the end. Therefore $F|U$ is a smooth embedding. ■

Definition 5.1: Section of $\pi : M \to N$

If $\pi : M \to N$ is a continuous map, a *section of* π is a continuous right inverse for π, i.e $\sigma : N \to M$, $\sigma \in C(N, M)$, $\pi \circ \sigma = \mathrm{id}_N$.

A *local section for* π is a continuous function σ from an open set $U \subseteq V$ into M with $\pi \circ \sigma = \mathrm{id}_U$.

Proposition 5.2: Characterization of smooth submersion

Let $\pi : M \to N$ be smooth, then π is a smooth submersion iff every point of M is in the image of a smooth local section of π.

Proof. Suppose π is a smooth submersion, and fix $p \in M$, by the Rank Theorem Proposition 4.3, and Equation (16), π has the coordinate representation

$$\hat{\pi}(x^1, \ldots, x^n, x^{n+1}, x^m) = (x^1, \ldots, x^n)$$

between two open sets $U \subseteq M$ and $V \subseteq N$, (it really does not matter). Now, define

$$\sigma : V \to M, (x^1, \ldots, x^n) \mapsto \underbrace{(x^1, \ldots, x^n, 0, \ldots, 0)}_{\mathbb{R}^m} \in U$$

the charts by assumption are centered, and $\pi \circ \sigma$ is clearly smooth (check coordinate-wise), so σ reaches p. Conversely, recall if the composition of maps $(g \circ f)$ is a surjection, then g is a surjection. Now, fix $p \in M$, this induces an open set V containing $\pi(p)$, and a smooth local section $\sigma_V : V \to M$. By Proposition 4.1, *the differential of a composition is equal to the composition of the differentials*

$$\mathrm{id}_{T_q N} = d(\mathrm{id}_U) = d(\pi)_{\sigma(q)} \circ d(\sigma)_q$$

so $d(\pi)_{\sigma q} = d(\pi)_p$ is a surjection and the proof is complete. ■

Regular values and level sets

If $F : \mathbf{X} \to \mathbf{Y}$, and $c \in \mathbf{Y}$, we call the $F^{-1}(\{y\})$ a *level set at* c, and c the *level value*. We often write $F^{-1}(y)$ in place of $F^{-1}(\{y\})$. If $\mathbf{Y} = \mathbb{R}^k$, then $F^{-1}(0)$ is the *zero set of* F.

Definition 6.1: Critical point of $F \in C^\infty(N, M)$

$p \in M$ is a *critical point* of F if the differential $dF_p : T_p N \to T_{F(p)} M$ fails to be surjective at p, otherwise p is a *regular point* of F.

Definition 6.2: Critical value of $f \in C^\infty(N, \mathbb{R})$

Let f be a test function on N, $c \in \mathbb{R}$ is a critical value if there exists a $p \in f^{-1}(c)$ where df_p is not surjective.

Otherwise c is called a regular value. Caveat: if c is not in the image of f, we call c a regular value as well. It is clear c is a regular value iff $c \notin f(M)$ or for every

$p \in f^{-1}(c)$, df_p is a surjection.

Notice $df_p : T_pN \to T_{f(p)}\mathbb{R}$ is not surjective $\iff$ all partials of the coordinate representation of f vanish. Since the matrix representation of df_p has the form

$$\mathcal{M}\{df_p\} = \begin{bmatrix} \frac{\partial f}{\partial x^1}\Big|_p & \cdots & \frac{\partial f}{\partial x^m}\Big|_p \end{bmatrix}$$

$\operatorname{rank} \mathcal{M}\{df_p\} \neq 1 \iff \frac{\partial f}{\partial x^j}\Big|_p = 0$ for $1 \leq j \leq m$.

Definition 6.3: Regular level set of $f \in C^\infty(N, \mathbb{R})$

If c is a regular value of f, then $f^{-1}(c)$ is called a regular level set.

Define $g = f - c$, the partials at $p \in N$ for both f and g agree, so the matrix representations of df_p and dg_p are identical (whereas their ranges might not be),

$$\begin{aligned} \mathcal{M}\{dg_p\} &= \begin{bmatrix} \frac{\partial g}{\partial x^1}\Big|_p & \cdots & \frac{\partial g}{\partial x^m}\Big|_p \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f - c}{\partial x^1}\Big|_p & \cdots & \frac{\partial f - c}{\partial x^m}\Big|_p \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x^1}\Big|_p & \cdots & \frac{\partial f}{\partial x^m}\Big|_p \end{bmatrix} \\ &= \mathcal{M}\{df_p\} \end{aligned}$$

Commentary

Proposition 4.1 roughly states that, if the differential of F at some point p is injective or surjective, then there exists a neighbourhood U about p such that $dF|U(p)$ is an injection or surjection. The continuity of the map $dF|U(p) \mapsto \Delta_{m\times m}(dF|U(p))$, induces a neighbourhood in the vector space of matrices about the differential $dF|U(p)$. This vector space is endowed with any of the equivalent norms on $\mathcal{M}(m \times n, \mathbb{R})$, which is equivalent to the entrywise 2-norm. Since all partials of the form $\frac{\partial \hat{F}^k}{\partial x^j}\Big|_{\hat{p}}$ are continuous, we take the intersection over all $n \times m$ partials such that $dF|U(p)$ is an injection or surjection. Finally, send this neighbourhood about $\hat{p}$ through to p by using the continuity of ϕ.

Chapter A: Review of Topology

Topological Spaces

This section will roughly follow Munkres text on Geeneral Topology, in particular we hope to cover Chapters 2, 3, 4 and 9. The rest of the Chapters should be covered proper by the subsequent section.

Definition 1.1: Topology

Let $\mathbf{X}$ be a non-empty set. A topology $\mathcal{T}$ on $\mathbf{X}$, sometimes denoted by $\mathcal{T}_{\mathbf{X}}$ is a family of subsets of $\mathbf{X}$,

- $\{\varnothing, \mathbf{X}\} \subseteq \mathcal{T}$,
- If U_1 and U_2 are elements of $\mathcal{T}$, so is their intersection.
- If $\{U_\alpha\}$ is an arbitrary family of sets in $\mathcal{T}$, their union is also contained in $\mathcal{T}$ as an element.

We call the elements of $\mathcal{T}$ open sets. The complements of elements in $\mathcal{T}$ are closed sets.

Basis of a Topology

Definition 2.1: Basis of a topology

A basis $\mathbb{B}$ is a family of subsets of $\mathbf{X}$, that satisfies:

- Every $x \in \mathbf{X}$ belongs (as an element) in some $V \in \mathbb{B}$.
- If B_1 and B_2 are basis elements, such that their intersection is non-empty. Then every $x \in B_1 \cap B_2$ induces a $B_3 \in \mathbb{B}$ with

$$x \in B_3 \subseteq B_1 \cap B_2$$

This roughly means a basis is 'finitely' fine at every point in x.

If $\mathbb{B}$ is a basis, it 'generates' a topology $\mathcal{T}$ through

$$\mathcal{T} = \left\{ U \subseteq \mathbf{X},\, \forall x \in U,\, x \in B \subseteq U \text{ for some } B \in \mathbb{B} \right\} \tag{18}$$

Notice this is equivalent to $\mathcal{T}$ is the collection of all unions of basis elements in $\mathbb{B}$.

Proposition 2.1

Let $\mathbb{B}$ be a basis as defined in Definition 2.1, then $\mathcal{T}$ as defined in Equation (18) is a valid topology on $\mathbf{X}$. And every member of $\mathcal{T}$ is and is precisely the union of elements in $\mathbb{B}$.

Proof. Every point in $\mathbf{X}$ belongs in some basis element, so $\mathbf{X} \in \mathcal{T}$, so does $\varnothing$. Next, if U_1 and U_2 are in $\mathcal{T}$, then

$$\begin{cases} x \in U_1 \leftrightsquigarrow x \in B_1 \subseteq U_1 \\ x \in U_2 \leftrightsquigarrow x \in B_2 \subseteq U_2 \end{cases} \implies x \in B_3 \subseteq B_1 \cap B_2 \subseteq U_1 \cap U_2$$

for some $B_3 \in \mathbb{B}$, so $\mathcal{T}$ is closed under finite intersections (perhaps after a standard induction argument).

If $\{U_\alpha\} \subseteq \mathcal{T}$, and x belongs in the union of all U_α, then $x \in B_\alpha \subseteq U_\alpha$, which is a subset of the entire union. So the union over U_α is again contained in $\mathcal{T}$, and $\mathcal{T}$ is a topology on $\mathbf{X}$.

It is worth noting that $\mathbb{B} \subseteq \mathcal{T}$. Finally, if $U \in \mathcal{T}$,

$$U = \bigcup_{x \in U} B_x$$

where B_x is the basis element taken to satisfy $x \in B_x \subseteq U$. Every point in U is included in some B_x, and hence is included in the union. For the reverse inclusion, notice the union of subsets of U is again a subset of U.

Now, if $E \subseteq \mathbf{X}$ is the union of basis elements in $\mathbb{B}$, if E is non-empty, then every point $x \in E$ belongs in some B_x. Recycling the previous argument, and we see that E is open in $\mathcal{T}$. If E is empty, we define the 'union' of no sets as the empty set. So $\mathcal{T}$ is precisely the collection of all unions of basis elements $\mathbb{B}$. ■

We are now in a position to compare the relative 'fineness' of topologies.

Definition 2.2: Fineness of topologies

If $\mathcal{T}'$ and $\mathcal{T}$ are both topologies on some non-empty set $\mathbf{X}$. We say $\mathcal{T}'$ is finer than $\mathcal{T}$, or $\mathcal{T}$ is coarser than $\mathcal{T}'$ if

$$\mathcal{T}' \supseteq \mathcal{T}$$

Proposition 2.2

If $\mathbb{B}$ and $\mathbb{B}'$ are bases for $\mathcal{T}'$ and $\mathcal{T}$, the following are equivalent:

- $\mathcal{T}'$ is finer than $\mathcal{T}$,
- If B is an arbitrary basis element in $\mathbb{B}$, then every point $x \in B$ induces a basis element in $\mathbb{B}'$ with
$$x \in \mathbb{B}' \subseteq \mathbb{B}$$

Proof. Suppose $\mathcal{T}'$ is finer than $\mathcal{T}$. Notice $\mathbb{B} \subseteq \mathcal{T}'$ as well. By Equation (18), each $x \in B$ induces a $B' \in \mathbb{B}'$

$$x \in B' \subseteq B$$

Conversely, fix any open set $U \in \mathcal{T}$, and for each $x \in U$,

$$x \in B' \subseteq B \subseteq U$$

Applying Definition 2.1 tells us U is open in $\mathcal{T}'$. ■

The last of the big three 'generating' definitions for topologies will be the sub-basis. It simply means the first condition (but not necessarily) the second, is satisfied in Definition 2.1

Definition 2.3: Sub-basis of a topology

A sub-basis $S \in \mathbb{P}(\mathbf{X})$ is a family of subsets of $\mathbf{X}$ that satisfies one property. Any point x in $\mathbf{X}$ belongs to at least one member of S.

A sub-basis can be upgraded to a basis by collecting all of its finite intersections.

Proposition 2.3

Let S be a sub-basis of $\mathbf{X}$, then the collection of all finite intersections of S forms a basis $\mathbb{B}$ of $\mathbf{X}$.

Proof. Every point in $\mathbf{X}$ lies in some element of S, hence in some element of $\mathbb{B}$. The second basis property is immediate, since $\mathbb{B}$ is closed under finite intersections. ■

Product Topology

We will start with products of a finite collection of topological spaces.

Definition 3.1: Finite Product of Topological Spaces

Let $(\mathbf{X}, \mathcal{T}_{\mathbf{X}})$ and $(\mathbf{Y}, \mathcal{T}_{\mathbf{Y}})$ be topological spaces. The product topology (denoted by $\mathcal{T}_{\mathbf{X}\times\mathbf{Y}}$) on $X \times Y$ is defined as the topology generated by the basis

$$\mathbb{B}_{\mathbf{X}\times\mathbf{Y}} = \left\{ U \times V,\ (U, V) \in \mathcal{T}_{\mathbf{X}} \times \mathcal{T}_{\mathbf{Y}} \right\} \tag{19}$$

Since bases are easier to describe than topologies, we have the following statement concerning the basis of the product topology.

Proposition 3.1

If $\mathbb{B}_{\mathbf{X}}$ and $\mathbb{B}_{\mathbf{Y}}$ are bases for $\mathcal{T}_{\mathbf{X}}$ and $\mathcal{T}_{\mathbf{Y}}$, then the product topology (as described in Definition 3.1) is also generated by

$$\mathcal{M} = \left\{ U \times V,\ (U, V) \in \mathbb{B}_{\mathbf{X}} \times \mathbb{B}_{\mathbf{Y}} \right\} \tag{20}$$

Proof. We will introduce (and use) the technique of 'double inclusion' by proving that the topologies generated are both finer than the other. Let us denote the topology generated by $\mathcal{M}$ in Equation (20) by $\mathcal{T}_{\mathcal{M}}$.

Since $\mathbb{B}_{\mathbf{X}} \times \mathbb{B}_{\mathbf{Y}} \subseteq \mathcal{T}_{\mathbf{X}} \times \mathcal{T}_{\mathbf{Y}}$, if $U \times V \in \mathcal{M}$ as in Equation (20), then we can pick the same 'open rectangle' again. We trivially have

$$x \in \underbrace{U \times V}_{\text{member of } \mathcal{T}_{\mathbf{X}} \times \mathcal{T}_{\mathbf{Y}}} \subseteq U \times V$$

and by Proposition 2.2, $\mathcal{T}_{\mathbf{X}\times\mathbf{Y}}$ is finer than $\mathcal{T}_{\mathcal{M}}$.

Fix any set $U \times V \in \mathbb{B}_{\mathbf{X}\times\mathbf{Y}}$, and if $(p, q) \in U \times V$, each coordinate induces basis elements from $\mathbb{B}_{\mathbf{X}}$ and By, more precisely:

$$\begin{cases} p \in U \implies p \in \text{Basis element of } \mathbb{B}_{\mathbf{X}} \subseteq U \\ q \in V \implies q \in \text{Basis element of } \mathbb{B}_{\mathbf{Y}} \subseteq V \end{cases} \implies (p, q) \in \underbrace{\quad}_{\text{in } \mathbb{B}_{\mathbf{X}}} \times \underbrace{\quad}_{\text{in } \mathbb{B}_{\mathbf{Y}}} \subseteq U \times V$$

by Proposition 2.2, $\mathcal{T}_{\mathcal{M}}$ is finer than $\mathcal{T}_{\mathbf{X}\times\mathbf{Y}}$ and $\mathcal{T}_{\mathbf{X}\times\mathbf{Y}} = \mathcal{T}_{\mathcal{M}}$. ■

Quotient Topology

Product Topology

The Cartesian Product of an arbitrary family of topological spaces, if equipped with the product topology, preserves a lot of the structure. If $\{X_\alpha\}_{\alpha \in A}$ is a family of topological spaces which are ____, then $\prod X_\alpha$ is ____. Replace ____ with:

1. Hausdorff, (Folland)
2. Regular,
3. Connected, (Munkres chp23, exercise 10)
4. First countable, if A is countable,
5. Second countable, if A is countable,
6. Compact (Tynchonoff's Theorem, Folland)

Connectedness

Definition 6.1: Connectedness

A topological space $\mathbf{X}$ is connected if U and V are disjoint open subsets whose union is $\mathbf{X}$, then at least one of U or V is empty.

See Folland Exercise 4.10 for more properties.

Definition 6.2: Path-connectedness

A topological space $\mathbf{X}$ is path-connected if for any two pair of points $x, y \in \mathbf{X}$. There exists a continuous function $f : [a, b] \to \mathbf{X}$, with $f(a) = x$ and $f(b) = y$.

Definition 6.3: Connected component

The connected components of $\mathbf{X}$ is the family of equivalence classes on $\mathbf{X}$, where $x \sim y$ if there is a connected subspace of $\mathbf{X}$ that contains both of them.

Proposition 6.1

Continuous functions map connected spaces to connected spaces (in the subspace topology).

Proof. Let $\mathbf{X}$ and $\mathbf{Y}$ be topological spaces and $f : \mathbf{X} \to \mathbf{Y}$ be continuous. If $f(\mathbf{X})$ is disconnected, then we can find U and V, open and disjoint in $\mathcal{T}_{f(\mathbf{X})}$ such that

$$U \cup V = f(\mathbf{X}) \implies f^{-1}(U) \cup f^{-1}(V) = \mathbf{X}$$

where $f^{-1}(f(\mathbf{X})) = \mathbf{X}$. Both $f^{-1}(U)$ and $f^{-1}(V)$ are open, non-empty, and are pairwise disjoint. So $\mathbf{X}$ is separated. ■

Proposition 6.2

Let $(\mathbf{X}_\alpha, \mathcal{T}_\alpha)$ be a family of connected topological spaces indexed by $\alpha \in A$. Then $\prod_{\alpha \in A} \mathbf{X}_\alpha$ is disconnected in the product topology.

Proof. We will attempt the contrapositive. Suppose $\prod_{\alpha \in A} \mathbf{X}_\alpha$ is disconnected, then ■

Interiors and closures

Definition 7.1: Interior of a set

A^o is defined to be the largest open subset of A,

$$A^o = \bigcup_{\substack{U \text{ open,} \\ U \subseteq A}} U$$

Corollary 7.1

The union of subsets of A is again a subset of A, therefore Corollary 7.1 implies $A^o \subseteq A$ for any $A \subseteq \mathbf{X}$.

Definition 7.2: Closure of a set

and $\overline{A}$ is the smallest closed superset of A,

$$\overline{A} = \bigcap_{\substack{K \text{ closed,} \\ A \subseteq K}} K$$

Proposition 7.1

The complement of the closure is the interior of the complement, or equivalently: $(\overline{A})^c = A^{co}$

Proof. Taking complements, and the substitution $U = K^c$ reads

$$\begin{aligned}
(\overline{A})^c &= \left[\bigcap_{\substack{K \text{ closed,} \\ A \subseteq K}} K \right]^c \\
&= \bigcup_{\substack{K \text{ closed,} \\ K^c \subseteq A^c}} K^c \\
&= \bigcap_{\substack{U \text{ open,} \\ U \subseteq A^c}} U \\
&= A^{co}
\end{aligned}$$

■

Remark 7.1

Personally, I remember this as pushing the complement inside and flipping the bar to a c!

Neighbourhoods

The concept of a neighbourhood allows us to characterize the interior of a set 'locally'.

Definition 8.1: Neighbourhood (not necessarily open)

A neighbourhood of $x \in \mathbf{X}$ is a set $U \subseteq \mathbf{X}$ where $x \in U^o$. The set of neighbourhoods for a point $x \in \mathbf{X}$ will sometimes be denoted by $\mathcal{N}(()x)$.

Proposition 8.1: Characterization of the interior

If $W = \left\{ x \in \mathbf{X}, \text{ there exists a neighbourhood } U \text{ of } x, U \subseteq A \right\}$, then $W = A^o$.

Proof. If $x \in A^o$, then A is a neighbourhood of x, and $A \subseteq A$, so $x \in W$. Conversely, if x is a member of W, it has a neighbourhood $U \subseteq A$ (not necessarily open). By monotonicity of the interior,

$$x \in U^o \subseteq A^o$$

and $x \in A^o$. ■

It is easy to see that A is open $\iff A^o = A \iff A$ is a neighbourhood of itself.

- The first equivalence follows from:

$$E \subseteq \mathbf{X} \implies E^o \subseteq E$$

  and if A is an open set, it is an open subset of itself, by Corollary 7.1 $A \subseteq A^o$. If $A^o = A$, then it suffices to show that A^o is open. Which it is, since it is the arbitrary union of open sets.

- To prove the second equivalence: suppose $A^o = A$, then each $x \in A$ has a neighbourhood contained (as a subset) in A, namely A itself. (This statement is hard to parse, the reader is encouraged to really work through this and be honest).

$$x \in A^o \subseteq A \implies A \subseteq A^o$$

  so A is a neighbourhood of itself. Conversely, if $A \subseteq A^o$, then $A = A^o$, since the reverse inclusion follows immediately from Corollary 7.1.

Adherent points

Similar to the neighbourhood, the concept of an adherent point of a set allows us to speak of the closure in more concrete terms. The following definition is key in understanding the relationship between the closure, interior, and the boundary.

Definition 9.1: Adherent point of a set

Let $A \subseteq \mathbf{X}$, $x \in \mathbf{X}$ is an adherent point of A if every neighbourhood U of x intersects A. In symbols,

$$U \cap A \neq \varnothing, \quad \forall U \in \mathcal{N}(x)$$

Proposition 9.1: Characterization of the closure

Let $A \subseteq X$, and let W be the set of adherent points of A, then $\overline{A} = W$

Proof. Suppose $x \notin W$, then there exists a neighbourhood U of x where

$$U \cap A = \varnothing \iff U \subseteq A^c$$

this is exactly the definition of the interior of A^c, so $x \in A^{co}$ and recall (from Proposition 7.1) that $(\overline{A})^c = A^{co}$, so $x \notin \overline{A}$. For the reverse inclusion, read the proof backwards, by flipping $\forall \to \exists$ within the set, and we see that

$$W^c = A^{co} = (\overline{A})^c$$

■

Dense and nowhere dense subsets

Definition 10.1: Dense subset

A subset of a topological space $E \subseteq \mathbf{X}$ is dense if $\overline{E} = \mathbf{X}$.

Definition 10.2: Nowhere dense subset

A subset of a topological space $E \subseteq \mathbf{X}$ is nowhere dense if $\overline{E}^o = \varnothing$.
This means E is dense in none of the (non-trivial) open subspaces of $\mathbf{X}$.

Proposition 10.1

E is dense in $\mathbf{X}$ iff for every non-empty, open set $U \subseteq \mathbf{X}$, $U \cap E \neq \varnothing$.

Proof of Proposition 10.1. Suppose E is dense, then $\overline{E} = \mathbf{X}$. Every point of $\mathbf{X}$ is an adherent point of E. Let $U \subseteq \mathbf{X}$ be a non-empty open set. If $x \in U$ then U is a neighbourhood of x, thus U intersects E. Conversely, suppose every non-empty open set U intersects E. Fix any point $x \in \mathbf{X}$, and any neighbourhood U of x. U has a non-empty interior (because it must contain x). But U^o is a non-empty open set, therefore $\varnothing \neq U^o \cap E \subseteq U \cap E$ ■

Proposition 10.2

Let $f : \mathbf{X} \to \mathbf{X}$ be a homeomorphism. E is nowhere dense iff $f(E)$ is nowhere dense.

Proof. Since f^{-1} is a homeomorphism, suppose $\overline{f^{-1}(E)}^o \neq \varnothing$, there exists a non-empty, open subset $U \subseteq \mathbf{X}$ with

$$\overline{f^{-1}(E)} \cap U = U$$

The direct image yields

$$f\left((\overline{f^{-1}(E)}) \cap U\right) = f(U)$$

since f is a bijection (injectivity is necessary here), it commutes with intersections.

$$f(\overline{f^{-1}(E)}) \cap f(U) = f\left((\overline{f^{-1}(E)}) \cap U\right) = f(U) \tag{21}$$

and f is continuous, so $f(\overline{A}) \subseteq \overline{f(A)}$ for any $A \subseteq \mathbf{X}$. For the reverse inclusion, f is a closed map, so $f(\overline{A})$ is a closed superset of $f(A)$ so

$$f(\overline{A}) = \overline{f(A)}$$

Take $A = f^{-1}(E)$, and $f(\overline{f^{-1}(E)}) = \overline{f(f^{-1}(E))} = \overline{E}$. From eq. (21), we see that

$$\overline{E} \cap f(U) = f(U)$$

$f(U)$ is a non-empty open subset of $\mathbf{X}$, since f is an open map, so E is not no-where dense. The reverse implication can be proven by replacing f with f^{-1}. ■

Urysohn's Lemma

Proposition 11.1: Folland Theorem 4.14

Suppose that A and B are disjoint closed subsets of the normal space X, and let $\Delta = \{k2^{-n} : n \geq 1 \text{ and } 0 < k < 2^n\}$ be the set of dyadic rationals in $(0, 1)$. There is a family $\{U_r : r \in \Delta\}$ of open sets such that

1. $A \subseteq U_r \subseteq B^c$ for every $r \in \Delta$,
2. $\overline{U_r} \subseteq U_s$ for $r < s$, and
3. For every $r < s$, $\overline{U}_r \subseteq U_s$

Proof. The goal of this proof is to show that for every $r \in \Delta$, there exists a open U_r that satisfies the above. As usual for these types of proofs we will proceed by induction. We can divide the problem by 'layers' (as I will hereinafter explain).

Let us suppose that for some $N \geq 1$ that all previous U_r in previous layers have been constructed properly, meaning if $r = k/2^n$, then for every $1 \leq n \leq N-1$, we have

$$r = \frac{k}{2^n},\ 1 \leq n \leq N-1,\ 1 \leq k \leq 2^{n-1}$$

And by 'constructed properly', we mean that for each U_r,

- $A \subseteq U_r \subseteq B^c$ and
- $U_r \in \mathcal{T}_X$

Then for this fixed layer $N \geq 1$, we only have to construct the $U_{k/2^N}$ for every odd k, this is because if k is an even number, then $k = 2j$ and $r = 2j/2^N = j/2^{N-1}$ and for this particular U_r is already constructed. So for every odd $k = 2j+1$, the sets of the form $U_{(k-1)/2^N}$ and $U_{(k+1)/2^N}$ are already defined, and satisfy

$$A \subseteq \overline{U}_{(k-1)/2^N} \subseteq U_{(k+1)/2^N} \subseteq B^c$$

For every $k-1 \neq 0$ and $k+1 \neq 1$. (We will consider these cases later). We claim that for every pair of open sets, $E_1, E_2 \in \mathcal{T}_X$, then there exists some open set $G \in \mathcal{T}_X$ such that if $(E_1, E_2) \in H \subseteq (\mathcal{T}_X \times \mathcal{T}_X)$ where H is defined as the set

$$H = \left\{(E_1, E_2) \subseteq (\mathcal{T}_X \times \mathcal{T}_X) : \overline{E_1} \cap E_2^c = \varnothing\right\}$$

Then there exists some $G = \mathcal{J}(E_1, E_2) \in \mathcal{T}_X$ such that

$$E_1 \subseteq \overline{E_1} \subseteq G \subseteq \overline{G} \subseteq E_2$$

Now consider any any $(E_1, E_2) \in H$, then this pair induces a pair of disjoint sets $\overline{E_1}$ and E_2^c since

$$\overline{E_1} \subseteq E_2 \implies \overline{E_1} \cap E_2^c = \varnothing$$

And by normality, there exists disjoint open sets G_1, G_2 such that

- $\overline{E_1} \subseteq G_1 \in \mathcal{T}_X$
- $E_2^c \subseteq G_2 \in \mathcal{T}_X$
- $G_1 \cap G_2 = \varnothing \implies G_1 \subseteq G_2^c \subseteq E_2$
- Since G_2^c is a closed set that contains G_1 as a subset, $\overline{G_1} \subseteq G_2^c \subseteq E_2$

It is at this point that we will make no further mention of G_2 (so we may discard the notion of G_2 in our minds). Let us now replace G with G_1 then it is an easy task to verify that $G = G_1 = \mathcal{J}(E_1, E_2)$ has the required properties.

Now define for every odd k, since $(U_{(k-1)/2^N}, U_{(k+1)/2^N}) \in H$ (we note in passing that $\mathcal{J}$ is not a function as the set G may not be unique).

$$U_{k/2^N} = \mathcal{J}\left(U_{(k-1)/2^N}, U_{(k+1)/2^N}\right)$$

Then, if $U_{(k-1)/2^N}$ and $U_{(k+1)/2^N}$ is 'well constructed' we have

$$A \subseteq \overline{U}_{(k-1)/2^N} \subseteq U_{(k+1)/2^N} \subseteq B^c$$

Therefore $U_{k/2^N} = \mathcal{J}(U_{(k-1)/2^N}, U_{(k+1)/2^N})$ sits 'right inbetween' the two sets so that

- $A \subseteq \overline{U}_{(k-1)/2^N} \subseteq U_{k/2^N}$ and
- $\overline{U}_{k/2^N} \subseteq U_{(k+1)/2^N} \subseteq B^c$

Combining the above two estimates will give us a 'well constructed' $U_{k/2^N}$ for every $k-1 \neq 0$ and $k+1 \neq 1$. Now let us deal with the remaining pathological cases.

If $k-1$ so happens to be 0, then no $r \in \Delta$ satisfies $r = 0/2^N$, and we substitute

$$\overline{U}_0 = A, \qquad \text{or alternatively, } U_0 = A^o$$

Then $U_0 \in \mathcal{T}_X$, $\overline{U}_0 = A \subseteq B^c$. It is at this point that we must mention that $0, 1 \notin \Delta$, so U_0 and U_1 do not have to obey the rules we have laid out for $U_{r\in\Delta}$.

Now if $k+1$ is equal to 2^N (this makes $r = (k+1)/2^N = 1$) we define

$$U_1 = B^c \in \mathcal{T}_X$$

With this, for every $0 \leq m \leq 2^N - 1, U_{m/2^N}$ must staisfy

$$\overline{U}_{m/2^N} \subseteq B^c = U_1$$

And the pair $(U_{(k-1)/2^N}, U_{(k+1)/2^N}) \in H$ (even for when $N = 1$, since $A = \overline{U}_0 \subsetneq U_1 = B^c$) and a corresponding $U_{k/2^N} = \mathcal{J}(\cdot, \cdot)$ such that

- $A \subseteq \overline{U}_{(k-1)/2^N} \subseteq U_{k/2^N}$
- $\overline{U}_{(k+1)/2^N} \subseteq B^c$

Now as a final step, we complete the base case for when $N = 1$. We would only have to construct for $k = 1$, since

$$U_{1/2} = \mathcal{J}(U_0, U_1) = \mathcal{J}(A, B^c)$$

Apply the induction step, and the proof is complete, at long last. ■

Proposition 11.2: Folland Theorem 4.15: Urysohn's Lemma

Urysohn's Lemma. Let X be a normal space, if A and B are disjoint closed subsets of X, then there exists a $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B.

Proof. Let $r \in \Delta$ be as in Lemma 4.14, and set U_r accordingly except for $U_1 = X$. Define

$$f(x) = \inf\{k : x \in U_k\}$$

Let us also write $W = \{k : x \in U_k\}$, Then for every $x \in A$ we have $f(x) = 0$, since by the construction of the 'onion' function in Lemma 4.14, for each $r \in \Delta \cap (0, 1)$,

$$x \in A \subseteq U_r \implies f(x) \leq r$$

Since $r > 0$ is arbitrary, and $0 \in W$, we can use a classic ε argument. If $f(x) > 0$ then there exists some $0 < r < f(x)$ by density of the dyadic rationals on the line, if $f(x) < 0$ then this implies that there exists some $f(x) < r < 0$ such that $x \in U_r$, but no $r \in \Delta$ can be negative, hence $f(x) = 0$.

Now, for every $x \in B$, since A and B are disjoint, and $A \subsetneq U_r \subsetneq B^c$, then for every $x \in B$ means that x is not a member of any U_r, but we set $U_1 = X$. Since none of the $r \in (0, 1)$ is a member of the set we are taking the infimum, and $x \in U_1 = X$. The ε argument follows: suppose for every $\varepsilon > 0$, $(1 - \varepsilon) \notin W$, and $1 \in W$, then $f(x) = 1$.

Since $x \in U_1 = X$, for every $x \in X$, $f(x) \leq 1$, and $f(x)$ cannnot be negative as $r > 0$ for every $r \in \Delta$. So $0 \leq f(x) \leq 1$. Now we have to show that this $f(x)$ is continuous. The remainder of the proof is divided into two parts. We would like to show that the inverse images of the half lines are open in X. So $f^{-1}((-\infty, \alpha)) \in \mathcal{T}$ and $f^{-1}((\alpha, +\infty)) \in \mathcal{T}$.

Suppose that $f(x) < \alpha$, so $\inf W < \alpha$, and using the density of Δ, there exists an r, $f(x) < r < \alpha$ such that $x \in U_r$ such that $x \in \bigcup_{r<\alpha} U_r$. So $f^{-1}((-\infty,\alpha)) \subseteq \bigcup_{r<\alpha} U_r$.

Fix an element $x \in \bigcup_{r<\alpha} U_r$, this induces an r such that $\inf W \leq r < \alpha$ therefore $f(x) < \alpha$, and $\bigcup_{r<\alpha} U_r \subseteq f^{-1}((-\infty,\alpha))$.
For the second case, suppose that $f(x) > \alpha$, then $\inf W > \alpha$, and there exists an r (by density) such that $\inf W > r > \alpha$ such that for every $k \in W$, $k \neq r$. Therefore $x \notin U_r$, but by density again, and using the property of the onion function: for every $s < r$ we get $\overline{U_s} \subseteq U_r$, taking complements (which reverses the estimate) — we have $x \notin \overline{U_s}$, but $\left(\overline{U_s}\right)^c$ is open in X. It immediately follows that

$$x \in f^{-1}((\alpha,+\infty)) \implies x \in (U_r)^c \subseteq \left(\overline{U_s}\right)^c \subseteq \bigcup_{s>\alpha} \left(\overline{U_s}\right)^c$$

So $f^{-1}((\alpha,+\infty))$ is a subset of $\bigcup_{s>\alpha} \left(\overline{U_s}\right)^c$. To show the reverse, fix an element x in the union, then this induces some $x \in \left(\overline{U_s}\right)^c \subseteq (U_s)^c$. Then for this $s > \alpha$, $(-\infty, s)$ contains no elements of W. This is because for every $p < s$ implies that $(U_s)^c \subseteq (U_p)^c$, so $p \notin W$. Our chosen s is a lower bound for W, and $\alpha < s \leq \inf W = f(x)$.

Since all of the inverse images from the generating set of $(\mathbb{R}, \mathcal{T}_{\mathbb{R}})$ are open in X, using Theorem 4.9 finishes the proof. ■

Notes on the construction of the countable 'onion' sequence within a normal space $\mathbf{X}$.

If $\mathbf{X}$ is a normal space, and A and B are disjoint closed subsets, then we can easily find an open U with

$$A \subseteq U \subseteq \overline{U} \subseteq B^c \tag{22}$$

We say that U hides in B^c if the closure of U is contained in B^c. Define $\Delta_n = \left\{k2^{-n},\ 1 < k < 2^n\right\}$, so that $\Delta_n \subseteq (0,1)$ for all $n \geq 1$. Notice

$$\Delta_1 \supseteq \cdots \supseteq \Delta_n \supseteq \Delta_{n+1}$$

and the even indices for Δ_{n+1} are contained in Δ_n. Suppose Δ_n is well defined, it suffices to choose the odd indices for Δ_{n+1}. If $r = j2^{-(n+1)}$, where j is odd, then r sits in between precisely two elements in $\Delta_n \cup \{0,1\}$. If r sits between an endpoint, then define $\overline{U_0} = A$, and $B^c = U_1$. And denote the closest left and neighbours by s, t respectively. If $s < r < t$, it is clear that $\overline{U_s}$ and U_t^c are disjoint closed sets.

Use the 'normal space' construction to obtain an superset of $\overline{U_s}$ that hides in U_t, denote this open set by U_r, and similar to Equation (22)

$$\overline{U_s} \subseteq U_r \subseteq \overline{U_r} \subseteq U_t$$

Now that the construction of this sequence is complete, we wish to prove Urysohn's Lemma. Let A and B be disjoint closed sets. And define

$$f(x) = \inf\left\{r \in \Delta \cup \{1\},\ x \in U_r\right\}$$

where $U_1 = \mathbf{X}$. So that $0 \leq f(x) \leq 1$ is immediate. If $x \in A$, then x is in all U_r, and by density of $\Delta \subseteq (0,1)$, we have $f(x) = 0$. Conversely, if $x \in B$ then $x \notin U_r$ for all $r \in \Delta$, if E denotes the indices in Δ where $x \in U_s$ when $s \in E$,

$$(-\infty,\ r) \subseteq E^c \iff E \subseteq [r,\ +\infty) \iff \inf(E) \geq r \tag{23}$$

Send $r \to 1$ and $f(x) = 1$. Thus $f|_A = 0$ and $f|_B = 1$.

To show continuity, it suffices to show that the inverse images of the open half $\left\{(x > \alpha),\ (x < \alpha)\right\}_{\alpha \in \mathbb{R}}$ lines are indeed open in $\mathbf{X}$. Let α be fixed. And if $x \in \{f < \alpha\}$, we can 'wiggle' the infimum towards the right (towards α), and using density of Δ within $(0,1)$, there exists a $r \in E$ that satisfies $f(x) < r < \alpha$. This is equivalent to

$$x \in \bigcup_{r<\alpha} U_r$$

If there exists an $r < \alpha$ st x belongs to U_r as an element, then $f(x) \leq r < \alpha$.

If $f(x) > \alpha$, then $(-\infty,\ \alpha) \subseteq E^c$, by Equation (23). Suppose $\alpha < 1$, otherwise $\{f > \alpha\} = \varnothing$. Wiggle $f(x)$ to the left and obtain an $r \in \Delta$, $\alpha < r < f(x)$ with $x \notin U_r$. By density again, take any $s < r$ by a small amount (st $s > \alpha$, $s \in \Delta$), and

$$\overline{U}_s \subseteq U_r \iff U_r^c \subseteq \overline{U}_s$$

so that $x \in \overline{U}_s^c$ for some $s > \alpha$. This is equivalent to

$$x \in \bigcup_{s>\alpha} \overline{U}_s^c$$

Conversely, if $x \notin \overline{U}_s^c$ for some $s > \alpha$, since $\{U_r\}$ (thus $\{\overline{U}_r\}$) is increasing, and $x \notin U_r$ for every $r \leq s$. Hence,

$$(-\infty,\ s] \subseteq E^c \iff E \subseteq (s,\ +\infty) \iff f(x) \geq s > \alpha$$

Compactness

Compactness is one of the most important concepts in topology and analysis.

Definition 12.1: Compact topological space

A topological space $\mathbf{X}$ is compact if every open covering $\{U_\alpha\}$ contains a finite subcover. That is, if $\{U_\alpha\}$ is an arbitrary collection of open sets, then

$$\mathbf{X} = \bigcup U_{\alpha \in A} \implies \bigcup_{j \leq n} U_{\alpha_j}$$

Definition 12.2: Compact set

$E \subseteq \mathbf{X}$ is compact if it is compact in the subspace topology.

Definition 12.3: Precompact set

$E \subseteq \mathbf{X}$ is precompact if its closure is compact (as a subset).

Definition 12.4: Paracompact space

A topological space $\mathbf{X}$ is paracompact if every open covering of $\mathbf{X}$ has a locally finite open refinement that covers $\mathbf{X}$.

Definition 12.5: Locally finite collection of sets

Let $\mathcal{A}$ be a collection of subsets of $\mathbf{X}$. It is called locally finite, if at every point $p \in \mathbf{X}$, we can find a neighbourhood U of p (not necessarily open), that intersects only finitely many members of $\mathcal{A}$. In symbols,

$$U \cap E = \varnothing \quad \text{for all but finitely many} \quad E \in \mathcal{A}$$

We do not require $\mathcal{A}$ to be a cover of $\mathbf{X}$, nor do we require $\mathcal{A}$ to be a collection of open sets.

Definition 12.6: Countably locally finite

A collection $\mathbb{B}$ is countably locally finite if it is the countable union of locally finite families.

$$\mathbb{B} = \overbrace{\bigcup_{\mathbb{N}}}^{\substack{\text{countable}\\ \text{union}}} \mathbb{B}_n, \quad \text{where each } \mathbb{B}_n \text{ is a locally finite collection}$$

Definition 12.7: Refinement

If $\mathcal{A}$ is a collection of sets, $\mathbb{B}$ is a refinement of $\mathcal{A}$ if every element $B \in \mathbb{B}$, induces an element $A \in \mathcal{A}$, such that $B \subseteq A$.

Remark 12.1: Intuition for refinements

If $\mathbb{B}$ is a refinement of $\mathcal{A}$, we can use the 'absolute continuity' muscle. For each element in $\mathbb{B}$ is dominated by some element (through subset inclusion) in $\mathcal{A}$. Recall, if ν and μ are non-negative measures, then $\nu \ll \mu$ if for every measurable set $E \in \mathcal{M}$, $\mu(E) = 0 \implies \nu(E) = 0$.

A refinement of a family of sets is another family of sets, whose elements are dominated by some other element in the un-refined family. *Refining families makes them 'smaller', cover less area.*

Proposition 12.1

Compact Hausdorff spaces are normal, compact subsets of Hausdorff spaces are closed, and closed subsets of compact sets are again compact.

Properties of Compact Spaces

Proposition 13.1

Let $\mathbf{X}$ and $\mathbf{Y}$ be topological spaces.

(a) If $F \in C(\mathbf{X}, \mathbf{Y})$, and $\mathbf{X}$ is compact, then $F(\mathbf{X})$ is compact.

(b) If $\mathbf{X}$ is compact and $F \in C(\mathbf{X}, \mathbb{R})$, then $F(\mathbf{X})$ is bounded, and F attains its supremum and infimum on $\mathbf{X}$.

(c) A finite union of compact subspaces of $\mathbf{X}$ is again compact.

(d) If $\mathbf{X}$ is Hausdorff, and A, B are disjoint, compact subspaces of $\mathbf{X}$, there exists open U and V, (see fig. 2).

(e) Every closed subset of a compact space is compact.

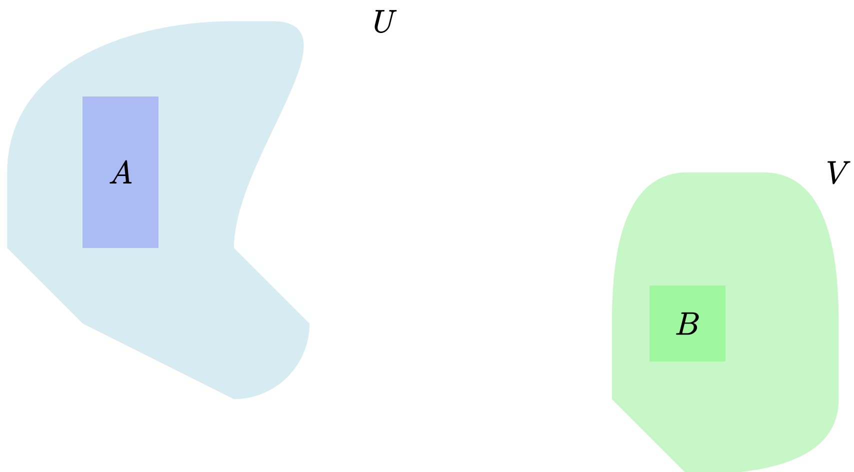

Figure 2: Closed sets A and B within open sets U and V, respectively.

(f) Every compact subset of a Hausdorff space is closed.

(g) Every compact subset of a metric space is bounded.

(h) Every finite product of compact spaces is compact.

(i) Every quotient of a compact space is compact.

Proof of Proposition 13.1 Part A. Let $f \in C(\mathbf{X}, \mathbf{Y})$ with $\mathbf{X}$ compact. Fix an open cover of $f(\mathbf{X})$ in the relative topology,

$$\{U_\alpha \cap f(\mathbf{X})\}_{\alpha \in A} \text{ covers } \mathbf{X},\ U_\alpha \text{ open in } \mathbf{Y}$$

So that $\bigcup f^{-1}(U_\alpha) = \bigcup f^{-1}(U_\alpha \cap f(\mathbf{X})) = \mathbf{X}$. Since $\{f^{-1}(U_\alpha)\}_{\alpha \in A}$ is an open cover for $\mathbf{X}$, this induces a finite subcollection of indices $\{\alpha_1, \ldots, \alpha_n\}$ with

$$\bigcup_{j=1}^{n} f^{-1}(U_{\alpha_j}) = \bigcup_{j=1}^{n} f^{-1}(U_{\alpha_j} \cap f(\mathbf{X}))$$

The direct image commutes with unions, therefore

$$f(\mathbf{X}) = f\left(\bigcup_{j=1}^{n} f^{-1}(U_\alpha \cap f(\mathbf{X}))\right) = \bigcup_{j=1}^{n} f\left(f^{-1}(U_{\alpha_j})\right) = \bigcup_{j=1}^{n} U_{\alpha_j}$$

■

Proof of Proposition 13.1 Part B. Let $\mathbf{X}$ be compact, and $f \in C(\mathbf{X}, \mathbb{R})$, so that $f(\mathbf{X}) \subseteq \mathbb{R}$ is compact. Compact subsets are closed and bounded in $\mathbb{R}$, let $A = \sup f(\mathbf{X})$ and $B = \inf f(\mathbf{X})$. Both A and B are accumulation points of $f(\mathbf{X})$, so $A = f(x)$ and $B = f(y)$ for some x, y in $\mathbf{X}$. ■

Proof of Proposition 13.1 Part C. Let $\mathbf{X}$ be a topological space, and $K_1, \ldots K_n$ be compact subspaces. Denote $K = \bigcup_{j=1}^{n} K_j$. Let $\{U_\alpha \cap K\}_{\alpha \in A}$ be an open cover for K, where U_α is open in $\mathbf{X}$. We can pass the argument to each individual K_j as follows. Let $1 \le j \le n$, then $\{U_\alpha \cap K_j\}_{\alpha \in A}$ is an oepn cover for K_j, so there exists as finite subcollection of indices $I_j \subseteq A$, (a finite subset of A) whose open sets cover K_j. Repeat this process for each j and

$$I = \bigcup_{j=1}^{n} I_j \text{ is a finite subset of } A$$

with $K_j \subseteq \bigcup_{\alpha \in I_j}(U_\alpha \cap K_j) \subseteq \bigcup_{\alpha \in I_j}(U_\alpha \cap K)$. Taking the union over all K_j reads

$$K = \bigcup_{j=1}^{n} K_j \subseteq \bigcup_{j=1}^{n} \bigcup_{\alpha \in I_j} (U_\alpha \cap K) = \bigcup_{\alpha \in I} U_\alpha \cap I$$

■

Proof of Proposition 13.1 Part D. Let $\mathbf{X}$ be Hausdorff. We first prove that compact subspaces of $\mathbf{X}$ are closed. Indeed, if K is compact in $\mathbf{X}$, fix any $x \in K^c$. Let y range through the elements of K, then $x \neq y$ induces a pair of disjoint open sets U_y and V_y, such that

- $x \in U_y$

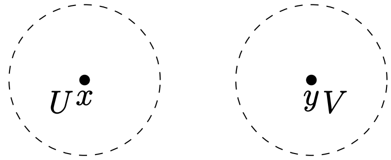

Figure 3: In a Hausdorff space, any two distinct points x and y can be separated by disjoint open neighbourhoods U and V.

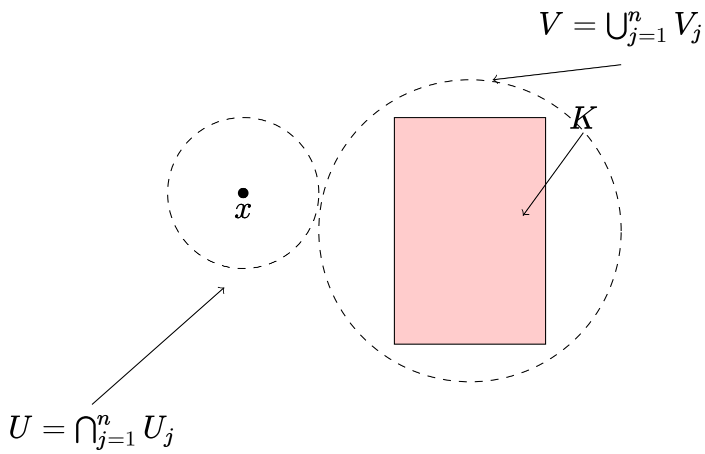

Figure 4: Compact sets are closed in Hausdorff spaces

- $y \in V_y$
- $U_y \cap V_y = \varnothing$
- See fig. 3

Let V_y range through all possible $y \in K$, So that $\{V_y\}_{y \in K}$ is an open cover. There exists a finite subcollection of 'anchor points' of K, $y_1, \dots, y_n$ that corresponds with $\{V_{y_j}\}_{j=1}^{n}$. A finite intersection of open sets is again open, so

$$U = \bigcap_{j=1}^{n} U_{y_j} \text{ is open}$$

Define $V = \bigcup_{j=1}^{n} V_{y_j}$, so $V \subseteq K$ and $U \cap V = \varnothing$ and $x \in U \subseteq K^c$ (see fig. 4). Therefore K is closed.

Finally, if A and B are disjoint compact sets, then each $x \in A \subseteq B^c$ induces neighbourhoods $x \in U_x$, and $B \subseteq V_x$ (see fig. 5), let x range through all the elements of A. By compactness of A, this produces a finite subcover, and

$$U = \bigcup_{j=1}^{n} U_{x_j} \quad V = \bigcap_{j=1}^{n} V_{x_j}$$

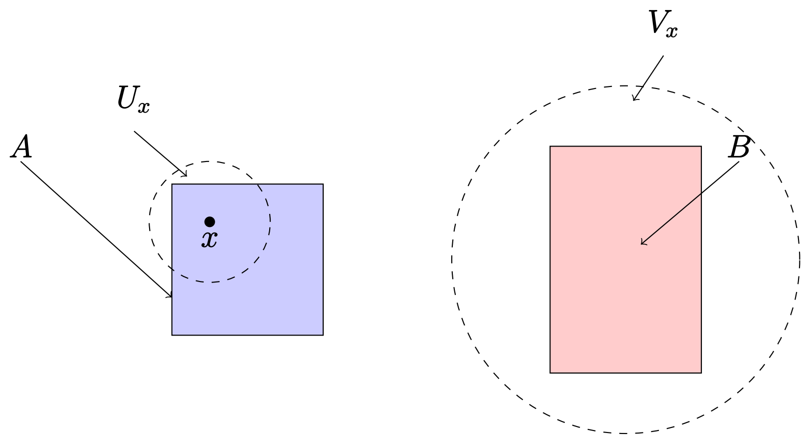

Figure 5: Closed sets A and B, point x in A, and disjoint neighbourhoods U around x and V around B.

are disjoint open sets that contain A and B respectively.

■

Proof of Proposition 13.1 Part E. Let $K \subseteq \mathbf{X}$ be a closed set of a compact space. Let $\{U_\alpha \cap K\}$ be an open cover for K, where each U_α is open in $\mathbf{X}$. We can append an extra set K^c which is open in $\mathbf{X}$. The collection

$$W = \{U_\alpha\} \cup \{K^c\} \text{ covers } \mathbf{X}$$

so there exists a finite subcollection of $W_1, \ldots, W_n$ that cover $\mathbf{X}$ (since $\mathbf{X}$ is compact by itself). Remove K^c from this finite subcollection if it exists, and take the intersection with K for each element W_j, and

$$\{W_1 \cap K, \ldots, W_n \cap K\} = \{U_1 \cap K, \ldots, U_n \cap K\} \text{ covers } K$$

so K is compact. ■

Proof of Proposition 13.1 Part F. Proven in Part D. ■

Proof of Proposition 13.1 Part G. let $K \subseteq \mathbf{X}$ be a compact subset of the metric space $(\mathbf{X}, d)$. Compact subsets of $\mathbf{X}$ are totally bounded, and hence bounded. ■

Proof of Proposition 13.1 Part H. See Tynchonoff's Theorem in Folland Chapter 4. ■

Proof of Proposition 13.1 Part I. Let $\mathbf{X}$ and $\mathbf{Y}$ be topological spaces and $\pi : \mathbf{X} \to \mathbf{Y}$ be a quotient map. So that $\mathbf{Y}$ is endowed with the quotient topology. So that π is a surjective continuous map. and $\pi(\mathbf{X}) = \mathbf{Y}$. Apply Part A, and we see that $\mathbf{Y}$ is compact. ■

Locally Compact Hausdorff Spaces

Compactness is an intrinsic topological property (in the subspace topology). We see from Proposition 4.25 that compact Hausdorff spaces are normal, which gives a sufficient condition for us to approximate and extend any continuous function; and allows us to extend certain 'local' properties to 'global' properties.

If given a Hausdorff space, not necessarily compact, the natural question is to ask 1) whether a topological space has 'enough' compact subsets to work with, and 2) whether we can embed a given topological space in a larger one to force it to be compact.

Definition 14.1: LCH space

Let $\mathbf{X}$ be a Hausdorff space. We call $\mathbf{X}$ a LCH space if every point $p \in \mathbf{X}$ admits a compact neighbourhood. That is, a compact set K whose interior contains p.

We note in passing that the above definition differs slightly from the usual 'local' definitions.

Definition 14.2: Locally connected

Let $\mathbf{X}$ be a topological space, it is locally connected if for every $x \in \mathbf{X}$, and open neighbourhood U containing x, there exists a connected, open neighbourhood V of x such that $x \in V \subseteq U$.

Definition 14.3: Locally path-connected

Let $\mathbf{X}$ be a topological space, it is locally connected if for every $x \in \mathbf{X}$, and open neighbourhood U containing x, there exists a path-connected, open neighbourhood V of x such that $x \in V \subseteq U$.

Definition 14.4: Local homeomorphism

$\mathbf{X}$ locally homeomorphic to $\mathbb{R}^n$ if every point $x \in \mathbf{X}$ belongs to a coordinate chart (U, ϕ), where U is an open neighbourhood of x and ϕ is a homeomorphism from $U \to \phi(U) \subseteq \mathbb{R}^n$.

Definition 14.5: Local diffeomorphism

Let M be a smooth manifold and $F \in C^\infty(M, N)$. F is a local diffeomorphism if every $p \in M$ in its domain induces a neighbourhood $U \subseteq M$ with $F|U : U \to F(U)$ is a diffeomorphism (in the sense of two open sub-manifolds).

Chapter B: Abstract Algebra

Groups

Definition 1.1: Semigroups, Monoids

A non-empty set G equipped with an associative binary operation $G \times G \to G$ is called a semigroup. For every $a, b, c \in G$, we have

$$a(bc) = (ab)c \tag{24}$$

A *monoid* is a semigroup G which contains a *two-sided identity* element $e \in G$ such that $ae = ea$ for all $a \in G$. (not necessarily unique)

Monoids admit unique two-sided identities.

Lemma 1.1: Monoids: unique identity

Let e and i be two-sided identities for a monoid G, then

Proof.

$$e = ei = i$$

■

Definition 1.2: Group

A semigroup G is a group if every element $a \in G$ admits a two-sided inverse a^{-1}. (not necessarily unique)

$$aa^{-1} = a^{-1}a = e$$

Proposition 1.1: Properties of Groups (Hungerford: Theorem 1.2)

Let G be a group with identity e, which is unique by lemma 1.1. Then

(i) $c \in G$ and $cc = c$ implies $c = e$.

(ii) Left/Right cancellation:

$$\begin{cases} ab = ac \implies b = c \\ ba = ca \implies b = c \end{cases}$$

(iii) If $a \in G$, its two-sided inverse is unique.

(iv) Let $a \in G$, then the inverse of its two-sided inverse (uniqueness guaranteed by iii), is a itself; or $(a^{-1})^{-1} = a$.

(v) If $a, b \in G$, then the following equations in x, y admit unique solutions

$$\begin{cases} ax = b \\ ya = b \end{cases}$$

Proof of Proposition 1.1.

Proof of Part (i):

$$cc = c \implies (cc)c^{-1} = cc^{-1} \implies c(cc^{-1}) = e \implies ce = c = e$$

Proof of Part (ii): First claim:

$$\begin{aligned} ab = ac &\implies a^{-1}(ab) = a^{-1}(ac) \\ &\implies (a^{-1}a)b = (a^{-1}a)c \implies eb = ec \implies b = c \end{aligned}$$

Second claim is the same, just cancel from the right using $aa^{-1} = e$ and associativity.

Proof of Part (iii): Suppose b and c are two-sided inverse for a, it follows from Part ii that

$$ab = ac \implies b = c = a^{-1}$$

Proof of Part (iv): From Part iii, the two-sided inverses of group elements exist and are unique, and $a^{-1}a = aa^{-1}$ so a is an inverse for a^{-1}, and it is the only inverse.

Proof of Part (v): First equation: write $ax = b = a(a^{-1}b)$, left-cancelling reads $x = a^{-1}b$, uniqueness follows from Part ii. Second equation is similar. ∎

Lemma 1.2: Group: equality lemma

For any pair of elements $a, b \in G$, $a = b \iff ab^{-1} = e$.

Proof. ($\implies$): $a = b \implies ab^{-1} = bb^{-1} = e$. ($\impliedby$): $ab^{-1} = e \implies a(b^{-1}b) = eb \implies a = eb = b$. ∎

Proposition 1.2: Semigroup: upgrade to group (Hungerford Proposition 1.3)

Let G be a semigroup, G is also a group iff both of the conditions below hold

- Existence of a left-identity: there exists $e \in G$ for every $a \in G$, $ea = a$.

- Existence of left-inverses: for every $a \in G$, there exists a $a^{-1} \in G$ with $a^{-1}a = e$, where e is any left-identity element.

Proof. ($\Longleftarrow$) is trivial. Suppose both conditions hold, notice the proof for Proposition 1.1 Part (i) we only used left-cancellation. $cc = c \implies e$. To prove a^{-1} is also a right-inverse for a, we can force it as follows:

$$(aa^{-1})(aa^{-1}) = a(a^{-1}a)a^{-1} = aea^{-1} = e \implies aa^{-1} = e$$

and a^{-1} is also a right-inverse, so every element $a \in G$ admits a two-sided inverse denoted by a^{-1}. To show e is also a right-identity for any arbitrary element $a \in G$,

$$\begin{aligned} ae &= a(a^{-1}a) && \text{left inverse} \\ &= (aa^{-1})a && \text{associativity} \\ &= ea && \text{right inverse} \\ &= a && \text{left identity} \end{aligned}$$

■

Proposition 1.3: Semigroup: upgrade to group (Hungerford Proposition 1.4)

Let G be a semigroup, G is a group iff for every pair of elements $a, b \in G$, the equations in x and y

$$\begin{cases} ax = b \\ ya = b \end{cases} \tag{25}$$

have solutions (not necessarily unique).

Proof. If G is a group, the existence of the solutions to eq. (25) follow from Proposition 1.1. We will attempt the contrapositive. Suppose G has no left identity, for every $e \in G$ we can always find an element $a \in G$ such that $ea \neq a$, but this is precisely the (first) equation for $a = a$ and $b = a$.

Now suppose G has a left identity element (not necessarily unique). Fix $e \in G$ as any left-identity, and suppose there is an element $a \in G$ with no left inverse, so for every $b \in G$, $ba \neq e$. But b is precisely the solution to the (second) equation with parameters $a = a$ and $b = e$. The negation of Proposition 1.2 is precisely the negation of Proposition 1.3, and the proof is complete. ■

Proposition 1.4: Hungerford Theorem 1.5

Let $R/\sim$ be an equivalence relation on a group G, such that it 'preserves' the group multiplication. More precisely,

$$\begin{cases} a_1 \sim a_2 \\ b_1 \sim b_2 \end{cases} \implies a_1 b_1 \sim a_2 b_2$$

Then the set G/R of all equivalence classes of G under R is a monoid under the binary operation defined by

$$(\overline{a})(\overline{b}) = \overline{ab} \qquad \text{reads: the product of two classes is the class containing the product of any pair of elements from the two classes} \tag{26}$$

where $\overline{a}$ denotes the equivalence class containing a. If G is a group, so is G/R, if G is an abelian group, so is G/R.

Proof. First, notice the binary operation in Equation (26) is well defined. It is independent of the equivalence class representatives chosen, as we have restriction on R that 'forces' the operation on G/R to be well defined. Indeed, let $\overline{a}$ and $\overline{b}$ be elements of G/R, if $a_1, a_2 \in \overline{a}$, and $b_1, b-2 \in \overline{b}$, by definition of R:

$$a_1 \sim a_2 \quad \text{and} \quad b_1 \sim b_2$$

by Equation (26), $a_1 b_1 \sim a_1 b_2 \implies \overline{a_1 b_1} = \overline{a_2 b_2}$.

Associativity is proven similarly, fix $\overline{a}, \overline{b}, \overline{c} \in G/R$, we pass the argument to any of the representatives of the three classes, so

$$(\overline{a}\overline{b})\overline{c} \triangleq \overline{ab}\overline{c} = \overline{(ab)c} = \overline{a(bc)} \triangleq \overline{a}\overline{bc} = \overline{a}(\overline{b}\overline{c})$$

Pass the argument to the representatives, let e denote the identity element in G, it is easily shown that $\overline{e}$ is the identity element in G/R, similarly for two-sided inverses and commutativity of the binary operation. ∎

Homomorphisms and Subgroups

Definition 2.1: Homomorphism

Let G and H be semigroups, $f : G \to H$ is a semi-group *homomorphism* if for all $a, b \in G$,

$$f(ab) = f(a)f(b) \tag{27}$$

Definition 2.2: Monomorphism

Injective homomorphism.

Definition 2.3: Epimorhpism

Surjective homomorphism.

Definition 2.4: Isomorphism

Bijective homomorphism.

Definition 2.5: Endomorphism

Homomorphism for which the domain and codmain (not the range) are equal; i.e $H = G$.

Definition 2.6: Automorphism

Bijective endomorphism.

Definition 2.7: Kernel of a homomorphism

The kernel of $f \in \mathrm{Hom}(G, H)$ is defined

$$\mathrm{Ker}\, f = \left\{ a \in G,\ f(a) = e \in H \right\} \tag{28}$$

as the set of elements in G that get sent to the identity of H.

Proposition 2.1: Hungerford Theorem 2.3

Let G and H be groups and let $f \in \mathrm{Hom}(G, H)$. Denote the identity elements of G and H by e_G and e_H

(i) $f(e_G) = e_H$,

(ii) $f(a^{-1}) = (f(a))^{-1}$ for every $a \in G$.

(iii) f is a monomorphism iff $\ker f = \{e_G\}$,

(iv) f is an isomorphism iff there exists a homomorphism $f^{-1} : H \to G$ that is also a two-sided inverse for f. In symbols:

$$f \circ f^{-1} = \mathrm{id}_H \quad \text{and} \quad f^{-1} \circ f = \mathrm{id}_G \tag{29}$$

Proof of Proposition 2.1.

Proof of Part (i): We will use Proposition 1.1 (i). Since $f(e_G) = f(e_G e_G) = f(e_G)f(e_G)$ in H, we see that $f(e_G) = H$ and $e_G \in \operatorname{Ker} f$

Proof of Part (ii): Let $a \in G$ be arbitrary, using Part (i), we can 'pass the multiplication' between $f(a)$ and $f(a^{-1})$ into G,

$$f(a)f(a^{-1}) = f(e_G) = e_H \implies f(a^{-1}) = (f(a))^{-1}$$

Proof of Part (iii): Suppose $\ker f = e_G$. Let $a, b \in G$ such that $f(a) = f(b)$. The equality lemma Lemma 1.2 tells us $(f(a))^{-1} = f(b)$ and $b = a^{-1}$, so $a = b$ by the Lemma again; f is injective.

Conversely, suppose f is injective, Part (i) tell us $\{e_G\} \subseteq \ker f$. Suppose $a \in \ker f \subseteq G$, but $e_G \in \ker f$, so $f(a) = f(e_G) = e_H$ forces ae_G, and $\ker f = \{e_G\}$.

Proof of Part (iv): ($\Longleftarrow$) is trivial since the existence of a (functional) two-sided inverse is equivalent to bijectivity. Suppose f is an isomorphism, and define f^{-1} as its two-sided (functional) inverse, it suffices to show that $f^{-1} \in \operatorname{Hom}(H, G)$. Fix $f(a)$ and $f(b)$ as arbitrary elements in H. We can do this because f is a bijection, so every element in H has a unique 'representative' in G.

$$f^{-1}(f(a))\, f^{-1}(f(b)) = ab = f^{-1}(f(ab)) = f^{-1}(f(a)f(b))$$

■