3 Transversality

We continue to use the constant rank theorem to produce more manifolds, except now these will be cut out only *locally* by functions. Globally, they are cut out by intersecting with another submanifold. You should think that intersecting with a submanifold locally imposes a number of constraints equal to its codimension.

The problem is that the intersection of submanifolds need not be a submanifold; this is why the condition of transversality is so important - it guarantees that intersections are smooth.

Two subspaces $K, L \subset V$ of a vector space V are *transverse* when $K + L = V$, i.e. every vector in V may be written as a (possibly non-unique) linear combination of vectors in K and L. In this situation one can easily see that $\dim V = \dim K + \dim L - \dim K \cap L$, or equivalently

$$\mathrm{codim}(K \cap L) = \mathrm{codim} K + \mathrm{codim} L. \tag{49}$$

We may apply this to submanifolds as follows:

Definition 3.1. Let $K, L \subset M$ be regular submanifolds such that every point $p \in K \cap L$ satisfies

$$T_pK + T_pL = T_pM. \tag{50}$$

Then K, L are said to be *transverse* submanifolds and we write $K \pitchfork L$.

Proposition 3.2. *If $K, L \subset M$ are transverse submanifolds, then $K \cap L$ is either empty, or a submanifold of codimension* $\mathrm{codim} K + \mathrm{codim} L$.

Proof. Let $p \in K \cap L$. Then there are neighbourhoods U, V of p for which $K \cap U = f^{-1}(0)$ for 0 a regular value of a function $f : U \longrightarrow \mathbb{R}^{\mathrm{codim} K}$ and $L \cap V = g^{-1}(0)$ for 0 a regular value of a function $g : V \longrightarrow \mathbb{R}^{\mathrm{codim} L}$.

Then p must be a regular point for $(f, g) : U \cap V \longrightarrow \mathbb{R}^{\mathrm{codim} K + \mathrm{codim} L}$, since the kernel of its derivative at p is the intersection $\ker Df(p) \cap \ker Dg(p)$, which is exactly $T_pK \cap T_pL$, which has codimension $\mathrm{codim} K + \mathrm{codim} L$ by the transversality assumption, implying $D(f, g)(p)$ is surjective. Therefore $(f, g)^{-1}(0, 0) = f^{-1}(0) \cap g^{-1}(0) = K \cap L \cap U \cap V$ is a submanifold. Since this is true for all $p \in K \cap L$, we obtain that $K \cap L$ is a submanifold of M, as required. Since $T_p(K \cap L) = T_pK \cap T_pL$, we see that $K \cap L$ has codimension $\mathrm{codim} L + \mathrm{codim} K$. $\square$

Example 3.3 (Exotic spheres)**.** Consider the following intersections in $\mathbb{C}^5 \backslash 0$:

$$S_k^7 = \{z_1^2 + z_2^2 + z_3^2 + z_4^3 + z_5^{6k-1} = 0\} \cap \{|z_1|^2 + |z_2|^2 + |z_3|^2 + |z_4|^2 + |z_5|^2 = 1\}. \tag{51}$$

This is a transverse intersection, since the complex vector field

$$\tfrac{1}{2} z_1 \partial_{z_1} + \tfrac{1}{2} z_2 \partial_{z_2} + \tfrac{1}{2} z_3 \partial_{z_3} + \tfrac{1}{3} z_4 \partial_{z_4} + \tfrac{1}{6k-1} z_5 \partial_{z_5}$$

(or its real part, if you prefer), is tangent to the first submanifold but not to the second (which is a hypersurface). Note that if a complex coordinate

has real and imaginary parts $z = x + iy$, then its associated vector field $\partial_z = \frac{\partial}{\partial z}$ is given by
$$\tfrac{\partial}{\partial z} = \tfrac{1}{2}\left(\tfrac{\partial}{\partial x} - i\tfrac{\partial}{\partial y}\right),$$
so that $\partial_z(z) = 1$ and $\partial_z(\bar{z}) = 0$.

For $k = 1, \ldots, 28$ the intersection is a smooth manifold homeomorphic to S^7. These exotic 7-spheres were constructed by Brieskorn and represent each of the 28 diffeomorphism classes on S^7.

We may choose to phrase the previous transversality result in a slightly different way, in terms of the embedding maps k, l for K, L in M. Specifically, we say the maps k, l are transverse in the sense that $\forall a \in K, b \in L$ such that $k(a) = l(b) = p$, we have $\operatorname{im}(Dk(a)) + \operatorname{im}(Dl(b)) = T_pM$. The advantage of this approach is that it makes sense for any maps, not necessarily embeddings.

Definition 3.4. Two maps $f : K \longrightarrow M$, $g : L \longrightarrow M$ of manifolds are called *transverse* when $\operatorname{im}(Df(a)) + \operatorname{im}(Dg(b)) = T_pM$ for all a, b, p such that $f(a) = g(b) = p$.

Proposition 3.5. *If $f : K \longrightarrow M$, $g : L \longrightarrow M$ are transverse smooth maps, then $K\,{}_f{\times}_g L = \{(a, b) \in K \times L \ : \ f(a) = g(b)\}$ is naturally a smooth manifold equipped with commuting maps*

$$\begin{tikzcd}
K \times L \arrow[drr, "p_2"] \arrow[ddr, "p_1"'] & & \\
& K\,{}_f{\times}_g L \arrow[ul, "i"'] \arrow[r] \arrow[d] \arrow[dr, "f\cap g"] & L \arrow[d, "g"] \\
& K \arrow[r, "f"'] & M
\end{tikzcd} \tag{52}$$

where i is the inclusion and $f \cap g : (a, b) \mapsto f(a) = g(b)$.

The manifold $K\,{}_f{\times}_g L$ of the previous proposition is called the *fiber product* of K with L over M, and is a generalization of the intersection product. It is often denoted simply by $K \times_M L$, when the maps to M are clear.

Proof. Consider the graphs $\Gamma_f \subset K \times M$ and $\Gamma_g \subset L \times M$. To impose $f(k) = g(l)$, we can take an intersection with the diagonal submanifold

$$\Delta = \{(k, m, l, m) \in K \times M \times L \times M\}. \tag{53}$$

Step 1. We show that the intersection $\Gamma = (\Gamma_f \times \Gamma_g) \cap \Delta$ is transverse. Let $f(k) = g(l) = m$ so that $x = (k, m, l, m) \in \Gamma$, and note that

$$T_x(\Gamma_f \times \Gamma_g) = \{((v, Df(v)), (w, Dg(w))),\ v \in T_kK,\ \ w \in T_lL\} \tag{54}$$

whereas we also have

$$T_x(\Delta) = \{((v, m), (w, m)) \ : \ v \in T_kK,\ \ w \in T_lL,\ \ m \in T_pM\} \tag{55}$$

By transversality of f, g, any tangent vector $m_i \in T_pM$ may be written as $Df(v_i) + Dg(w_i)$ for some (v_i, w_i), $i = 1, 2$. In particular, we may decompose a general tangent vector to $M \times M$ as

$$(m_1, m_2) = (Df(v_2), Df(v_2)) + (Dg(w_1), Dg(w_1)) + (Df(v_1 - v_2), Dg(w_2 - w_1)), \tag{56}$$

leading directly to the transversality of the spaces (54), (55). This shows that Γ is a submanifold of $K \times M \times L \times M$.

Step 2. The projection map $\pi : K \times M \times L \times M \to K \times L$ takes Γ bijectively to $K{}_f{\times}_g L$. Since (54) is a graph, it follows that $\pi|_\Gamma : \Gamma \to K \times L$ is an injective immersion. Since the projection π is an open map, it also follows that $\pi|_\Gamma$ is a homeomorphism onto its image, hence is an embedding. This shows that $K{}_f{\times}_g L$ is a submanifold of $K \times L$. $\square$

Example 3.6. If $K_1 = M \times Z_1$ and $K_2 = M \times Z_2$, we may view both K_i as "fibering" over M with fibers Z_i. If p_i are the projections to M, then $K_1 \times_M K_2 = M \times Z_1 \times Z_2$, hence the name "fiber product".

Example 3.7. Let $L \subset M$ be a submanifold and let $f : K \to M$ be "transverse to L" in the sense that f is transverse to the embedding $\iota_L : L \to M$. This means that for each pair (k, l) such that $f(k) = l$, we have $Df(T_kK) + T_lL = T_lM$. Under this condition, the theorem implies that

$$f^{-1}(L) = \{k \in K \ : \ f(k) \in L\}$$

is a smooth submanifold of K (Why?) This is a generalization of the regular value theorem.

Example 3.8. Consider the Hopf map $p : S^3 \longrightarrow S^2$ given by composing the embedding $S^3 \subset \mathbb{C}^2 \backslash \{0\}$ with the projection $\pi : \mathbb{C}^2 \backslash \{0\} \longrightarrow \mathbb{C}P^1 \cong S^2$. Then for any point $q \in S^2$, $p^{-1}(q) \cong S^1$. Since p is a submersion, it is obviously transverse to itself, hence we may form the fiber product

$$S^3 \times_{S^2} S^3,$$

which is a smooth 4-manifold equipped with a map $p \cap p$ to S^2 with fibers $(p \cap p)^{-1}(q) \cong S^1 \times S^1$.

These are our first examples of nontrivial fiber bundles, which we shall explore later.

3.1 Stability

Transversality is a stable condition. In other words, if transversality holds, it will continue to hold for any small perturbation (of the submanifolds or maps involved). In a sense, stability says that transversal maps form an open set, and genericity says that this open set is dense in the space of maps. To make this precise, we would introduce a topology on the space of maps, something which we leave for another course.

Definition 3.9. We call a smooth map

$$F : M \times [0, 1] \to N \tag{57}$$

a smooth homotopy from f_0 to f_1, where $f_t = F \circ j_t$ and $j_t : M \to M \times [0, 1]$ is the embedding $x \mapsto (x, t)$.

Definition 3.10. A property of a smooth map $f : M \longrightarrow N$ is *stable* under perturbations when for any smooth homotopy f_t with $f_0 = f$, there exists an $\epsilon > 0$ such that the property holds for all f_t with $t < \epsilon$.

Proposition 3.11. *If M is compact, then the property of $f : M \to N$ being an immersion (or submersion) is stable under perturbations.*

Proof. If $f_t, t \in [0, 1]$ is a smooth homotopy of the immersion f_0, then in any chart around the point $p \in M$, the derivative $Df_0(p)$ has a $m \times m$ submatrix with nonvanishing determinant, for $m = \dim M$. By continuity, this $m \times m$ submatrix must have nonvanishing determinant in a neighbourhood around $(p, 0) \in M \times [0, 1]$. We can cover $M \times \{0\}$ by a finite number of such neighbourhoods, since M is compact. Choose ϵ such that $M \times [0, \epsilon)$ is contained in the union of these intervals, giving the result. The proof for submersions is identical. $\square$

Corollary 3.12. *If K is compact and $f : K \to M$ is transverse to the closed submanifold $L \subset M$ (this just means that f is transverse to the embedding $\iota : L \to M$), then the transversality is stable under perturbations of f.*

Proof. Let $F : K \times [0, 1] \to M$ be a homotopy with $f_0 = f$. We show that K has an open cover by neighbourhoods in which f_t is transverse for t in a small interval; we then use compactness to obtain a uniform interval.

First the points which do not intersect L: $F^{-1}(M \setminus L)$ is open in $K \times [0, 1]$ and contains $(K \setminus f^{-1}(L)) \times \{0\}$. So, for each $p \in K \setminus f^{-1}(L)$, there is a neighbourhood $U_p \subset K$ of p and an interval $I_p = [0, \epsilon_p)$ such that $F(U_p \times I_p) \cap L = \varnothing$.

Now, the points which do intersect L: L is a submanifold, so for each $p \in f^{-1}(L)$, we can find a neighbourhood $V \subset M$ containing $f(p)$ and a submersion $\psi : V \to \mathbb{R}^l$ cutting out $L \cap V$. Transversality of f and L is then the statement that ψf is a submersion at p. This implies there is a neighbourhood $\tilde{U}_p$ of $(p, 0)$ in $K \times [0, 1]$ where ψf_t is a submersion. Choose an open subset (containing $(p, 0)$) of the form $U_p \times I_p$, for $I_p = [0, \epsilon_p)$.

By compactness of K, choose a finite subcover of $\{U_p\}_{p \in K}$; the smallest ϵ_p in the resulting subcover gives the required interval in which f_t remains transverse to L. $\square$

Remark 3.13. Transversality of two maps $f : M \to N$, $g : M' \to N$ can be expressed in terms of the transversality of $f \times g : M \times M' \to N \times N$ to the diagonal $\Delta_N \subset N \times N$. So, if M and M' are compact, we get stability for transversality of f, g under perturbations of both f and g.

Remark 3.14. Local diffeomorphism and embedding are also stable properties.

3.2 Sard's theorem

Not only is transversality *stable*, it is actually *generic*, meaning that even if it does not hold, it can be made to hold by a small perturbation. The fundamental idea which allows us to prove that transversality is a generic condition is a the theorem of Sard showing that critical values of a smooth map $f : M \longrightarrow N$ (i.e. points $q \in N$ for which the map f and the inclusion $\iota : q \hookrightarrow N$ fail to be transverse maps) are *rare*. The following proof is taken from Milnor, based on Pontryagin.

The meaning of "rare" will be that the set of critical values is of *measure zero*, which means, in $\mathbb{R}^m$, that for any $\epsilon > 0$ we can find a sequence of balls in $\mathbb{R}^m$, containing $f(C)$ in their union, with total volume less than ϵ. Some easy facts about sets of measure zero: the countable union of measure zero sets is of measure zero, the complement of a set of measure zero is dense.

We begin with an elementary lemma describing the behaviour of measure-zero sets under differentiable maps.

Lemma 3.15. *Let $I^m = [0,1]^m$ be the unit cube, and $f : I^m \longrightarrow \mathbb{R}^n$ a C^1 map. If $m < n$ then $f(I^m)$ has measure zero. If $m = n$ and $A \subset I^m$ has measure zero, then $f(A)$ has measure zero.*

Proof. If $f \in C^1$, its derivative is bounded on I^m, so for all $x, y \in I^m$ we have
$$||f(y) - f(x)|| \leq K||y - x||, \tag{58}$$
for a constant[3] $K > 0$ depending only on f. So, the image of a ball of radius r in I^m is contained in a ball of radius Kr, which has volume proportional to r^n.

If $A \subset I^m$ has measure zero, then for each ϵ we have a countable covering of A by balls of radius r_k with total volume $c_m \sum_k r_k^m < \epsilon$. We deduce that $f(A_i)$ is covered by balls of radius Kr_k with total volume $K^n c_n \sum_k r_k^n$; since $n \geq m$ this goes to zero as $\epsilon \to 0$. We conclude that $f(A)$ is of measure zero.

If $m < n$ then f defines a C^1 map $I^m \times I^{n-m} \longrightarrow \mathbb{R}^n$ by pre-composing with the projection map to I^m. Since $I^m \times \{0\} \subset I^m \times I^{n-m}$ clearly has measure zero, its image must also. $\square$

Remark 3.16. If we considered the case $n < m$, the resulting sum of volumes may be larger in $\mathbb{R}^n$. For example, the projection map $\mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $(x, y) \mapsto x$ clearly takes the set of measure zero $y = 0$ to one of positive measure.

A subset $A \subset M$ of a manifold is said to have measure zero when its image in each chart of an atlas has measure zero. Lemma 3.15, together with the fact that a manifold is second countable, implies that the property is independent of the choice of atlas, and that it is preserved under equidimensional maps:

Corollary 3.17. *Let $f : M \to N$ be a C^1 map of manifolds where $\dim M = \dim N$. Then the image $f(A)$ of a set $A \subset M$ of measure zero also has measure zero.*

Corollary 3.18 (Baby Sard)**.** *Let $f : M \to N$ be a C^1 of manifolds where $\dim M < \dim N$. Then $f(M)$ (i.e. the set of critical values) has measure zero in N.*

Remark 3.19. Note that this implies that space-filling curves are not C^1.

[3]This is called a Lipschitz constant.

Now we investigate the measure of the critical values of a map $f : M \to N$ where $\dim M = \dim N$. The set of critical points need not have measure zero, but we shall see that

> **The variation of f is constrained along its critical locus since this is where Df drops rank. In fact, the set of critical *values* has measure zero.**

Theorem 3.20 (Equidimensional Sard)**.** *Let $f : M \to N$ be a C^1 map of n-manifolds, and let $C \subset M$ be the set of critical points. Then $f(C)$ has measure zero.*

Proof. It suffices to show result for the unit cube mapping to Euclidean space (using second countability, we can cover M by countable collection of charts $(U_i, \varphi_i)_{i\in I}$ with the property that $(\varphi_i^{-1}(I^n))_{i\in I}$ covers M. Since a countable union of measure zero sets is measure zero, we obtain the result). Let $f : I^n \longrightarrow \mathbb{R}^n$ a C^1 map, and let K be the Lipschitz constant for f on I^n, i.e.

$$||f(x) - f(y)|| \le K|x - y|, \quad \forall x, y \in I^n. \tag{59}$$

Let c be a critical point, so that the image of $Df(c)$ is a proper subspace of $\mathbb{R}^n$. Choose a hyperplane containing this subspace, translate it to $f(c)$, and call it H. Then

$$d(f(x), H) \le ||f(x) - f_c^{\mathrm{lin}}(x)||, \tag{60}$$

where $f_c^{\mathrm{lin}}(x) = f(c) + D_c f(x-c)$ is the linear approximation to f at c. By the definition of the derivative, for each $c \in C$, we have that $\forall \epsilon > 0, \exists \delta > 0$ such that

$$||f(x) - f_c^{\mathrm{lin}}(x)|| < \epsilon ||x - c|| \text{ for all } x \text{ s.t. } ||x - c|| < \delta.$$

Because f is C^1 and C is compact, we conclude that $\forall \epsilon > 0, \exists \delta > 0$ such that the inequality above holds for all $c \in C$.

Now we apply this: if $c \in C$ and $||x - c|| \le \delta$, then $f(x)$ is within a distance $\epsilon\delta$ from H and within a distance $K\delta$ of $f(c)$, so lies within a parallelopiped of volume

$$(2\epsilon\delta)(2K\delta)^{n-1}. \tag{61}$$

Now subdivide I^n into h^n cubes of edge length h^{-1} with h sufficiently large that $h^{-1}\sqrt{n} < \delta$. Apply the argument for each small cube, in which $||x - c|| \le h^{-1}\sqrt{n} < \delta$. The number of cubes containing critical points is at most h^n, so this gives a total volume for $f(C)$ less than

$$(2\epsilon h^{-1}\sqrt{n})(2Kh^{-1}\sqrt{n})^{n-1}(h^n) = 2^n K\epsilon n^{n/2}. \tag{62}$$

Since ϵ can be chosen arbitrarily small, $f(C)$ has measure zero. $\square$

The argument above will not work for $\dim N < \dim M$; we need more control on the function f. In particular, one can find a C^1 function $I^2 \longrightarrow \mathbb{R}$ which fails to have critical values of measure zero. (Hint: find a C^1 function $f : \mathbb{R} \to \mathbb{R}$ with critical values containing the Cantor set $C \subset [0, 1]$. Compose $f \times f$ with the sum $\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and note that $C+C = [0, 2]$.) As a result, Sard's theorem in general requires more differentiability of f.

Theorem 3.21 (Big Sard's theorem)**.** *Let $f : M \longrightarrow N$ be a C^k map of manifolds of dimension m, n, respectively. Let C be the set of critical points. Then $f(C)$ has measure zero if $k > \frac{m}{n} - 1$.*

Proof. As before, it suffices to show for $f : I^m \longrightarrow \mathbb{R}^n$. We do an induction on m – note that the theorem holds for $m = 0$.

Define $C_1 \subset C$ to be the set of points x for which $Df(x) = 0$. Define $C_i \subset C_{i-1}$ to be the set of points x for which $D^j f(x) = 0$ for all $j \leq i$. So we have a descending sequence of closed sets:

$$C \supset C_1 \supset C_2 \supset \cdots \supset C_k. \tag{63}$$

We will show that $f(C)$ has measure zero by showing

1. $f(C_k)$ has measure zero,
2. each successive difference $f(C_i \backslash C_{i+1})$ has measure zero for $i \geq 1$,
3. $f(C \backslash C_1)$ has measure zero.

Step 1: For $x \in C_k$, Taylor's theorem gives the estimate

$$||f(x + t) - f(x)|| \leq c||t||^{k+1}, \tag{64}$$

where c depends only on I^m and f.

Subdivide I^m into h^m small cubes with edge h^{-1}; then any point in in the small cube I_0 containing x may be written as $x + t$ with $||t|| \leq h^{-1}\sqrt{m}$. As a result, $f(I_0)$ is contained by a cube of edge $ah^{-(k+1)}$, with $a = 2cm^{(k+1)/2}$ independent of the small cube size. At most h^m cubes are necessary to cover C_k, and their images have total volume less than

$$h^m (ah^{-(k+1)})^n = a^n h^{m-(k+1)n}. \tag{65}$$

Assuming that $k > \frac{m}{n} - 1$, this tends to 0 as we increase the number of cubes.

Step 2: For each $x \in C_i \backslash C_{i+1}$, $i \geq 1$, there is a $i + 1^{th}$ partial, say wlog $\partial^{i+1} f_1 / \partial x_1 \cdots \partial x_{i+1}$, which is nonzero at x. Therefore the function

$$w(x) = \partial^i f_1 / \partial x_2 \cdots \partial x_{i+1} \tag{66}$$

vanishes on C_i but its partial derivative $\partial w / \partial x_1$ is nonvanishing near x. Then

$$(w(x), x_2, \ldots, x_m) \tag{67}$$

forms an alternate coordinate system in a neighbourhood V around x by the inverse function theorem (the change of coordinates is of class C^k), and we have trapped C_i inside a hyperplane. The restriction of f to $w = 0$ in V is clearly critical on $C_i \cap V$ and so by induction on m we have that $f(C_i \cap V)$ has measure zero. Cover $C_i \setminus C_{i+1}$ by countably many such neighbourhoods V.

Step 3: Let $x \in C \backslash C_1$. Note that we won't necessarily be able to trap C in a hypersurface. But, since there is some partial derivative, wlog $\partial f_1 / \partial x_1$, which is nonzero at x, so defining $w = f_1$, we have that

$$(w(x), x_2, \ldots, x_m) \tag{68}$$

is an alternative coordinate system in some neighbourhood V of x (the coordinate change is a diffeomorphism of class C^k). In these coordinates, the hyperplanes $w = t$ in the domain are sent into hyperplanes $y_1 = t$ in the codomain, and so f can be described as a family of maps f_t whose domain and codomain has dimension reduced by 1. Since $w = f_1$, the derivative of f in these coordinates can be written

$$Df = \begin{pmatrix} 1 & 0 \\ * & Df_t \end{pmatrix}, \tag{69}$$

and so a point $x' = (t, p)$ in V is critical for f if and only if p is critical for f_t. Therefore, the critical values of f consist of the union of the critical values of f_t on each hyperplane $y_1 = t$ in the codomain. Since the domain of f_t has dimension reduced by one, by induction it has critical values of measure zero. So the critical values of f intersect each hyperplane in a set of measure zero, and by Fubini's theorem this means they have measure zero. Cover $C \setminus C_1$ by countably many such neighbourhoods. $\square$

Remark 3.22. Note that $f(C)$ is measurable, since it is the countable union of compact subsets (the set of critical values is not necessarily closed, but the set of critical points is closed and hence a countable union of compact subsets, which implies the same of the critical values.)

To show the consequence of Fubini's theorem directly, we can use the following argument. First note that for any covering of $[a, b]$ by intervals, we may extract a finite subcovering of intervals whose total length is $\leq 2|b - a|$. To see this, first choose a minimal subcovering $\{I_1, \ldots, I_p\}$, numbered according to their left endpoints. Then the total overlap is at most the length of $[a, b]$. Therefore the total length is at most $2|b - a|$.

Now let $B \subset \mathbb{R}^n$ be compact, so that we may assume $B \subset \mathbb{R}^{n-1} \times [a, b]$. We prove that if $B \cap P_c$ has measure zero in the hyperplane $P_c = \{x^n = c\}$, for any constant $c \in [a, b]$, then it has measure zero in $\mathbb{R}^n$.

If $B \cap P_c$ has measure zero, we can find a covering by open sets $R_c^i \subset P_c$ with total volume $< \epsilon$. For sufficiently small α_c, the sets $R_c^i \times [c - \alpha_c, c + \alpha_c]$ cover $B \cap \bigcup_{z \in [c - \alpha_c, c + \alpha_c]} P_z$ (since B is compact). As we vary c, the sets $[c - \alpha_c, c + \alpha_c]$ form a covering of $[a, b]$, and we extract a finite subcover $\{I_j\}$ of total length $\leq 2|b - a|$.

Let R_j^i be the set R_c^i for $I_j = [c - \alpha_c, c + \alpha_c]$. Then the sets $R_j^i \times I_j$ form a cover of B with total volume $\leq 2\epsilon|b - a|$. We can make this arbitrarily small, so that B has measure zero.

3.3 Brouwer's fixed point theorem

Corollary 3.23. *Let M be a compact manifold with boundary. There is no smooth map $f : M \longrightarrow \partial M$ leaving ∂M pointwise fixed. Such a map is called a smooth retraction of M onto its boundary.*

Proof. Such a map f must have a regular value by Sard's theorem, let this value be $y \in \partial M$. Then y is obviously a regular value for $f|_{\partial M} = \mathrm{Id}$ as well, so that $f^{-1}(y)$ must be a compact 1-manifold with boundary given by $f^{-1}(y) \cap \partial M$, which is simply the point y itself. Since there is no compact 1-manifold with a single boundary point, we have a contradiction. $\square$

For example, this shows that the identity map $S^n \to S^n$ may not be extended to a smooth map $f : \overline{B(0,1)} \to S^n$.

Lemma 3.24. *Every smooth map of the closed n-ball to itself has a fixed point.*

Proof. Let $D^n = \overline{B(0,1)}$. If $g : D^n \to D^n$ had no fixed points, then define the function $f : D^n \to S^{n-1}$ as follows: let $f(x)$ be the point in S^{n-1} nearer to x on the line joining x and $g(x)$.

This map is smooth, since $f(x) = x + tu$, where

$$u = ||x - g(x)||^{-1}(x - g(x)), \tag{70}$$

and t is the positive solution to the quadratic equation $(x+tu)\cdot(x+tu) = 1$, which has positive discriminant $b^2 - 4ac = 4(1 - |x|^2 + (x \cdot u)^2)$. Such a smooth map is therefore impossible by the previous corollary. $\square$

Theorem 3.25 (Brouwer fixed point theorem)**.** *Any continuous self-map of D^n has a fixed point.*

Proof. The Weierstrass approximation theorem says that any continuous function on $[0,1]$ can be uniformly approximated by a polynomial function in the supremum norm $||f||_\infty = \sup_{x\in[0,1]} |f(x)|$. In other words, the polynomials are dense in the continuous functions with respect to the supremum norm. The Stone-Weierstrass is a generalization, stating that for any compact Hausdorff space X, if A is a subalgebra of $C^0(X, \mathbb{R})$ such that A separates points $(\forall x, y, \exists f \in A \ : f(x) \neq f(y))$ and contains a nonzero constant function, then A is dense in C^0.

Given this result, approximate a given continuous self-map g of D^n by a polynomial function p' so that $||p' - g||_\infty < \epsilon$ on D^n. To ensure p' sends D^n into itself, rescale it via

$$p = (1 + \epsilon)^{-1}p'. \tag{71}$$

Then clearly p is a D^n self-map while $||p - g||_\infty < 2\epsilon$. If g had no fixed point, then $|g(x) - x|$ must have a minimum value μ on D^n, and by choosing $2\epsilon = \mu$ we guarantee that for each x,

$$|p(x) - x| \geq |g(x) - x| - |g(x) - p(x)| > \mu - \mu = 0. \tag{72}$$

Hence p has no fixed point. Such a smooth function can't exist and hence we obtain the result. $\square$