Fixed Point Theorems of Banach and Brouwer

Daniel Son

Abstract

Given a topological space X and a mapping $f : X \to X$, a fixed point is defined to be a point such that $f(x) = x$. Fixed points are likely to exist for most functions. In this paper, we present two such conditions on f that guarantees a fixed point, using the concept of contraction and retraction.

1 Banach's Fixed Point Theorem

1.1 Preliminaries: Metric Spaces and Banach Spaces

In order to understand the fixed point theorems, we present some preliminaries.

Definition 1.1 (Metric Spaces)**.** *A space (X, d) is called a metric space if X is a set, and $d : X \times X \to \mathbb{R}_{\geq 0}$ is a function that satisfies the following properties. For all $x, y, z \in X$,*

- $d(x, y) = 0$ *if and only if* $x = y$.
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality)

Also, we call a metric space to be complete if all Cauchy Sequences in the spaces necessarily converge.

The metric $d(x, y)$ endows a distance in an abstract space X which is a set. Examples of metric spaces include $\mathbb{R}, \mathbb{R}^n$ equipped with the Eucledian metric.

In applications to physics, an important complete metric space is a Banach Space. We define a norm of a vector and the Banach Space below.

Definition 1.2 (Norm and Normed Space)**.** *Let V be a vector space, and $\| \cdot \| : V \to \mathbb{R}_{\geq 0}$ a function that maps a vector to a nonnegative real number. The function is a norm if the following conditions hold.*

- $\|x\| = 0$ *if and only if* $x = 0$ (Positive Definite)
- $\|\alpha x\| = |\alpha|\|x\|$ (Homogeneity)
- $\|x\| + \|y\| \geq \|x + y\|$ (Triangle Inequality)

If V is endowed with a metric $d : V \times V \to \mathbb{R}_{\geq 0}$ defined as

$$d(x, y) = \|x - y\| \tag{1.1}$$

then V is called a normed vector space.

Definition 1.3 (Banach Space)**.** *A normed vector space V is called a Banach Space if $(V, \|\cdot\|)$ is a complete metric space.*

In fact, it is possible to show normed spaces are necessarily metric spaces. We present a brief proof.

Proposition 1.1. *Normed spaces are metric spaces.*

Proof. We verify that the metric defined by the norm indeed satisfies the three conditions of the metric space. $d(x, y) = 0$ if and only if $\|x - y\| = 0$ if and only if $x - y = 0$ or $x = y$. Also,

$$d(x, y) \;=\; \|x - y\| \;=\; \|y - x\||-1| \;=\; d(y, x) \tag{1.2}$$

where the second equality is achieved by homogeneity. Finally, the triangle equality can be established as follows.

$$d(x, y) + d(y, z) \;\geq \|x - y\| + \|y - z\| \geq \|x - z\| = d(x, z) \tag{1.3}$$

□

1.2 Proof of Banach's FPT

Definition 1.4 (Contraction)**.** *Let (X, d) be a complete metric space, and $T : M \subset X \to X$ be a mapping from a closed subset M to M. T is called a contraction if there exists a real number $k \in (0, 1)$ that satisfies the following.*

$$d(Tx, Ty) \;\leq\; kd(x, y) \quad \forall x, y \in M \tag{1.4}$$

The Banach's fixed point theorem[1] guarantees a unique existance of a unique fixed point of T and convergence under consecutive iterations of T, given that T is a contraction.

Theorem 1.1 (Contraction Mapping Theorem)**.** *Let (X, d) be a complete metric space and $T : M \subset X \to M$ be a contraction where M is a closed subset. Then,*

- *the equation $Tx = x$ has a unique solution, i.e. there is a unique fixed point;*
- *the sequence (x_n) defined iteratively as $x_{n+1} = Tx_n$ converges for every $x_0 \in M$.*

Proof. We show that the iterative sequence is Cauchy. The completeness of the space X will guarantee that the sequence converges. Moreover, since the subset $M \subset X$ is closed, it must be the case that the converging point, namely $\bar{x}$, must be within M. Arguing by a way of contradiction, we show that $\bar{x}$ is a unique fixed point.

[1]Contraction Mapping Theorem , or CMT in short, is a different name for Banach's fixed point theorem

Choose any $\epsilon > 0$. For two entries x_n and x_{n+m}, we show that large enough n can always guarantee that the distance between the two points can be less than ϵ. By induction, it is straightforward to verify

$$d(x_{n+1}, x_n) \;\leq\; k^n d(x_1, x_0) \tag{1.5}$$

and by invoking the triangle inequality, we obtain

$$\begin{aligned} d(x_{n+m}, x_n) \leq& \left(\sum_{j=0}^{m-1} k^j\right) k^n d(x_1, x_0) && (1.6)\\ =& \frac{1-k^m}{1-k} k^n d(x_1, x_0) && (1.7)\\ \leq& \frac{d(x_1, x_0)}{1-k} k^n \leq \epsilon && (1.8) \end{aligned}$$

The last inequality is attained by an arbitrarily large choice of n, since $k \in (0,1)$.

Let $\bar{x}$ be the point of convergence, and assume for a contradiction that this point is not a fixed point, i.e. $Tx \neq x$. Since X is a metric space, $d(x, Tx) > 0$ necessarily. However, by the triangle inequality, we observe the following.

$$\begin{aligned} d(x_n, \bar{x}) + d(x_n, x_{n+1}) + d(x_{n+1}, T\bar{x}) \;&\geq d(\bar{x}, T\bar{x})\\ d(x_n, \bar{x}) + d(x_n, x_{n+1}) + d(Tx_n, T\bar{x}) \;&\geq d(\bar{x}, T\bar{x})\\ (k+1)d(x_n, \bar{x}) + d(x_n, x_{n+1}) \;&\geq\; d(\bar{x}, T\bar{x}) \qquad (1.9) \end{aligned}$$

The LHS of the final inequality converges to zero by the Cauchyness and convergence of sequence $(x)_n$. However, the RHS is a fixed positive real number, which is a contradiction.

It remains to deonstrate uniqueness. Suppose, on the contrary, that there exists two unique fixed points $a, b \in M$, such that

$$Ta \;=\; a \quad \text{and} \quad Tb \;=\; b. \tag{1.10}$$

Since T is a contraction, $d(Ta, Tb)$ is strictly less than $d(a, b)$. However, the two quantities describe distance between the same pair of points, and must be equal, a contradiction. $\square$

2 Brouwer's Fixed Point Theorem

2.1 Preliminaries: Lagrangians and Mollifiers

We will use the concept of Lagrangians to prove Brouwer's version of the FPT. In order to proceed, we need to understand the method of by-parts integration in higher dimensions.

Theorem 2.1 (By-parts in higher dimensions). *Suppose $\Omega \subset \mathbb{R}^n$ is a smooth domain, i.e. the boundary is continuously differentiable. Define a test function*

$\phi : \overline{\Omega} \to \mathbb{R}^m$ that vanishes around the boundary, i.e. $\forall x \in \partial\Omega$, $\phi(x) = 0$. For any set of smooth functions $\{F_{ij} : \Omega \to \mathbb{R} | 1 \le i \le m, 1 \le j \le n\}$,

$$\int_\Omega \sum_{j=1}^n \sum_{i=1}^m F_{ij} \partial_i \phi_i dx = -\int_\Omega \sum_{j=1}^n \sum_{i=1}^m \partial_j \left(F_{ij}\right) \phi_i dx \tag{2.1}$$

Proof. The regular by-parts formula is derived by the boundary conditions in $\mathbb{R}$. The divergence theorem simplifies a volume integral into a surface integral over the boundary. Define a vector field $G : \Omega \to \mathbb{R}^n$ as follows.

$$G(x) := \sum_{i=1}^m F_{ij}(x)\phi_i(x) \tag{2.2}$$

The divergence of $G(x)$ over the domain Ω must equal to the surface integral.

$$\int_\Omega \nabla \cdot G(x) dx = \int_{\partial\Omega} G(x) \cdot \nu dx \tag{2.3}$$

Here, ν is the normal vector at the point x on the boundary. Since the test function vanishes at the boundary, the entirety of $G(x)$ vanishes at the boundary. Therefore

$$\int_\Omega \nabla \cdot G(x) dx = 0. \tag{2.4}$$

Compute the jth partial of $G(x)$.

$$\partial_j G(x) = \sum_{i=1}^m \partial_j \left(F_{ij}(x)\right) \phi_i(x) + \sum_{i=1}^m F_{ij}(x) \partial_j \phi_i(x) \tag{2.5}$$

Add up the jth partials and to obtain the vanishing divergence integral. The integrals can be split into two parts, which yields the following.

$$\int_\Omega \sum_{j=1}^n \sum_{i=1}^m F_{ij}(x) \partial_j \phi_i(x) dx = -\int_\Omega \sum_{j=1}^n \sum_{i=1}^m \partial_j (F_{ij}(x)) \phi_i(x) dx \tag{2.6}$$

□

Definition 2.1 (Energy Functional)**.** *Let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be a smooth input function defined over a smooth domain Ω. The Larangian is defined as a functional that takes $x \in \Omega$, the function $u(x)$, and the Jacobian matrix $\nabla u(x)$, i.e. $L : \mathbb{R}^n \times \mathbb{R}^{m \times n} \times \Omega$. We define the energy functional as the following integral.*

$$I(u) := \int_\Omega L(\nabla u(x), u(x), x) dx, \tag{2.7}$$

For simplicity of notation, denote z_j to be the jth entry of $x \in \Omega$ and p_{ij} as the ith row of the jth colomn of the Jacobian matrix $\nabla u(x)$

Theorem 2.2 (Euler-Lagrange Equation)**.** *Let $u : \Omega \to \mathbb{R}^m$ be a minimizer or a maximizer of the energy functional with a fixed boundary condition. That is, $u(x)$ agrees with the boundary function $g : \partial\Omega \to \mathbb{R}^m$ over the domain $\partial\Omega$. Then, $u(x)$ must satisfy the following equations for all $i \in [m]$.*

$$L_{z_i}(\nabla u, u, x) \;=\; \sum_{j=1}^{n} \partial_j \left(L_{p_{i,j}}(\nabla u, u, x) \right) \tag{2.8}$$

Proof. Take a test function $\phi : \overline{\Omega} \to \mathbb{R}^m$ that will simulate the perturbations. The test function must satisfy

1. $\phi(x) = 0\ \forall x \in \partial\Omega$, the test function vanishes at the boundary;
2. $\mathrm{supp}(\phi) \subset \Omega \setminus \partial\Omega$, the test function has a support strictly included in the interior of Ω;
3. $\phi \in C^\infty(\overline{\Omega})^m$, the test function is smooth.

Perturb the energy functional by $\epsilon > 0$.

$$\mathcal{I}(\epsilon) \;:=\; I(u + \epsilon\phi) \tag{2.9}$$

If the function u indeed maximize energy, then the derivative of energy with respect to ϵ must be zero, regardless of the choice of the test function ϕ. Therefore $\frac{d}{d\epsilon}\mathcal{I}(\epsilon)\Big|_{\epsilon=0} = 0$, and we compute this derivative as a sum of partial derivatives.

$$\begin{aligned}\frac{d}{d\epsilon}\mathcal{I}(\epsilon)\Big|_{\epsilon=0} &= \int_\Omega \sum_{j=1}^{n}\sum_{i=1}^{m} L_{p_{ij}}(\nabla u(x) + \epsilon\nabla\phi(x), u(x) + \epsilon\phi(x), x)\partial_j\phi_i dx\Big|_{\epsilon=0} \\ &\quad + \int_\Omega \sum_{i=1}^{m} L_{z_j}(\nabla u(x) + \epsilon\nabla\phi(x), u(x) + \epsilon\phi(x), x)\phi_i dx\Big|_{\epsilon=0} \\ &= \int_\Omega \sum_{j=1}^{n}\sum_{i=1}^{m} L_{p_{ij}}(\nabla u(x), u(x), x)\partial_j\phi_i dx + \int_\Omega \sum_{i=1}^{m} L_{z_i}(\nabla u(x), u(x), x)\phi_i dx\end{aligned} \tag{2.10}$$

Invoke the by-parts formula by setting $F_{ij} = L_{p_{ij}}(\nabla u(x), u(x), x)$.

$$\int_\Omega \sum_{j=1}^{n}\sum_{i=1}^{m} L_{p_{i,j}}(\nabla u(x), u(x), x)\partial_j\phi_i dx \;=\; -\int_\Omega \sum_{j=1}^{n}\sum_{i=1}^{m} \partial_j \left[L_{p_{i,j}}(\nabla u(x), u(x), x) \right] \phi_i dx \tag{2.11}$$

Equate the expression obtained from (2.10) to zero. Interchange the order of the finite double sum.

$$\int_\Omega \sum_{i=1}^{m}\sum_{j=1}^{n} \partial_j \left[L_{p_{i,j}}(\nabla u(x), u(x), x) \right] \phi_i dx \;=\; \int_\Omega \sum_{i=1}^{m} L_{z_i}(\nabla u(x), u(x), x)\phi_i dx \tag{2.12}$$

The equality holds for any test function ϕ. Therefore, it must be the case that the coefficints of each ϕ_i in the integrand must agree for all $i \in m$. Thus,

$$\sum_{j=1}^{n} \partial_j \left[L_{p_{i,j}}(\nabla u(x), u(x), x) \right] \;=\; L_{z_i}(\nabla u(x), u(x), x) \tag{2.13}$$

□

In practice, x is usually set to be time. With this analogy in mind, the result reads that the partial of the Lagrangian with respect to a generalized coordinate equals to the time-divergence of the gradient of the Lagrangian over the velocities associated with the given coordinate. [2] Also, such functions $u(x)$ that satisfy the Euler Lagrange equations are called **stationary points**.

Definition 2.2 (Null Lagrangians)**.** *If every smooth function defined over the domain Ω is a stationary point, then we call L to be a Null Lagrangian.*

Theorem 2.3 (Energies of a Null Lagrangian are determined by the boundary values)**.** *Let $u, v : \overline{\Omega} \to \mathbb{R}^m$ be two stationary points that agree on the boundary, i.e. $u(x) = v(x) \forall x \in \partial\Omega$. Then, $I(u) = I(v)$.*

Proof. The crux of the argument is to consider the function $u(x) - v(x)$ like a test function, and invoke the by-parts formula as shown in the proof of the Euler Lagrange Equations.

Suppose $u(x)$ and $v(x)$ agree on the boundary of Ω. Define

$$\mathcal{I}(\tau) \;=\; I(\tau u(x) + (1-\tau)v(x)). \tag{2.14}$$

Take the derivative of $\mathcal{I}$ with respect to τ and invoke the by-parts formula for higher dimensions.

$$\mathcal{I}'(\tau) \;=\; \begin{array}{l} \displaystyle\int_\Omega \sum_{j=1}^{n}\sum_{i=1}^{m} L_{p_{i,j}}(\tau\nabla u + (1-\tau)\nabla v, \tau u + (1-\tau)v, x)(\partial_j u_i - \partial_j v_i)dx \\ \displaystyle\qquad + \int_\Omega \sum_{i=1}^{m} L_{z_i}(\tau\nabla u + (1-\tau)\nabla v, \tau u + (1-\tau)v, x)(u_i - v_i)dx \end{array} \tag{2.15}$$

$$= \int_\Omega \sum_{i=1}^{m} \left(-\sum_{j=1}^{n} \partial_j L_{p_{i,j}} + L_{z_i} \right) dx \;=\; 0 \tag{2.16}$$

The last equality follows since $\tau u(x) + (1-\tau)v(x)$ satisfies the boundary conditions, and L is a null Lagrangian, which implies that the Euler-Lagrange equation holds for any function including $\tau u(x) + (1-\tau)v(x)$.

Consequently, $\mathcal{I}(0) = \mathcal{I}(1)$ or $I(u) = I(v)$ as desired. □

[2] Of course, in reality, time is one dimensional. The analogy of x as time serves as a nice mneonic, assuming that x is a multi-dimensional time quantity.

Without proof, we present a result that will be used to prove Brouwer's FPT. The proof of this theorem is by linear algebra.

Proposition 2.1 (Determinants are Null Lagrangians)**.** *Suppose the variation* u *is defined as* $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ *such that the jacobian matrix* ∇u *is a square matrix. Then, the Lagrangian*

$$L(\nabla(u), u, x) \ := \ \det(\nabla u) \tag{2.17}$$

is a Null-Lagrangian.

Moreover, we present a special function in order to approximate continuous functions to a smooth function.

Definition 2.3 (Mollifiers)**.** *Let* $\phi : B_1(0) \subseteq \mathbb{R}^n \to \mathbb{R}$ *be a function defined as*

$$\phi(x) \ := \ C \exp\left(\frac{1}{|x|^2 - 1}\right) \tag{2.18}$$

. The function ϕ *is called a standard mollifier, and the constant* C *is adjusted such that the function integrates to 1 over the n-ball of radius 1. i.e.*

$$\int_{B_1(0)} \phi(x)dx \ = \ 1 \tag{2.19}$$

Also, ϕ_ϵ *is defined as*

$$\phi_\epsilon(x) \ = \ \frac{1}{\epsilon^n} f\left(\frac{1}{\epsilon}\right). \tag{2.20}$$

The Mollifier can be considered as a concentrated point mass at the origin in $\mathbb{R}^n$. The one dimensional mollifier is the delta function. Just like the delta function, the mollifier satisfies the shifting property.

Proposition 2.2 (Shifting property of the mollifier)**.** *Define the convolution of a mollifier with some function* $f : \mathbb{R}^n \to \mathbb{R}$ *as*

$$f_\epsilon(x) \ := \ [f * \phi_\epsilon](x) \ := \ \int_{\mathbb{R}^n} f(y)\phi_\epsilon(x-y)dy \tag{2.21}$$

If f *is uniformly continuous, then as* $\epsilon \to 0$*,* $f_\epsilon(x)$ *converges uniformly to* $f(x)$*.*

Proof. We wish to bound the following quantity for arbitrary $x \in \mathbb{R}^n$.

$$|f_\epsilon(x) - f(x)| \ = \ \left| \int_{\mathbb{R}^n} f(y)\phi_\epsilon(x-y)dy - f(x)\int_{\mathbb{R}^n} \phi_\epsilon(y)dy \right| \tag{2.22}$$

Relabel the arguments of the first integral, and absorb $f(x)$ under the integral sign for the second integral.

$$= \ \int_{\mathbb{R}^n} f(x-y)\phi_\epsilon(y)dy - \int_{\mathbb{R}^n} f(x)\phi_\epsilon(y)dy \ = \ \int_{\mathbb{R}^n} |f(x-y) - f(x)|\phi_\epsilon(y)dy \tag{2.23}$$

The support of the mollifier is an epsilon ball centerd at the origin. Invoking uniform continuity of f, we can claim that for all $y \in \mathbb{R}^n$ such that $|y| < \epsilon$, given any $\delta > 0$,

$$|f(x-y) - f(x)| \ < \ \delta. \tag{2.24}$$

Use this to bound the last expression in (2.23).

$$|f_\epsilon(x) - f(x)| \ = \ \max_{y \in B_\epsilon(0)} |f(x-y) - f(x)| \int_{B_\epsilon(0)} \phi_\epsilon(y) dy \ < \ \delta \tag{2.25}$$

We conclude that $f_\epsilon(x) \to f(x)$ uniformly. □

2.2 Proof of Brouwer's FPT

Finally, we have collected all the ingredients to prove Brouwer's Fixed Point Theorem. We first define a retraction, and prove the retraction principle which states that there cannot exist a retraction from a closed ball to its boundary. In the last lecture, we showed that the retraction principle implies Brouwer's FPT. We prove the retraction principle and leave the proof of Brouwer's FPT as an exercise. [3]

Definition 2.4 (Retraction)**.** *Consider two topological spaces A, B where $A \subseteq B$ and the topology of A is the subset topology induced by B. A continuous function $f : B \to A$ is called a retraction if $f|_A = Id_A$, i.e. for every $x \in A$, $f(x) = x$.*

For example, given spaces $R_{discrete}$ and $Z_{discrete}$, the floor function $f(x) = \lfloor x \rfloor$ is a retraction. Any function between the two spaces are continuous which demonstrates that f is continuous, and the floor function is preserved at the integer points.

Lemma 2.1 (Retraction Principle)**.** *Let $B := \overline{B}_1(0) \in \mathbb{R}^n$ be a closed ball in the euclidian space $\mathbb{R}^n$. The closed ball does not admit a retraction to its boundary, i.e. a continuous function $f : B \to \partial B$ such that $f(x) = x$ at $x \in \partial B$ does not exist.*

Proof. Assume such a function f exists. First, we decuce a contradiction under the condition that f is smooth. Secondly, we approximate an arbitrary continuous function to a smooth function using the mollifier. Invoking our result from the first part proves the theorem.

Part I. Let f be smooth. We evaluate the energy functional over the determinant Lagrangian. Since $f : B \to \mathbb{R}^n$ and $Id : B \to \mathbb{R}^n$ agrees at the boundary, the energy functional must agree.

$$\int_B \det(\nabla f) dx \ = \ \int_B \det(\mathrm{Id}_n) dx \ = \ \int_B dx \ > \ 0 \tag{2.26}$$

[3] Consider the ray from x to $f(x)$ and take the intersection with the boundary to obtain a continuous function.

Since f is a retraction to the boundary, the image of any $x \in B$ must have a norm of x. i.e.

$$f(x) \cdot f(x) \;=\; 1. \tag{2.27}$$

Apply a total derivative both sides. By the chain rule,

$$2Df(x) \cdot f(x) \;=\; 0 \quad \text{or} \quad \nabla f(x)^\intercal f(x) \;=\; 0 \tag{2.28}$$

[4] This implies that the Jacobian matrix ∇f has a zero eigenvalue, hence a vanishing determinant. Therefore, the integral in (**??**) must be zero, a contracition.

Part II. Let f be continuous but not necessarily smooth. Extend $f(x)$ to be defined over the entirety of $\mathbb{R}^n$ by letting $f(x) = x$ for $x \in \mathbb{R}^n \setminus B$. Note that the extended function is uniformly continuous, since the identity is uniformly continuous and continuous functions defined over compact intervals are continuous. Define a new function

$$\tilde{f}(x) \;=\; \frac{[f * \phi_\epsilon](2x)}{|[f * \phi_\epsilon](2x)|} \tag{2.29}$$

which for small enough ϵ is a smooth retraction to the boundary. By part 1, such functions cannot exist

□

Theorem 2.4 (Brouwer's Fixed Point Theorem)**.** *Denote a closed ball in $\mathbb{R}^n$ as $B := B_1(0)$. Suppose $f : B \to B$ is a continuous mapping. There must exist a fixed point $x \in B$ such that $f(x) = x$.*

Remark 2.1. *In fact Brouwer's FPT implies the retraction principle, and therefore the two statements are equivalent. To prove this, assume that there exists a retraction $r(x)$ and consider the fixed point of $(x - r(x))/2$. By considering the norm of the new function, the fixed point must lie in the boundary. However, at the boundary, the new function evaluates identically to zero, which is a contradiction.*

References

[1] Hunter, J. *Linear Systems and Eigenvalues.* Chapter 5. Retrieved from `https://www.math.ucdavis.edu/~hunter/book/ch5.pdf`

[2] Belk, J. *Banach Spaces: Measure Theory Notes.* Retrieved from `https://e.math.cornell.edu/people/belk/measuretheory/BanachSpaces.pdf`

[3] Department of Mathematics, University of Oxford. *Lecture Notes 2022.* Retrieved from `https://courses.maths.ox.ac.uk/pluginfile.php/22783/mod_resource/content/1/lecture-notes-2022.pdf`

[4] Terek, Ivo. *Introduction to Lorentz Geometry: Curves and Surfaces.* Springer, 2021.

[4] The second expression can be verified by componentwise expansion.