Notes for Advanced Topics in Functional Analysis

Augusto Del Zotto
email: augusto.delzotto@uzh.ch

Spring semester 2026 - UZH

Contents

1 Differential calculus - 1 - 16/02/2026 **3**
1.1 Introduction . . . 3
1.2 Fréchet and Gateaux derivative . . . 3

2 Differential calculus - 2 - 17/02/2026 **7**
2.1 Higher order derivatives - Schwartz Theorem . . . 9

3 Differential calculus - 3 - 23/02/2026 **12**
3.1 Taylor formula in Banach spaces . . . 13
3.2 Polynomials and inverse Taylor formula . . . 15

4 Differential calculus - 4 - 24/02/2026 **17**
4.1 Contraction principle . . . 19
4.2 Inverse and implicit function theorems . . . 19

5 Differential calculus - 5 - 02/03/2026 **21**
5.1 Smooth dependence w.r.t. initial data . . . 24

6 Nemitski operators – Nis-Erik Bohne - 09/03/2026 **26**
6.1 General setting . . . 26
6.2 Continuity of Nemitski operators . . . 26
6.3 Differentiability of Nemitski operators . . . 27
6.4 Technical results . . . 31

7 Bifurcation Theory 1 - 10/03/2026 **32**
7.1 Bifurcation points . . . 32
7.2 Bifurcation from simple eigenvalues . . . 33

8 Bifurcation Theory 2 - 16/03/2026 **36**
8.1 Time-periodic solutions for second-order systems . . . 40

9 Water waves 1 - 17/03/2026 **44**
9.1 The water wave system . . . 44
9.2 Zakharov variables and Dirichlet–Neumann operator . . . 47

10 Water waves 2 - 23/03/2026 **50**
10.1 Travelling waves . . . 50
10.2 Existence of small periodic travelling waves . . . 54

11 Hopf bifurcation 1 - 24/03/2026 **58**
11.1 Abstract Hopf bifurcation . . . 58
11.2 Dynamical Hopf bifurcation . . . 61

12 Hopf bifurcation 2 - 30/03/2026 **63**
12.1 A Van der Pol-type example . . . 65

13 Brouwer degree 1 - 13/04/2026 **68**
13.1 Construction of a degree map . . . 71

14 The Brouwer mapping degree 2 - 14/04/2026 **74**
14.1 Generalization to continuos functions . . . 75
14.2 Step 1: admitting critical values. . . . 76

15 The Brouwer mapping degree 3 - 21/04/2026 **81**
15.1 Step 2: admitting continuous functions. . . . 81
15.2 Application: Bröwer fixed point . . . 83

16 The Brouwer mapping degree 4 - 27/04/2026 **87**

17 The Leray-Schauder degree 1 - 04/05/2026 **90**
17.1 Compact mappings . . . 91

18 The Leray-Schauder degree 2 - 05/05/2026 **96**
18.1 Definition of the degree map . . . 96

19 The Leray-Schauder degree 3 - 11/05/2026 **101**
19.1 Application: Schauder fixed point . . . 101
19.2 Application: integral and differential equations . . . 104

20 Stationary Navier-Stokes 1 - 18/05/2026 **107**
20.1 Introduction to Navier–Stokes . . . 107
20.2 Stationary solutions and weak formulation . . . 107
20.3 Plan of the argument . . . 108
20.4 Sobolev spaces toolbox . . . 109

21 Stationary Navier-Stokes 2 - 19/05/2026 **112**
21.1 Functional setting and the nonlinear operator . . . 112
21.2 Existence by the Leray-Schauder principle . . . 114
21.3 Uniqueness for small data . . . 115

Bibliography

Selection of topics from the following books:

- A. Ambrosetti, G. Prodi - *A Primer of Nonlinear Analysis*;
- G. Teschl - *Topics in Real and Functional Analysis*;
- K. Deimling - *Nonlinear Functional Analysis*.

1 Differential calculus - 1 - 16/02/2026

1.1 Introduction

Definition 1.1 (Normed vector space and Banach space)**.** Let X be a vector space over a field K ($\mathbb{R}$ or $\mathbb{C}$) and let $\|\cdot\|_X : X \to \mathbb{R}$ be a function on X such that

1. $\|x+y\|_X \leq \|x\|_X + \|y\|_X \qquad \forall x, y \in X,$
2. $\|\lambda x\|_X = |\lambda| \ \|x\|_X \qquad \forall \lambda \in K, \ \forall x \in X,$
3. $\|x\|_X = 0 \iff x = 0 \qquad \forall x \in X.$

Then $\|\cdot\|_X$ is a *norm* on X and $(X, \|\cdot\|_X)$ is a *normed vector space*. Moreover, if $\|\cdot\|_X$ is *complete*, i.e.

$$\forall \{x_n\} \subset X \text{ Cauchy sequence } \exists x \in X \text{ s.t. } \lim_{n\to\infty} \|x_n - x\|_X = 0,$$

then $(X, \|\cdot\|_X)$ is a *Banach space*.

Example 1.2. Examples.

- $(\mathbb{R}, |\cdot|)$ is a Banach space.
- $(\mathbb{R}, \|\cdot\|_p)$ is a Banach space.
- $(\mathbb{Q}, |\cdot|)$ is a normed vector space.
- $(C([0,1]), \|\cdot\|_\infty)$ is a Banach space.
- $(C([0,1]), \|\cdot\|_1)$ is a normed vector space.
- $(L^p([0,1]), \|\cdot\|_p)$ is a Banach space.
- $(\ell^p, \|\cdot\|_{\ell^p})$ is a Banach space (ℓ^p space of sequences).

Exercise 1.3. Show that $(C([0,1]), \|\cdot\|_1)$ is not a Banach space.

Exercise 1.4. Let

$$\mathbb{R}_0 := \Big\{ \{x_i\}_{i\in\mathbb{N}} : \ \exists N \in \mathbb{N} \text{ s.t. } x_i = 0 \ \forall i > N \Big\}, \qquad \|x\|_\infty := \sup_i |x_i|. \tag{1.1}$$

Show that $(\mathbb{R}_0, \|\cdot\|_\infty)$ is not a Banach space.

1.2 Fréchet and Gateaux derivative

Let X, Y be Banach spaces with norms $\|\cdot\|_X, \|\cdot\|_Y$ respectively, $U \subset X$ an open subset, and $F : U \to Y$ a functional.

Definition 1.5 (Continuous maps)**.** We write $F \in \mathcal{C}(X, Y)$ if F is a continuous map from X to Y, i.e.

$$F(x_n) \to F(x) \text{ in the } Y\text{-norm} \qquad \forall\, x_n \to x \text{ in the } X\text{-norm}.$$

Moreover, for an open subset $U \subseteq X$, we denote by $\mathcal{C}(U, Y)$ the set of continuous maps from U to Y.

Definition 1.6. Given X, Y Banach spaces, we define

$$\mathcal{L}(X, Y) := \{A : X \to Y : A \text{ linear and continuous}\} \tag{1.2}$$

as the space of linear and continuous maps between X and Y. For an element $A \in \mathcal{L}(X, Y)$ we also define the norm of A as

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|_Y}{\|x\|_X}. \tag{1.3}$$

Exercise 1.7. Assume Y is a Banach space. Show that $\mathcal{L}(X,Y)$ (endowed with $\|\cdot\|$) is a Banach space.

Remark. If $X = Y$, we denote $\mathcal{L}(X,Y) = \mathcal{L}(X,X) = \mathcal{L}(X)$.

We now generalize the usual definition of differentiability of maps in Euclidean spaces.

Definition 1.8 (Fréchet differential)**.** F is said to be *Fréchet differentiable* (F-differentiable, or simply differentiable) at $x \in U$ if and only if there exists $A \in \mathcal{L}(X,Y)$ such that

$$\lim_{\|h\|_X \to 0} \frac{\|F(x+h) - F(x) - A(h)\|_Y}{\|h\|_X} = 0. \tag{1.4}$$

Equivalently,

$$F(x+h) - F(x) = Ah + o(\|h\|_X) \qquad (\|h\|_X \to 0).$$

We define $A := dF(x)$ as the differential of F at x. Moreover, if F is F-differentiable for all $x \in U$, we say that F is F-differentiable in U.

We discuss now this notion of the differential, presenting some properties of the F-differential, together with some examples and exercises. We begin by showing that the differential is unique, analogously to the finite dimensional case.

Proposition 1.9. *The differential is unique.*

Proof. Assume by contradiction that there exist $A_1, A_2 \in \mathcal{L}(X,Y)$ such that

$$\lim_{\|h\|_X \to 0} \frac{\|F(x+h) - F(x) - A_i(h)\|_Y}{\|h\|_X} = 0, \qquad i = 1,2,$$

and $A_1 \neq A_2$. Then there exists $h_0 \in X$, $h_0 \neq 0$, such that $A_1 h_0 \neq A_2 h_0$. Let

$$a := \|A_1 h_0 - A_2 h_0\|_Y > 0, \qquad \text{and consider } h = t h_0,\ t \in \mathbb{R}.$$

Then

$$\frac{\|A_1 h - A_2 h\|_Y}{\|h\|_X} = \frac{\|A_1(th_0) - A_2(th_0)\|_Y}{\|th_0\|_X} = \frac{a}{\|h_0\|_X} \neq 0.$$

On the other hand,

$$\begin{aligned} 0 < \frac{\|A_1 h - A_2 h\|_Y}{\|h\|_X} &= \frac{\big\|\big(F(x+h) - F(x) - A_1 h\big) - \big(F(x+h) - F(x) - A_2 h\big)\big\|_Y}{\|h\|_X} \\ &\leq \frac{\|F(x+h) - F(x) - A_1 h\|_Y}{\|h\|_X} + \frac{\|F(x+h) - F(x) - A_2 h\|_Y}{\|h\|_X} \longrightarrow 0, \end{aligned}$$

as $t \to 0$, a contradiction. □

Exercise 1.10. Prove that if F is F-differentiable in x, then F is also continuous in x.

Example 1.11.

1. Let $F(x) = y_0 + Tx$, for $y_0 \in Y$ and $T \in \mathcal{L}(X,Y)$. Then $dF(x) = T$.
2. Let X be a Hilbert space and consider $F(x) = \|x\|_X^2$. Then F is F-differentiable and $dF(x) = 2\langle x, \cdot\rangle \in \mathcal{L}(X,\mathbb{R})$. Indeed,

$$\begin{aligned} F(x+h) - F(x) &= \langle x+h, x+h\rangle - \langle x, x\rangle \\ &= \langle x, x\rangle + 2\langle x, h\rangle + \langle h, h\rangle - \langle x, x\rangle \\ &= 2\langle x, h\rangle + o(\|h\|_X). \end{aligned}$$

3. Let $B : X \times Y \to Z$ be a bilinear form, with X, Y, Z Banach spaces. Then B is differentiable and

$$dB(x,y) = B(\cdot, y) + B(x, \cdot).$$

Exercise 1.12. Let E be a Banach space and define $X := \mathcal{L}(E,E) = \mathcal{L}(E)$ (with the operator norm $\|\cdot\|_X$).

1. Prove that the subset $U := \{T \in \mathcal{L}(E) : \ T \text{ invertible}\}$ is open;

2. Prove that the functional $F : U \to X$ defined as $F(T) = T^{-1}$ is F-differentiable in U and compute its differential.

From now on, we will refer to F-differentiability simply as differentiability. We state now some properties expected from the comparison with the finite dimensional case.

Proposition 1.13. *We have the following properties.*

1. *Let $F, G : U \to Y$ be differentiable at x_0. Then $F + G$ is differentiable at x_0 and*

$$d(F + G)(x_0) = dF(x_0) + dG(x_0).$$

2. *Let $F : U \to Y$ be differentiable at $x_0 \in U \subset X$ and $G : V \to Z$ differentiable at $F(x_0) \in V \subset Y$. Then $G \circ F$ is differentiable at x_0 and*

$$d(G \circ F)(x_0) = dG(F(x_0)) \circ dF(x_0).$$

3. *If $Y = \mathbb{R}$, then $F \cdot G$ is differentiable and*

$$d(F \cdot G)(x_0) = dF(x_0)\, G(x_0) + dG(x_0)\, F(x_0).$$

Proof. (*Proof only of (2).*) Define $H(x) := G \circ F(x)$. For h small,

$$F(x_0 + h) = F(x_0) + dF(x_0)[h] + o(\|h\|_X).$$

Since G is differentiable at $y_0 := F(x_0)$, we get

$$\begin{aligned} H(x_0 + h) &= G(F(x_0 + h)) \\ &= G\big(F(x_0) + dF(x_0)[h] + o(\|h\|_X)\big) \\ &= G(F(x_0)) + dG(F(x_0))\big[dF(x_0)[h] + o(\|h\|_X)\big] + o(\|h\|_X) \\ &= H(x_0) + \big(dG(F(x_0)) \circ dF(x_0)\big)[h] + o(\|h\|_X). \end{aligned}$$

Hence H is differentiable at x_0 with $dH(x_0) = dG(F(x_0)) \circ dF(x_0)$. □

Definition 1.14. Let $F : U \subseteq X \to Y$ be differentiable in U. We say that $F \in C^1(U, Y)$ iff the map

$$U \ni x \longmapsto dF(x) \in \mathcal{L}(X, Y)$$

is continuous.

Analogously to what we have in finite dimensions, we present now a weaker form of differentiability that mimics the directional derivative in $\mathbb{R}^n$. This is the concept of Gateaux differential (G-differential).

Definition 1.15 (Gateaux derivatives)**.** Let $F : U \to Y$ and $x \in U$. We say F is G-differentiable at x if there exists $A \in \mathcal{L}(X, Y)$ such that for all $h \in X$ it holds

$$\lim_{t \to 0} \frac{F(x + th) - F(x)}{t} = Ah. \tag{1.5}$$

The map A is uniquely determined and we call it the G-differential of F at x, denoting it by $d^G F(x)$.

Exercise 1.16. Prove that G-differentiability does not imply continuity.

Exercise 1.17.

1. Prove that $d^G F(x)$ is unique;

2. Assume that F is F-differentiable, prove that then F is also G-differentiable;

3. Prove that the converse of the previous exercise does not hold.

Exercise 1.18. Let $f : \mathbb{R} \to \mathbb{R}$, $f \in C^1$. Let $X = C([0,1])$ with the norm

$$\|u\|_X = \|u\|_\infty = \sup_{t\in[0,1]} |u(t)|.$$

Define $F : X \to X$ as $F(u) = f \circ u$, i.e. $F(u)(t) = f(u(t))$. Prove that F is differentiable on X and compute its differential.

Exercise 1.19. Define $F : L^2([0,1]) \to L^2([0,1])$ by

$$u(x) \longmapsto F(u)(x) = \sin(u(x)).$$

Prove that

- F is Lipschitz continuous,
- F is G-differentiable at 0,
- F is not F-differentiable at 0.

2 Differential calculus - 2 - 17/02/2026

The question whether there is a connection or not between F-differentiable and G-differentiable is answered in the following Theorem.

Theorem 2.1. *Let X, Y be Banach spaces, $U \subset X$ open, and $F : U \to Y$. Assume F is G-differentiable in U and that $d^G F : U \to \mathcal{L}(X, Y)$ is continuous at some $x_0 \in U$. Then F is F-differentiable at x_0 and*

$$dF(x_0) = d^G F(x_0).$$

Before proving it, we state and prove a result equivalent to the Mean Value Theorem for finite dimensional spaces.

Theorem 2.2 (Mean value)**.** *Let X, Y be Banach spaces, $U \subset X$ open, and let $F : U \to Y$ be G-differentiable in U. Consider $u, v \in U$ such that the segment connecting them is entirely contained in U, i.e.*

$$[u, v] := \{\varepsilon u + (1 - \varepsilon)v : \varepsilon \in [0, 1]\} \subset U.$$

Then

$$\|F(u) - F(v)\|_Y \leq \sup_{w \in [u,v]} \left\|d^G F(w)\right\|_{\mathcal{L}(X,Y)} \|u - v\|_X . \tag{2.1}$$

Proof. If $F(u) = F(v)$ there is nothing to prove, so assume $F(u) \neq F(v)$ and set $y := F(u) - F(v) \neq 0$. By a standard corollary of the Hahn–Banach theorem, there exists $\psi \in Y^*$ such that

$$\|\psi\|_{Y^*} = 1, \qquad \psi(y) = \|y\|_Y = \|F(u) - F(v)\|_Y .$$

Define the path $\gamma : [0, 1] \to U$ by

$$\gamma(t) := (1 - t)v + tu \in [u, v] \subset U,$$

and the scalar function $h : [0, 1] \to \mathbb{R}$ by

$$h(t) := \psi\big(F(\gamma(t))\big).$$

We claim that h is differentiable on $(0, 1)$ and that

$$h'(t) = \psi\big(d^G F(\gamma(t))[u - v]\big) \qquad \forall t \in (0, 1). \tag{2.2}$$

Indeed, for $\tau \to 0$ we have $\gamma(t + \tau) = \gamma(t) + \tau(u - v)$ and hence

$$\begin{aligned} \frac{h(t + \tau) - h(t)}{\tau} &= \psi\left(\frac{F(\gamma(t + \tau)) - F(\gamma(t))}{\tau}\right) \\ &= \psi\left(\frac{F(\gamma(t) + \tau(u - v)) - F(\gamma(t))}{\tau}\right) \xrightarrow[\tau \to 0]{} \psi\big(d^G F(\gamma(t))[u - v]\big), \end{aligned}$$

which proves (2.2). In particular, h is continuous on $[0, 1]$ and differentiable on $(0, 1)$, so by the classical (one-dimensional) mean value theorem there exists $\theta \in (0, 1)$ such that

$$h(1) - h(0) = h'(\theta).$$

Therefore,

$$\begin{aligned} \|F(u) - F(v)\|_Y &= \psi(F(u) - F(v)) \\ &= h(1) - h(0) = h'(\theta) \\ &= \psi\big(d^G F(\gamma(\theta))[u - v]\big) \leq \|\psi\|_{Y^*} \left\|d^G F(\gamma(\theta))\right\|_{\mathcal{L}(X,Y)} \|u - v\|_X \\ &\leq \sup_{w \in [u,v]} \left\|d^G F(w)\right\|_{\mathcal{L}(X,Y)} \|u - v\|_X , \end{aligned}$$

which is exactly (2.1). □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. Fix $x_0 \in U$ as in the statement. For h small enough so that $x_0 + h \in U$, define

$$R(h) := F(x_0 + h) - F(x_0) - d^G F(x_0)[h].$$

We claim that

$$R(h) = o(\|h\|_X) \qquad \text{as } \|h\|_X \to 0,$$

which implies that F is F-differentiable at x_0 with $dF(x_0) = d^G F(x_0)$ (by uniqueness of the Fréchet differential).

First observe that R is G-differentiable and, for every $k \in X$,

$$d^G R(h)[k] = d^G F(x_0 + h)[k] - d^G F(x_0)[k].$$

Apply Theorem 2.2 to the map R on the segment $[0, h] \subset X$. Since

$$[0, h] = \{th : t \in [0, 1]\} \quad \text{and} \quad x_0 + [0, h] \subset U$$

for $\|h\|_X$ small, we obtain

$$\begin{aligned} \|R(h)\|_Y &= \|R(h) - R(0)\|_Y \\ &\le \sup_{t\in[0,1]} \left\|d^G R(th)\right\|_{\mathcal{L}(X,Y)} \|h\|_X \\ &= \sup_{t\in[0,1]} \left\|d^G F(x_0 + th) - d^G F(x_0)\right\|_{\mathcal{L}(X,Y)} \|h\|_X . \end{aligned}$$

Dividing by $\|h\|_X$ (for $h \neq 0$) yields

$$\frac{\|R(h)\|_Y}{\|h\|_X} \le \sup_{t\in[0,1]} \left\|d^G F(x_0 + th) - d^G F(x_0)\right\|_{\mathcal{L}(X,Y)} .$$

By continuity of $d^G F$ at x_0, the right-hand side converges to 0 as $\|h\|_X \to 0$, hence

$$\lim_{\|h\|_X \to 0} \frac{\|R(h)\|_Y}{\|h\|_X} = 0,$$

which is exactly the Fréchet differentiability of F at x_0 with differential $dF(x_0) = d^G F(x_0)$. □

Proposition 2.3. *Let X, Y be Banach spaces, $U \subset X$ open and connected, and let $F : U \to Y$ be G-differentiable in U with $d^G F = 0$ on U. Then F is constant on U.*

Proof. Fix $x_0 \in U$ and define

$$V := \{x \in U : \ F(x) = F(x_0)\}.$$

Then $V \neq \emptyset$ (since $x_0 \in V$). We prove that V is both closed and open in U.

1. V **is closed in** U**.** Let $(x_n) \subset V$ and assume $x_n \to x$ in U. Since U is open, there exists $r > 0$ such that $B_X(x, r) \subset U$. For every $y \in B_X(x, r)$, the segment $[x, y] \subset B_X(x, r) \subset U$, hence by Theorem 2.2,

$$\|F(y) - F(x)\|_Y \le \sup_{w\in[x,y]} \left\|d^G F(w)\right\|_{\mathcal{L}(X,Y)} \|y - x\|_X = 0.$$

Thus $F(y) = F(x)$ for every $y \in B_X(x, r)$, i.e. F is constant on $B_X(x, r)$. For n large enough, $x_n \in B_X(x, r)$, and therefore

$$F(x) = F(x_n) = F(x_0).$$

Hence $x \in V$, so V is closed in U.

2. V **is open in** U

   Let $x \in V$. Since U is open, there exists $r > 0$ such that $B_X(x,r) \subset U$. Fix $y \in B_X(x,r)$ and set $h := y - x$. Then $[x, x+h] = [x,y] \subset B_X(x,r) \subset U$. Applying the Mean Value Theorem 2.2 gives

$$\|F(y) - F(x)\|_Y \leq \sup_{w \in [x,y]} \left\|d^G F(w)\right\|_{\mathcal{L}(X,Y)} \|y - x\|_X . \tag{2.3}$$

   But $d^G F = 0$ on U, hence the supremum is 0 and therefore $F(y) = F(x)$. Since $x \in V$, $F(x) = F(x_0)$, so $F(y) = F(x_0)$ and thus $y \in V$. We have shown $B_X(x,r) \subset V$, i.e. V is open in U.

The set V is nonempty, open and closed in the connected set U. Hence $V = U$, i.e. F is constant on U.

□

Definition 2.4 (Partial derivatives)**.** Let X, Y, Z be normed spaces and $F : X \times Y \to Z$. We say that F has partial derivative with respect to $x \in X$ at $(x_0, y_0) \in X \times Y$ if the map

$$x \longmapsto F(x, y_0)$$

is F-differentiable at x_0. We denote its differential by $\partial_x F(x_0, y_0) \in \mathcal{L}(X, Z)$. Analogously, we define the partial derivative w.r.t. y, denoted by $\partial_y F(x_0, y_0) \in \mathcal{L}(Y, Z)$.

Exercise 2.5. Assume F admits partial derivatives $\partial_x F$ and $\partial_y F$ in a neighbourhood of (x_0, y_0) which are continuous as maps

$$\partial_x F : X \times Y \to \mathcal{L}(X, Z), \qquad \partial_y F : X \times Y \to \mathcal{L}(Y, Z).$$

Prove that F is F-differentiable at (x_0, y_0) and that

$$dF(x_0, y_0)[h, k] = \partial_x F(x_0, y_0)[h] + \partial_y F(x_0, y_0)[k].$$

2.1 Higher order derivatives - Schwartz Theorem

Definition 2.6 (Twice differentiability)**.** Let $F : U \subset X \to Y$. We say that F is twice differentiable if the differential of F, $dF : U \to \mathcal{L}(X, Y)$, is differentiable. For $x \in U$, we denote the second differential by $d^2F(x) \in \mathcal{L}(X, \mathcal{L}(X, Y))$.

It is convenient to see $d^2F(x)$ as a bilinear map on X, namely an element of the set

$$\mathcal{L}_2(X, Y) := \{B : X \times X \to Y : \ B \text{ is bilinear and continuous}\}.$$

For this reason, let us prove the following result.

Proposition 2.7. *$\mathcal{L}(X, \mathcal{L}(X, Y))$ is isometrically isomorphic to $\mathcal{L}_2(X, Y)$.*

Proof. Define $\Phi : \mathcal{L}(X, \mathcal{L}(X, Y)) \to \mathcal{L}_2(X, Y)$ by

$$\Phi(A)(x_1, x_2) := A(x_1)[x_2], \qquad A \in \mathcal{L}(X, \mathcal{L}(X, Y)).$$

Then $\Phi(A)$ is bilinear, and for all $x_1, x_2 \in X$,

$$\|\Phi(A)(x_1, x_2)\|_Y = \|A(x_1)[x_2]\|_Y \leq \|A\|_{\mathcal{L}(X,\mathcal{L}(X,Y))} \|x_1\|_X \|x_2\|_X ,$$

so $\Phi(A) \in \mathcal{L}_2(X, Y)$ and $\|\Phi(A)\| \leq \|A\|$.

Conversely, define $\Psi : \mathcal{L}_2(X, Y) \to \mathcal{L}(X, \mathcal{L}(X, Y))$ by

$$\Psi(B)(x_1)[x_2] := B(x_1, x_2), \qquad B \in \mathcal{L}_2(X, Y).$$

Then $\Psi(B)$ is linear, and

$$\|\Psi(B)(x_1)\|_{\mathcal{L}(X,Y)} = \sup_{\|x_2\|_X \leq 1} \|B(x_1, x_2)\|_Y \leq \|B\| \|x_1\|_X ,$$

hence $\Psi(B) \in \mathcal{L}(X, \mathcal{L}(X, Y))$ and $\|\Psi(B)\| \leq \|B\|$.

Finally, it is immediate that $\Psi(\Phi(A)) = A$ and $\Phi(\Psi(B)) = B$. Moreover, from the two inequalities above we get

$$\|\Phi(A)\| \leq \|A\| \leq \|\Phi(A)\| , \qquad \|\Psi(B)\| \leq \|B\| \leq \|\Psi(B)\| ,$$

so both maps are isometries and inverse to each other. □

Exercise 2.8. Let $B : X \times X \to Y$ be bilinear. Show that TFAE:

1. B is continuous at $(0,0)$;
2. B is continuous at any point (x_1, x_2);
3. B is bounded on $\|x_1\|_X \le 1$, $\|x_2\|_X \le 1$;
4. there exists a constant $M > 0$ such that $\|B(x_1,x_2)\|_Y \le M \|x_1\|_X \|x_2\|_X$ for all $x_1, x_2 \in X$.

Remark 2.9. The space $\mathcal{L}_2(X,Y)$ is a vector space with norm

$$\|B\| := \sup_{x_1\neq 0,\, x_2\neq 0} \frac{\|B(x_1,x_2)\|_Y}{\|x_1\|_X \|x_2\|_X} = \inf\{M > 0 : \text{ item (4) holds}\}. \tag{2.4}$$

Example 2.10. Consider $F(u) = u^2$ defined on $X = C([0,1])$. Then $F : X \to X$ is twice differentiable and

$$dF(u)[h] = 2uh, \qquad d^2F(u)[h,k] = 2hk.$$

Definition 2.11. We define $\mathcal{L}_2^s(X,Y)$ to be the set of bilinear operators from $X \times X$ to Y that are symmetric, namely

$$B(x_1,x_2) = B(x_2,x_1) \qquad \forall x_1, x_2 \in X.$$

Exercise 2.12. Prove that $\mathcal{L}_2^s(X,Y)$ is a closed subset of $\mathcal{L}_2(X,Y)$.

Theorem 2.13 (Schwartz Theorem)**.** *If $F : U \to Y$ is twice differentiable at $x \in U$, then $d^2F(x)$ is symmetric.*

Proof. Fix $h, k \in X$. We prove that

$$d^2F(x)[h,k] = d^2F(x)[k,h].$$

Define, for h fixed and k variable,

$$G_h(k) := F(x+h+k) - F(x+k) - F(x+h) + F(x) - d^2F(x)[h,k].$$

Note that $G_h(0) = 0$.

Step 1. Since dF is F-differentiable at x with differential $d^2F(x)$, for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big\|dF(x+m) - dF(x) - d^2F(x)[m,\cdot]\big\|_{\mathcal{L}(X,Y)} \le \varepsilon \|m\|_X \qquad \forall m \in X,\ \|m\|_X < \delta. \tag{2.5}$$

For $\|h\|_X, \|k\|_X < \delta$, the map $k \mapsto G_h(k)$ is F-differentiable and, for $\ell \in X$,

$$\begin{aligned} dG_h(k)[\ell] &= dF(x+h+k)[\ell] - dF(x+k)[\ell] - d^2F(x)[h,\ell] \\ &= \Big(dF(x+h+k) - dF(x) - d^2F(x)[h+k,\cdot]\Big)[\ell] \\ &\quad - \Big(dF(x+k) - dF(x) - d^2F(x)[k,\cdot]\Big)[\ell]. \end{aligned}$$

Taking norms and using (2.5) with $m = h+k$ and $m = k$ gives

$$\|dG_h(k)\|_{\mathcal{L}(X,Y)} \le \varepsilon \|h+k\|_X + \varepsilon \|k\|_X \le 2\varepsilon(\|h\|_X + \|k\|_X).$$

Applying Theorem 2.2 to G_h on the segment $[0,k]$ yields

$$\begin{aligned} \|G_h(k)\|_Y = \|G_h(k) - G_h(0)\|_Y &\le \sup_{t\in[0,1]} \|dG_h(tk)\|_{\mathcal{L}(X,Y)} \ \|k\|_X \\ &\le 2\varepsilon(\|h\|_X + \|k\|_X) \ \|k\|_X \le 2\varepsilon(\|h\|_X + \|k\|_X)^2. \end{aligned}$$

Step 2. Define similarly

$$G_k(h) := F(x+h+k) - F(x+h) - F(x+k) + F(x) - d^2F(x)[k,h].$$

Repeating Step 1 with the roles of h and k swapped, for $\|h\|_X, \|k\|_X < \delta$ we get

$$\|G_k(h)\|_Y \le 2\varepsilon(\|h\|_X + \|k\|_X)^2.$$

Observing that

$$G_h(k) - G_k(h) = d^2F(x)[k,h] - d^2F(x)[h,k],$$

we obtain, for $\|h\|_X, \|k\|_X < \delta$,

$$\left\|d^2F(x)[h,k] - d^2F(x)[k,h]\right\|_Y \le \|G_h(k)\|_Y + \|G_k(h)\|_Y \le 4\varepsilon(\|h\|_X + \|k\|_X)^2. \tag{2.6}$$

Step 3. Let $h, k \in X$ be arbitrary and fix $\varepsilon > 0$. Let $\delta > 0$ be as in (2.5). If $h = k = 0$ we are done; otherwise set

$$\tilde{h} := \frac{\delta}{2(\|h\|_X + \|k\|_X)}\, h, \qquad \tilde{k} := \frac{\delta}{2(\|h\|_X + \|k\|_X)}\, k.$$

Then $\left\|\tilde{h}\right\|_X + \left\|\tilde{k}\right\|_X = \delta/2$, in particular $\left\|\tilde{h}\right\|_X, \left\|\tilde{k}\right\|_X < \delta$, so (2.6) applies to $(\tilde{h}, \tilde{k})$:

$$\left\|d^2F(x)[\tilde{h},\tilde{k}] - d^2F(x)[\tilde{k},\tilde{h}]\right\|_Y \le 4\varepsilon\Big(\left\|\tilde{h}\right\|_X + \left\|\tilde{k}\right\|_X\Big)^2 = \varepsilon\delta^2.$$

By bilinearity of $d^2F(x)$,

$$d^2F(x)[\tilde{h},\tilde{k}] = \left(\frac{\delta}{2(\|h\|_X + \|k\|_X)}\right)^2 d^2F(x)[h,k], \quad d^2F(x)[\tilde{k},\tilde{h}] = \left(\frac{\delta}{2(\|h\|_X + \|k\|_X)}\right)^2 d^2F(x)[k,h].$$

Hence

$$\left(\frac{\delta}{2(\|h\|_X + \|k\|_X)}\right)^2 \left\|d^2F(x)[h,k] - d^2F(x)[k,h]\right\|_Y \le \varepsilon\delta^2,$$

and therefore

$$\left\|d^2F(x)[h,k] - d^2F(x)[k,h]\right\|_Y \le 4\varepsilon(\|h\|_X + \|k\|_X)^2.$$

Since $\varepsilon > 0$ is arbitrary, we conclude $\left\|d^2F(x)[h,k] - d^2F(x)[k,h]\right\|_Y = 0$, i.e. $d^2F(x)[h,k] = d^2F(x)[k,h]$. □

3 Differential calculus - 3 - 23/02/2026

We do not necessarily have to stop at two derivatives.

Definition 3.1. Let $\mathcal{L}_n(X,Y)$ be the space of n-linear and continuous operators $B : X^n \to Y$. For $B \in \mathcal{L}_n(X,Y)$ we define

$$\|B\| := \sup_{\|x_1\|_X \le 1, \dots, \|x_n\|_X \le 1} \|B(x_1,\dots,x_n)\|_Y \tag{3.1}$$

$$= \sup_{x_1,\dots,x_n \neq 0} \frac{\|B(x_1,\dots,x_n)\|_Y}{\|x_1\|_X \cdots \|x_n\|_X} \tag{3.2}$$

$$= \inf\Big\{ M > 0 : \ \|B(x_1,\dots,x_n)\|_Y \le M \|x_1\|_X \cdots \|x_n\|_X \ \ \forall x_1,\dots,x_n \in X \Big\}. \tag{3.3}$$

Moreover, we define $\mathcal{L}^s_n(X,Y)$ as the space of symmetric n-linear continuous operators, i.e.

$$B(x_1,\dots,x_n) = B(x_{\sigma(1)},\dots,x_{\sigma(n)}) \qquad \forall \sigma \in S_n.$$

We can infer from the discussion above the following (true) facts.

1. The characterization of $\mathcal{L}_2$ holds also for $\mathcal{L}_n$ (TFAE exercise).
2. If Y is Banach, then also $\mathcal{L}_n(X,Y)$ is Banach.
3. There is an isometric isomorphism
$$\mathcal{L}\big(X, \mathcal{L}_{n-1}(X,Y)\big) \simeq \mathcal{L}_n(X,Y),$$
given by $A \mapsto B_A$ with $B_A(x_1,\dots,x_n) := A(x_1)[x_2,\dots,x_n]$.
4. If F is differentiable n times, then $d^nF(x) \in \mathcal{L}^s_n(X,Y)$ for every x where $d^nF(x)$ exists.

Exercise 3.2. Let $X = C([0,1])$ and $F : X \to X$ be defined by $F(u) = u^n$ (pointwise). Compute d^kF for all $k \in \mathbb{N}$.

Exercise 3.3. Let $f : X \to \mathbb{R}$ and assume f has a (local) minimum at x_0. Prove that $df(x_0) = 0$ and that

$$d^2f(x_0)[h,h] \ge 0 \qquad \forall h \in X.$$

Proof. Fix $h \in X$ and define $\varphi(t) := f(x_0 + th)$ for t small. Since x_0 is a local minimum of f, we have that $t = 0$ is a local minimum of φ, hence $\varphi'(0) = 0$ and $\varphi''(0) \ge 0$. On the other hand,

$$\varphi'(0) = df(x_0)[h], \qquad \varphi''(0) = d^2f(x_0)[h,h],$$

so $df(x_0)[h] = 0$ for all h (i.e. $df(x_0) = 0$) and $d^2f(x_0)[h,h] \ge 0$ for all h. □

Remark 3.4. If X is finite dimensional, the counterpart of the previous exercise holds true: if $df(x_0) = 0$ and

$$d^2f(x_0)[h,h] > 0 \qquad \forall h \neq 0,$$

then x_0 is a (strict) local minimum.

Exercise 3.5. Let $\ell^2(\mathbb{N}) = \{x = (x_n)_{n\in\mathbb{N}} : \ \|x\|^2 = \sum_{n\in\mathbb{N}} |x_n|^2 < \infty\}$ and let $f : \ell^2(\mathbb{N}) \to \mathbb{R}$ be defined by

$$f(x) = \sum_{n\ge1} \frac{x_n^2}{n^3} - \sum_{n\ge1} x_n^4. \tag{3.4}$$

Show that:

1. $f \in C^2$;
2. $df(0) = 0$;
3. $d^2f(0) > 0$ (i.e. $d^2f(0)[h,h] > 0$ for all $h \neq 0$);
4. $x = 0$ is **not** a local minimum.

3.1 Taylor formula in Banach spaces

Definition 3.6 (Integration)**.** Let X be a Banach space and $[a,b] \subset \mathbb{R}$.

- A *simple function* is a map $f : [a,b] \to X$ for which there exist a partition $a = t_0 < t_1 < \cdots < t_m = b$ and vectors $x_1, \ldots, x_m \in X$ such that

$$f(t) = \sum_{i=1}^{m} x_i \, \chi_{(t_{i-1},t_i]}(t).$$

- For such f we define its (Riemann) integral by

$$\int_a^b f(t)\mathrm{d}t := \sum_{i=1}^{m} (t_i - t_{i-1})\, x_i \in X.$$

- The map $f \mapsto \int_a^b f$ is linear on the space of simple functions, and satisfies the bound

$$\left\| \int_a^b f(t)\mathrm{d}t \right\|_X \leq (b-a)\ \|f\|_\infty\,, \qquad \|f\|_\infty := \sup_{t\in[a,b]} \|f(t)\|_X\,.$$

- Since continuous functions can be uniformly approximated by simple functions, the integral extends uniquely by continuity to a bounded linear operator on $C([a,b],X)$ (still denoted by $\int_a^b$), and for $f \in C([a,b],X)$ one has

$$\int_a^b f(t)\mathrm{d}t = \lim_{|\Pi|\to 0} \sum_{i=1}^{m} f(\xi_i)\,(t_i - t_{i-1}), \qquad \xi_i \in (t_{i-1}, t_i).$$

Proposition 3.7 (Properties of the integral)**.** *The following properties hold true:*

- $$\left\| \int_a^b f(t)\mathrm{d}t \right\|_X \leq \|f\|_\infty\, |b-a| \tag{3.5}$$

  for $f \in C([a,b],X)$;

- *If $A \in \mathcal{L}(X,Y)$ and $f \in C([a,b],X)$, then*

$$A\Big(\int_a^b f(t)\mathrm{d}t\Big) = \int_a^b Af(t)\mathrm{d}t. \tag{3.6}$$

Proof. The estimate (3.5) holds for simple functions by construction, and then extends to $C([a,b],X)$ by uniform approximation. For (3.6), first check it for simple functions by linearity:

$$A\Big(\sum_{i=1}^{m}(t_i - t_{i-1})x_i\Big) = \sum_{i=1}^{m}(t_i - t_{i-1})Ax_i.$$

Then extend again by uniform approximation and continuity of A. □

Remark 3.8. Identification of $\mathcal{L}(\mathbb{R},X)$ with X: if $L \in \mathcal{L}(\mathbb{R},X)$, then $L(t) = t\,L(1)$ for all $t \in \mathbb{R}$, and

$$\mathcal{L}(\mathbb{R},X) \ni L \longleftrightarrow L(1) \in X$$

is a linear isometric isomorphism.

Theorem 3.9 (Fundamental Theorem of Calculus)**.** *Let $F \in C^1([a,b],X)$. Then*

$$F(t) = F(a) + \int_a^t dF(s)\mathrm{d}s. \tag{3.7}$$

Conversely, if $f \in C([a,b],X)$ then $F(t) = \int_a^t f(s)\mathrm{d}s \in C^1([a,b],X)$ and $dF(t) = f(t)$.

Proof. Step 1: the converse statement. Let $f \in C([a,b],X)$ and define

$$G(t) := \int_a^t f(s)\mathrm{d}s.$$

Fix $t \in (a,b)$ and $\varepsilon \to 0$ such that $t+\varepsilon \in [a,b]$. Then

$$\begin{aligned}\frac{G(t+\varepsilon)-G(t)}{\varepsilon} - f(t) &= \frac{1}{\varepsilon}\int_t^{t+\varepsilon} f(s)\mathrm{d}s - f(t) \\ &= \frac{1}{\varepsilon}\int_t^{t+\varepsilon} \big(f(s)-f(t)\big)\mathrm{d}s.\end{aligned}$$

Hence, using (3.5),

$$\left\|\frac{G(t+\varepsilon)-G(t)}{\varepsilon} - f(t)\right\|_X \le \sup_{s\in[t,t+\varepsilon]} \|f(s)-f(t)\|_X \xrightarrow[\varepsilon\to 0]{} 0$$

by continuity of f. Therefore G is differentiable at t and $G'(t) = f(t)$; the continuity of G' follows from continuity of f. Thus $G \in C^1([a,b],X)$ and $dG(t) = f(t)$.

Step 2: the direct statement. Let $F \in C^1([a,b],X)$ and set $f := dF \in C([a,b],X)$. Define

$$H(t) := F(t) - F(a) - \int_a^t f(s)\mathrm{d}s.$$

Then $H \in C^1([a,b],X)$ and

$$dH(t) = dF(t) - f(t) = 0 \qquad \forall t \in [a,b].$$

Hence H is constant on $[a,b]$, and since $H(a) = 0$ we conclude $H(t) = 0$ for all t, i.e. (3.7). □

Theorem 3.10 (Taylor formula)**.** *Let $F : U \subset X \to Y$, with $F \in C^{k+1}(U,Y)$. Then,*

$$F(u+h) = F(u) + dF(u)[h] + \frac{1}{2}d^2F(u)[h,h] + \cdots + \frac{1}{k!}d^kF(u)[h,\dots,h] + R(u,h), \tag{3.8}$$

where

$$R(u,h) = \frac{1}{k!}\int_0^1 (1-t)^k d^{k+1}F(u+th)[h,\dots,h]\mathrm{d}t \tag{3.9}$$

and

$$\lim_{\|h\|_X\to 0} \frac{\|R(u,h)\|_Y}{\|h\|_X^{k+1}} = 0. \tag{3.10}$$

Proof. Assume $u, u+h \in U$ and define $\gamma(t) := u + th$ for $t \in [0,1]$, so that $\gamma([0,1]) \subset U$ for $\|h\|_X$ small enough. Set

$$\varphi(t) := F(\gamma(t)) = F(u+th) \in Y.$$

By the chain rule, $\varphi \in C^{k+1}([0,1],Y)$ and, for $j = 1,\dots,k+1$,

$$\varphi^{(j)}(t) = d^jF(u+th)[\underbrace{h,\dots,h}_{j \text{ times}}].$$

Apply the one-dimensional Taylor formula (with integral remainder) to φ between 0 and 1:

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!}\varphi''(0) + \cdots + \frac{1}{k!}\varphi^{(k)}(0) + \frac{1}{k!}\int_0^1 (1-t)^k\varphi^{(k+1)}(t)\mathrm{d}t.$$

Using $\varphi(1) = F(u+h)$, $\varphi(0) = F(u)$ and the expressions for $\varphi^{(j)}(t)$ gives (3.8)–(3.9). Finally, since

$$\left\|d^{k+1}F(u+th)[h,\dots,h]\right\|_Y \le \left\|d^{k+1}F(u+th)\right\|_{\mathcal{L}_{k+1}(X,Y)} \|h\|_X^{k+1},$$

we obtain

$$\|R(u,h)\|_Y \le \frac{1}{k!}\Big(\sup_{t\in[0,1]}\left\|d^{k+1}F(u+th)\right\|_{\mathcal{L}_{k+1}(X,Y)}\Big)\,\|h\|_X^{k+1}\int_0^1 (1-t)^k\mathrm{d}t.$$

Since $\int_0^1(1-t)^k\mathrm{d}t = \frac{1}{k+1}$ and $d^{k+1}F$ is continuous at u, the supremum stays bounded and (3.10) follows. □

3.2 Polynomials and inverse Taylor formula

Definition 3.11 (Monomial)**.** Let $n \in \mathbb{N}$ and let $A \in \mathcal{L}^s_n(X,Y)$. The associated *monomial of degree* n is the map $p : X \to Y$ defined by

$$p(x) := A[x,\dots,x] =: A[x]^n.$$

Definition 3.12 (Space of monomials)**.** For $k \in \mathbb{N}$, we define the space of (continuous) monomials of degree k by

$$\mathcal{P}^k(X,Y) := \Big\{ p : X \to Y : \ \exists A \in \mathcal{L}^s_k(X,Y) \text{ s.t. } p(x) = A[x]^k \ \ \forall x \in X \Big\}. \tag{3.11}$$

We endow $\mathcal{P}^k(X,Y)$ with the norm

$$\|p\| := \sup_{\|x\|_X \le 1} \|p(x)\|_Y = \sup_{x \neq 0} \frac{\|p(x)\|_Y}{\|x\|_X^k} = \inf\Big\{ M > 0 : \ \|p(x)\|_Y \le M \, \|x\|_X^k \ \ \forall x \in X \Big\}.$$

Remark 3.13. If $p \in \mathcal{P}^k(X,Y)$, then p is homogeneous of degree k:

$$p(\lambda x) = \lambda^k p(x) \qquad \forall \lambda \in \mathbb{R}, \ \forall x \in X.$$

Moreover, p is continuous (indeed C^∞ as a polynomial map).

Proposition 3.14 (Polarization formula)**.** *Let $n \in \mathbb{N}$, let $A \in \mathcal{L}^s_n(X,Y)$ and set $p(x) := A[x]^n$. Then for all $x_1,\dots,x_n \in X$,*

$$A[x_1,\dots,x_n] = \frac{1}{2^n n!} \sum_{\varepsilon_1,\dots,\varepsilon_n \in \{\pm 1\}} \varepsilon_1 \cdots \varepsilon_n \, p(\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n). \tag{3.12}$$

Proof. Set

$$S := \sum_{\varepsilon_1,\dots,\varepsilon_n \in \{\pm 1\}} \varepsilon_1 \cdots \varepsilon_n \, p(\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n) \in Y.$$

We compute S by iterating the one-dimensional fundamental theorem of calculus.

First, separate the sum with respect to $\varepsilon_1 = \pm 1$:

$$S = \sum_{\varepsilon_2,\dots,\varepsilon_n} \varepsilon_2 \cdots \varepsilon_n \Big(p(x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n) - p(-x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n) \Big).$$

For fixed $\varepsilon_2,\dots,\varepsilon_n$, define $g(t) := p(t x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n)$. Then

$$g(1) - g(-1) = \int_{-1}^{1} g'(t)\mathrm{d}t = \int_{-1}^{1} dp(t x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n)[x_1]\mathrm{d}t.$$

Therefore

$$S = \sum_{\varepsilon_2,\dots,\varepsilon_n} \varepsilon_2 \cdots \varepsilon_n \int_{-1}^{1} dp(t_1 x_1 + \varepsilon_2 x_2 + \cdots + \varepsilon_n x_n)[x_1]\mathrm{d}t_1.$$

Repeat the same argument for $\varepsilon_2,\dots,\varepsilon_n$ (each time replacing a difference with an integral of a derivative). After n steps one obtains

$$S = \int_{[-1,1]^n} d^n p(t_1 x_1 + \cdots + t_n x_n)[x_1,\dots,x_n]\mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

Since $p(x) = A[x]^n$ with A symmetric n-linear, the n-th differential is constant and equals

$$d^n p(\cdot) = n! \, A \in \mathcal{L}^s_n(X,Y),$$

hence

$$S = \int_{[-1,1]^n} n! \, A[x_1,\dots,x_n]\mathrm{d}t_1 \cdots \mathrm{d}t_n = 2^n n! \, A[x_1,\dots,x_n].$$

Dividing by $2^n n!$ yields (3.12). □

Example 3.15. Polarization formula for $n = 2$. If $p(x) = A[x, x]$ with $A \in \mathcal{L}_2^s(X, Y)$, then

$$A[x_1, x_2] = \frac{1}{8}\Big(p(x_1 + x_2) - p(x_1 - x_2) - p(-x_1 + x_2) + p(-x_1 - x_2)\Big).$$

Since p is homogeneous of degree 2, $p(-x) = p(x)$ and the formula reduces to the classical identity

$$A[x_1, x_2] = \frac{1}{4}\big(p(x_1 + x_2) - p(x_1 - x_2)\big).$$

Proposition 3.16. *There is a bijection $\phi : \mathcal{L}_n^s(X, Y) \to \mathcal{P}^n(X, Y)$ given by*

$$\phi(A)(x) = A[x]^n.$$

Moreover, for $p = \phi(A)$,

$$\|p\| \le \|A\| \le \frac{n^n}{n!}\,\|p\|\,. \tag{3.13}$$

Proof. Step 1: ϕ is well-defined and injective. If $A \in \mathcal{L}_n^s(X, Y)$ then $p(x) := A[x]^n$ is a continuous monomial of degree n, hence $p \in \mathcal{P}^n(X, Y)$. If $\phi(A_1) = \phi(A_2)$, then the associated monomials coincide, and by the polarization formula (3.12) we recover $A_1 = A_2$, hence ϕ is injective.

Step 2: ϕ is surjective. If $p \in \mathcal{P}^n(X, Y)$, by definition there exists $A \in \mathcal{L}_n^s(X, Y)$ with $p(x) = A[x]^n$, hence $p = \phi(A)$.

Step 3: the norm inequalities. First, for $\|x\|_X \le 1$,

$$\|p(x)\|_Y = \|A[x]^n\|_Y \le \|A\|\ \|x\|_X^n \le \|A\|\,,$$

so $\|p\| \le \|A\|$.

For the converse bound, fix $x_1, \dots, x_n \in X$ with $\|x_i\|_X \le 1$. By the polarization formula and homogeneity of p,

$$\begin{aligned}
\|A[x_1, \dots, x_n]\|_Y &\le \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n \in \{\pm 1\}} \|p(\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n)\|_Y \\
&\le \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n} \|p\|\ \|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\|_X^n \\
&\le \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n} \|p\|\ (\|x_1\|_X + \cdots + \|x_n\|_X)^n \\
&\le \frac{1}{2^n n!} \sum_{\varepsilon_1, \dots, \varepsilon_n} \|p\|\ n^n = \frac{n^n}{n!}\,\|p\|\,.
\end{aligned}$$

Taking the supremum over $\|x_i\|_X \le 1$ yields $\|A\| \le \frac{n^n}{n!}\,\|p\|$, completing (3.13). □

4 Differential calculus - 4 - 24/02/2026

Theorem 4.1 (Converse Taylor formula)**.** *Let X, Y be Banach spaces, $U \subset X$ open, and $F : U \to Y$. Fix $n \in \mathbb{N}$. Assume that there exist maps*

$$a_j : U \to \mathcal{L}^s_j(X, Y), \qquad j = 1, \dots, n,$$

which are continuous and such that for every $x \in U$ and every $h \in X$ with $x + h \in U$ one has

$$F(x+h) = F(x) + a_1(x)[h] + \cdots + \frac{1}{n!}\, a_n(x)[h, \dots, h] + R_n(x, h), \tag{4.1}$$

where the remainder is locally uniform *in the following sense: for every $x_0 \in U$ there exists a neighborhood $V(x_0) \subset U$ such that*

$$\forall \varepsilon > 0\ \exists \delta > 0 : \ \big(x \in V(x_0),\ \|h\|_X \le \delta,\ x + h \in U\big) \implies \|R_n(x, h)\|_Y \le \varepsilon \, \|h\|^n_X \,. \tag{4.2}$$

Then $F \in C^n(U, Y)$ and

$$d^j F(x) = a_j(x) \qquad \forall x \in U,\ \forall j = 1, \dots, n.$$

Proof. We argue by induction on n.

Step 1: the case $n = 1$. Assumption (4.1) reads

$$F(x + h) = F(x) + a_1(x)[h] + o(\|h\|_X) \quad \text{locally uniformly in } x,$$

which is exactly the definition of Fréchet differentiability of F at x with $dF(x) = a_1(x)$. Since a_1 is continuous, we get $F \in C^1(U, Y)$.

Step 2: induction step. Assume the statement holds up to $n - 1$ (with $n \ge 2$). Since (4.2) implies $R_n(x, h) = o(\|h\|^{n-1}_X)$ locally uniformly, from (4.1) we also have a Taylor-type expansion up to order $n - 1$ with locally uniform remainder. By the induction hypothesis,

$$F \in C^{n-1}(U, Y) \qquad \text{and} \qquad d^j F(x) = a_j(x) \quad \forall j = 1, \dots, n-1.$$

It remains to recover $d^n F$.

Claim. For every $x \in U$,

$$\big\| d^{n-1}F(x + h) - d^{n-1}F(x) - a_n(x)[h, \cdot, \dots, \cdot] \big\|_{\mathcal{L}_{n-1}(X,Y)} = o(\|h\|_X) \qquad \text{as } \|h\|_X \to 0. \tag{4.3}$$

This implies that $d^{n-1}F$ is Fréchet differentiable at x and

$$d(d^{n-1}F)(x) = a_n(x) \in \mathcal{L}\big(X, \mathcal{L}_{n-1}(X, Y)\big) \simeq \mathcal{L}_n(X, Y),$$

hence $F \in C^n(U, Y)$ and $d^n F(x) = a_n(x)$. Since a_n is continuous by assumption, we also get $F \in C^n(U, Y)$.

Proof of the claim. Fix $x \in U$ and take $h, k \in X$ small with

$$\frac{\|h\|_X}{2} \le \|k\|_X \le \|h\|_X \,, \qquad x + h \in U, \qquad x + h + k \in U. \tag{4.4}$$

Apply (4.1) at the point $x + h$ with increment k:

$$F(x+h+k) = F(x+h) + dF(x+h)[k] + \cdots + \frac{1}{(n-1)!}\, d^{n-1}F(x+h)[k, \dots, k] + \frac{1}{n!} a_n(x+h)[k, \dots, k] + R_n(x+h, k). \tag{4.5}$$

Apply (4.1) at the point x with increment $h + k$:

$$F(x+h+k) = F(x) + dF(x)[h+k] + \cdots + \frac{1}{(n-1)!}\, d^{n-1}F(x)[h+k, \dots, h+k] + \frac{1}{n!} a_n(x)[h+k, \dots, h+k] + R_n(x, h+k). \tag{4.6}$$

Subtracting (4.6) from (4.5) and grouping terms as a polynomial in k yields an identity of the form

$$0 = g_0(h) + g_1(h)[k] + \cdots + g_{n-1}(h)[k,\dots,k] + g_n(h)[k,\dots,k], \tag{4.7}$$

where $g_j(h) \in \mathcal{L}^s_j(X,Y)$ for $j = 0,\dots,n-1$ and, in particular,

$$g_{n-1}(h) = \frac{1}{(n-1)!}\Big(d^{n-1}F(x+h) - d^{n-1}F(x) - a_n(x)[h,\cdot,\dots,\cdot]\Big) \in \mathcal{L}^s_{n-1}(X,Y). \tag{4.8}$$

Moreover, the last coefficient $g_n(h)[k,\dots,k]$ collects the terms involving $a_n(x+h)-a_n(x)$ and the remainders R_n. Using (4.4), continuity of a_n, and the locally uniform estimate (4.2), we have

$$\|(a_n(x+h)-a_n(x))[k,\dots,k]\|_Y = o(\|h\|_X^n), \quad \|R_n(x+h,k)\|_Y = o(\|h\|_X^n), \quad \|R_n(x,h+k)\|_Y = o(\|h\|_X^n).$$

Hence (4.7) implies

$$g_0(h) + g_1(h)[k] + \cdots + g_{n-1}(h)[k,\dots,k] = o(\|h\|_X^n). \tag{4.9}$$

Vandermonde argument. Fix $k \in X$ satisfying (4.4). Choose distinct scalars $\lambda_0, \lambda_1, \dots, \lambda_{n-1} \in [0,1]$ and set $k_\ell := \lambda_\ell k$. Substituting $k = k_\ell$ into (4.9) gives, for every $\ell = 0,\dots,n-1$,

$$g_0(h) + \lambda_\ell g_1(h)[k] + \lambda_\ell^2 g_2(h)[k,k] + \cdots + \lambda_\ell^{n-1} g_{n-1}(h)[k,\dots,k] = o(\|h\|_X^n).$$

This is a linear system with Vandermonde matrix

$$\begin{pmatrix} 1 & \lambda_0 & \lambda_0^2 & \cdots & \lambda_0^{n-1} \\ 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{n-1} & \lambda_{n-1}^2 & \cdots & \lambda_{n-1}^{n-1} \end{pmatrix},$$

which is invertible since $\lambda_\ell \neq \lambda_m$ for $\ell \neq m$. Therefore we deduce, in particular,

$$g_{n-1}(h)[k,\dots,k] = o(\|h\|_X^n) \qquad \text{whenever } \frac{\|h\|_X}{2} \le \|k\|_X \le \|h\|_X. \tag{4.10}$$

Now take $v \in X$ with $\|v\|_X = 1$ and set $k := \|h\|_X\, v$. Then (4.4) holds for $\|h\|_X$ small and (4.10) gives

$$g_{n-1}(h)[v,\dots,v] = \frac{1}{\|h\|_X^{n-1}}\, g_{n-1}(h)[k,\dots,k] = \frac{o(\|h\|_X^n)}{\|h\|_X^{n-1}} = o(\|h\|_X).$$

Taking the supremum over $\|v\|_X \le 1$ yields

$$\|g_{n-1}(h)\|_{\mathcal{L}_{n-1}(X,Y)} = o(\|h\|_X),$$

and using (4.8) this is exactly (4.3). The claim (and hence the theorem) follows. □

Exercise 4.2. Compute the determinant of the Vandermonde matrix

$$V(\lambda_0,\dots,\lambda_{n-1}) := \begin{pmatrix} 1 & \lambda_0 & \lambda_0^2 & \cdots & \lambda_0^{n-1} \\ 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{n-1} & \lambda_{n-1}^2 & \cdots & \lambda_{n-1}^{n-1} \end{pmatrix},$$

and deduce that the Vandermonde matrix is invertible if and only if $\lambda_l \neq \lambda_m$ for all $l \neq m$.

4.1 Contraction principle

Definition 4.3 (Contraction and fixed points)**.** Let X be a Banach space, $C \subset X$ and $F : C \to C$. A point $\bar{x} \in C$ is a *fixed point* of F if $F(\bar{x}) = \bar{x}$. We say that F is a *contraction* if there exists $\theta \in [0, 1)$ such that

$$\|F(x) - F(y)\|_X \leq \theta \|x - y\|_X \qquad \forall x, y \in C.$$

For $n \in \mathbb{N}$ we set $F^n := \underbrace{F \circ \cdots \circ F}_{n \text{ times}}$ and $F^0 := \mathrm{Id}$.

Theorem 4.4 (Contraction principle)**.** *Let X be a Banach space, $C \subset X$ a nonempty closed set, and let $F : C \to C$ be a contraction with constant $\theta \in [0, 1)$. Then F has a unique fixed point $\bar{x} \in C$. Moreover, for every $x_0 \in C$, the sequence $x_{n+1} := F(x_n)$ satisfies $x_n \to \bar{x}$ and*

$$\|F^n(x_0) - \bar{x}\|_X \leq \frac{\theta^n}{1-\theta} \|F(x_0) - x_0\|_X \qquad \forall n \in \mathbb{N}. \tag{4.11}$$

Proof. Uniqueness. If $\bar{x}_1, \bar{x}_2$ are fixed points, then

$$\|\bar{x}_1 - \bar{x}_2\|_X = \|F(\bar{x}_1) - F(\bar{x}_2)\|_X \leq \theta \|\bar{x}_1 - \bar{x}_2\|_X ,$$

hence $(1 - \theta) \|\bar{x}_1 - \bar{x}_2\|_X \leq 0$ and thus $\bar{x}_1 = \bar{x}_2$.

Existence. Fix $x_0 \in C$ and define $x_{n+1} := F(x_n)$. By contraction,

$$\|x_{n+1} - x_n\|_X = \|F(x_n) - F(x_{n-1})\|_X \leq \theta \|x_n - x_{n-1}\|_X \leq \cdots \leq \theta^n \|x_1 - x_0\|_X .$$

Hence for $u > m$,

$$\|x_u - x_m\|_X \leq \sum_{j=m+1}^{u} \|x_j - x_{j-1}\|_X \leq \sum_{j=m+1}^{u} \theta^{j-1} \|x_1 - x_0\|_X \leq \frac{\theta^m}{1-\theta} \|x_1 - x_0\|_X ,$$

and the right-hand side tends to 0 as $m \to \infty$. Thus (x_n) is a Cauchy sequence. Since C is closed and X is complete, there exists $\bar{x} \in C$ such that $x_n \to \bar{x}$. By continuity of F,

$$F(\bar{x}) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = \bar{x},$$

so $\bar{x}$ is a fixed point.

Finally, estimate (4.11) follows by letting $u \to \infty$ in

$$\|x_u - x_n\|_X \leq \frac{\theta^n}{1-\theta} \|x_1 - x_0\|_X$$

and recalling $x_n = F^n(x_0)$ and $x_1 - x_0 = F(x_0) - x_0$. □

4.2 Inverse and implicit function theorems

We now employ the contraction principle to local invertibility for mappings between Banach spaces. We introduce some definitions.

Definition 4.5 (Invertibility for linear operators)**.** Let X, Y be Banach spaces and let $A \in \mathcal{L}(X, Y)$. We say that A is *invertible* if there exists $B \in \mathcal{L}(Y, X)$ such that

$$AB = \mathrm{Id}_Y \qquad \text{and} \qquad BA = \mathrm{Id}_X.$$

In this case B is unique and it is denoted by A^{-1}. We write

$$\mathrm{Inv}(X, Y) := \{A \in \mathcal{L}(X, Y) : A \text{ is invertible}\}.$$

Definition 4.6 (Local invertibility for continuous mappings)**.** Let X, Y be Banach spaces, $U \subset X$ open, and $F \in C(U, Y)$. We say that F is *locally invertible at* $\bar{x} \in U$ if there exist open sets $W \subset U$ and $V \subset Y$ such that $\bar{x} \in W$, $F(\bar{x}) \in V$, and a map

$$G \in C(V, W)$$

satisfying

$$F \circ G = \mathrm{Id}_V \qquad \text{and} \qquad G \circ F = \mathrm{Id}_W.$$

In this case, G is called a *local inverse* of F at $\bar{x}$ and we write $G = F^{-1}$ (on V).

Remark 4.7. We have from the definition the two following properties:

- *Transitivity:* let X_1, X_2, X_3 be Banach spaces, $U_1 \subset X_1$, $U_2 \subset X_2$ open, and

  $$F_1 \in C(U_1, X_2), \qquad F_2 \in C(U_2, X_3).$$

  Assume that F_1 is locally invertible at $\bar{x} \in U_1$ and that F_2 is locally invertible at $F_1(\bar{x}) \in U_2$. Then $F_2 \circ F_1$ is locally invertible at $\bar{x}$, and on suitable neighborhoods one has

  $$(F_2 \circ F_1)^{-1} = F_1^{-1} \circ F_2^{-1}.$$

- *Stability:* if F is locally invertible at $\bar{x}$ with local inverse $G : V \to W$, then F is locally invertible at every $x \in W$. Indeed, for each such x we have $F(x) \in V$ and the same identities

  $$F \circ G = \mathrm{Id}_V, \qquad G \circ F = \mathrm{Id}_W$$

  show that G is also a local inverse of F at x (possibly after restricting to smaller neighborhoods of x and $F(x)$).

5 Differential calculus - 5 - 02/03/2026

Theorem 5.1 (Inverse function theorem)**.** *Let X, Y be Banach spaces, $U \subset X$ open, $F \in C^1(U, Y)$, and fix $u_0 \in U$. Assume that $dF(u_0) \in \mathcal{L}(X, Y)$ is invertible. Then F is locally invertible at u_0. Moreover, the local inverse F^{-1} is of class C^1 and satisfies*

$$d(F^{-1})(y) = \Big(dF(F^{-1}(y))\Big)^{-1} \qquad \text{for every } y \text{ in the domain of } F^{-1}.$$

Proof. **Step 1: reduction to a normalized setting.** Set $y_0 := F(u_0)$ and $A := dF(u_0)^{-1} \in \mathcal{L}(Y, X)$. Consider the translated/rescaled map

$$\widetilde{F}(x) := A\big(F(u_0 + x) - y_0\big), \qquad x \in \widetilde{U} := \{x \in X : \ u_0 + x \in U\}.$$

Then $\widetilde{U} \subset X$ is open, $\widetilde{F} \in C^1(\widetilde{U}, X)$,

$$\widetilde{F}(0) = 0, \qquad d\widetilde{F}(0) = A\, dF(u_0) = \mathrm{Id}_X.$$

If $\widetilde{F}$ is locally invertible at 0, then F is locally invertible at u_0 by undoing the affine changes of variables. Hence, from now on we assume

$$Y = X, \qquad u_0 = 0, \qquad F(0) = 0, \qquad dF(0) = \mathrm{Id}_X.$$

Step 2: fixed point formulation. Define

$$G(x) := F(x) - x, \qquad x \in U.$$

Then $G(0) = 0$ and

$$dG(0) = dF(0) - \mathrm{Id}_X = 0.$$

By continuity of dG at 0, there exists $s > 0$ such that $\overline{B_s(0)} \subset U$ and

$$\sup_{x \in \overline{B_s(0)}} \|dG(x)\|_{\mathcal{L}(X,X)} \leq \frac{1}{2}. \tag{5.1}$$

Fix $y \in \overline{B_{s/2}(0)}$ and consider

$$T_y(x) := y - G(x), \qquad x \in \overline{B_s(0)}.$$

For any $x_1, x_2 \in \overline{B_s(0)}$, by the mean value estimate along the segment $[x_1, x_2]$ (one can write $G(x_2) - G(x_1) = \int_0^1 dG(x_1 + t(x_2 - x_1))[x_2 - x_1]\mathrm{d}t$),

$$\|T_y(x_2) - T_y(x_1)\|_X = \|G(x_2) - G(x_1)\|_X \leq \sup_{z \in [x_1, x_2]} \|dG(z)\|_{\mathcal{L}(X,X)} \ \|x_2 - x_1\|_X \leq \frac{1}{2} \|x_2 - x_1\|_X ,$$

where we used (5.1). Thus T_y is a contraction on $\overline{B_s(0)}$. Moreover, for $x \in \overline{B_s(0)}$ we have

$$\|T_y(x)\|_X \leq \|y\|_X + \|G(x) - G(0)\|_X \leq \frac{s}{2} + \sup_{z \in [0,x]} \|dG(z)\|_{\mathcal{L}(X,X)} \ \|x\|_X \leq \frac{s}{2} + \frac{1}{2} s = s,$$

so $T_y(\overline{B_s(0)}) \subset \overline{B_s(0)}$.

By the Banach fixed point theorem, there exists a unique $x \in \overline{B_s(0)}$ such that $x = T_y(x)$, i.e.

$$x = y - G(x) \quad \Longleftrightarrow \quad F(x) = y.$$

Hence F is bijective from $B_s(0)$ onto $B_{s/2}(0)$, with inverse

$$F^{-1} : B_{s/2}(0) \to B_s(0), \qquad F^{-1}(y) = \text{the unique fixed point of } T_y.$$

Step 3: continuity (in fact Lipschitz) of the inverse. Let $y_1, y_2 \in B_{s/2}(0)$ and set $x_i := F^{-1}(y_i)$. Using $x_i = T_{y_i}(x_i)$ we obtain

$$x_2 - x_1 = T_{y_2}(x_2) - T_{y_1}(x_1) = (y_2 - y_1) + (G(x_1) - G(x_2)).$$

Taking norms and using (5.1),

$$\|x_2 - x_1\|_X \le \|y_2 - y_1\|_X + \|G(x_2) - G(x_1)\|_X \le \|y_2 - y_1\|_X + \frac{1}{2}\|x_2 - x_1\|_X,$$

hence

$$\left\|F^{-1}(y_2) - F^{-1}(y_1)\right\|_X = \|x_2 - x_1\|_X \le 2\|y_2 - y_1\|_X. \tag{5.2}$$

Step 4: differentiability of the inverse and formula for the differential. First we note that $dF(x)$ is invertible for x sufficiently close to 0: since $dF(0) = \mathrm{Id}_X$ and dF is continuous, shrinking s if needed we can assume

$$\sup_{x \in \overline{B_s(0)}} \|dF(x) - \mathrm{Id}_X\|_{\mathcal{L}(X,X)} < 1.$$

Then each $dF(x)$ is invertible on $\overline{B_s(0)}$ (Neumann series), and the inverses are uniformly bounded on this set.

Fix $y \in B_{s/2}(0)$ and set $x := F^{-1}(y) \in B_s(0)$. Let $h \in X$ be such that $y + h \in B_{s/2}(0)$ and set $x_h := F^{-1}(y + h)$. Then

$$F(x_h) - F(x) = h.$$

Since F is (Fréchet) differentiable at x, we can write

$$F(x_h) = F(x) + dF(x)[x_h - x] + r(x_h - x), \qquad \frac{\|r(k)\|_X}{\|k\|_X} \xrightarrow[k\to 0]{} 0.$$

Thus

$$h = dF(x)[x_h - x] + r(x_h - x).$$

Apply $dF(x)^{-1}$ to get

$$x_h - x - dF(x)^{-1}[h] = -\,dF(x)^{-1}\big[r(x_h - x)\big].$$

Divide by $\|h\|_X$ and use (5.2) (which gives $\|x_h - x\|_X \le 2\|h\|_X$):

$$\frac{\left\|x_h - x - dF(x)^{-1}[h]\right\|_X}{\|h\|_X} \le \left\|dF(x)^{-1}\right\|_{\mathcal{L}(X,X)} \frac{\|r(x_h - x)\|_X}{\|x_h - x\|_X} \frac{\|x_h - x\|_X}{\|h\|_X} \xrightarrow[h\to 0]{} 0.$$

Hence F^{-1} is differentiable at y and

$$d(F^{-1})(y) = dF(F^{-1}(y))^{-1}.$$

Step 5: C^1 regularity of the inverse. We have a composition of continuous maps

$$y \longmapsto F^{-1}(y) \longmapsto dF(F^{-1}(y)) \longmapsto \big(dF(F^{-1}(y))\big)^{-1}.$$

The first arrow is continuous by (5.2), the second is continuous because dF is continuous, and the inversion map is continuous on $\mathrm{Inv}(\mathcal{L}(X,X))$. Therefore $y \mapsto d(F^{-1})(y)$ is continuous, i.e. $F^{-1} \in C^1$. □

Example 5.2 (C^1 regularity cannot be dropped)**.** Let $X = Y = C([-1,1])$ with $\|u\|_X = \|u\|_\infty$ and define $F : X \to Y$ by

$$F(u)(t) = \varphi(u(t)),$$

where $\varphi : [-1,1] \to \mathbb{R}$ is continuous, nondecreasing, differentiable at 0 with $\varphi'(0) = 1$, but φ has "plateaux" accumulating at 0 as follows. For every $n \ge 1$ set

$$a_n := \frac{1}{n} - \frac{1}{4n^2}, \qquad b_n := \frac{1}{n} + \frac{1}{4n^2}, \qquad I_n := [a_n, b_n] \cap [0,1],$$

and assume

$$\varphi(t) = \frac{1}{n} \quad \text{for all } t \in I_n.$$

(On each gap (b_{n+1}, a_n) one may take φ affine so that φ is continuous and nondecreasing, and set $\varphi(t) = t$ for $t \le 0$.)

Show that F is differentiable at 0 and $dF(0) = \mathrm{Id}_X$, but F is *not* locally invertible at 0.

Proof. **Step 1: differentiability of F at 0 and $dF(0) = \mathrm{Id}_X$.** By construction one can arrange (and we assume) that there exists $C > 0$ such that, for all $t \in [-1, 1]$,

$$|\varphi(t) - t| \le Ct^2. \tag{5.3}$$

(Indeed, on each plateau I_n one has $|\varphi(t) - t| = |1/n - t| \le 1/(4n^2)$, and on the affine gaps the maximum deviation from the identity is attained at the endpoints; in both cases the deviation is $O(1/n^2) = O(t^2)$.)

Let $u \in X$. Using (5.3),

$$\|F(u) - u\|_\infty = \sup_{t\in[-1,1]} |\varphi(u(t)) - u(t)| \le C \sup_{t\in[-1,1]} |u(t)|^2 = C \|u\|_\infty^2 .$$

Hence

$$\frac{\|F(u) - F(0) - \mathrm{Id}_X[u]\|_\infty}{\|u\|_\infty} = \frac{\|F(u) - u\|_\infty}{\|u\|_\infty} \le C \|u\|_\infty \xrightarrow[u\to 0]{} 0,$$

so F is Fréchet differentiable at 0 and $dF(0) = \mathrm{Id}_X$.

Step 2: F is not locally surjective near 0. For each $n \ge 1$ define $v_n \in Y$ by

$$v_n(t) := \frac{1}{n} + \frac{t}{n^2}, \qquad t \in [-1, 1].$$

Then $\|v_n\|_\infty \le \frac{1}{n} + \frac{1}{n^2} \to 0$ as $n \to \infty$.

We claim that $v_n \notin \mathrm{Ran}(F)$ for every n. Assume by contradiction that there exists $u_n \in X$ with $F(u_n) = v_n$, i.e.

$$\varphi(u_n(t)) = \frac{1}{n} + \frac{t}{n^2} \qquad \forall t \in [-1, 1].$$

If $t > 0$, then $\varphi(u_n(t)) > 1/n$. Since φ is nondecreasing and $\varphi \equiv 1/n$ on $[a_n, b_n]$, it follows that necessarily $u_n(t) \ge b_n = \frac{1}{n} + \frac{1}{4n^2}$. Similarly, if $t < 0$ then $\varphi(u_n(t)) < 1/n$, hence $u_n(t) \le a_n = \frac{1}{n} - \frac{1}{4n^2}$.

Therefore, for every $t > 0$ and $t < 0$ we have the one-sided bounds

$$u_n(t) \ge \frac{1}{n} + \frac{1}{4n^2} \quad (t > 0), \qquad u_n(t) \le \frac{1}{n} - \frac{1}{4n^2} \quad (t < 0).$$

Letting $t \downarrow 0$ and $t \uparrow 0$ gives

$$\liminf_{t\downarrow 0} u_n(t) \ \ge \ \frac{1}{n} + \frac{1}{4n^2}, \qquad \limsup_{t\uparrow 0} u_n(t) \ \le \ \frac{1}{n} - \frac{1}{4n^2},$$

which is incompatible with continuity of u_n at 0. This contradiction shows $v_n \notin \mathrm{Ran}(F)$.

Step 3: failure of local invertibility. If F were locally invertible at 0, then (in particular) it would be locally surjective: there would exist $r > 0$ such that $B_r(0) \subset \mathrm{Ran}(F)$. But for n large we have $\|v_n\|_\infty < r$ while $v_n \notin \mathrm{Ran}(F)$, a contradiction. Hence F is not locally invertible at 0 despite F being differentiable at 0 with $dF(0) = \mathrm{Id}_X$. □

Theorem 5.3 (Implicit function)**.** *Let X, Y, Z be Banach spaces, $W \subset X \times Y$ open, and $F : W \to Z$ with $F \in C^1(W, Z)$. Fix $(u_0, v_0) \in W$ such that*

$$F(u_0, v_0) = 0, \qquad D_Y F(u_0, v_0) \in \mathcal{L}(Y, Z) \text{ is invertible.}$$

Then there exist open sets $U \subset X$, $V \subset Y$ with $u_0 \in U$ and $v_0 \in V$ and a unique map $G : U \to V$, $G \in C^1(U, Y)$ such that

- *$F(u, G(u)) = 0$ for all $u \in U$;*
- *if $F(u, v) = 0$ with $(u, v) \in U \times V$, then $v = G(u)$.*

Moreover, for every $u \in U$,

$$dG(u)[h] = -\Big(D_Y F(u, G(u))\Big)^{-1} \Big(D_X F(u, G(u))[h]\Big), \qquad h \in X.$$

Proof. Consider the *graph map*

$$H : W \subset X \times Y \longrightarrow X \times Z, \qquad H(u,v) := (u, F(u,v)).$$

Clearly $H \in C^1(W, X \times Z)$. Its differential at (u_0, v_0) is the block operator

$$dH(u_0,v_0)[h,k] = \big(h,\ D_X F(u_0,v_0)[h] + D_Y F(u_0,v_0)[k]\big), \qquad (h,k) \in X \times Y.$$

Equivalently, in block form (with respect to $X \times Y \to X \times Z$),

$$dH(u_0,v_0) = \begin{pmatrix} \mathrm{Id}_X & 0 \\ D_X F(u_0,v_0) & D_Y F(u_0,v_0) \end{pmatrix}.$$

Since $D_Y F(u_0,v_0)$ is invertible, $dH(u_0,v_0)$ is invertible with inverse

$$dH(u_0,v_0)^{-1} = \begin{pmatrix} \mathrm{Id}_X & 0 \\ -\big(D_Y F(u_0,v_0)\big)^{-1} D_X F(u_0,v_0) & \big(D_Y F(u_0,v_0)\big)^{-1} \end{pmatrix}.$$

By the inverse function theorem applied to H at (u_0,v_0), there exist neighborhoods $W_1 \subset W$ of (u_0,v_0) and $W_2 \subset X \times Z$ of $(u_0,0)$ and a C^1-diffeomorphism

$$\phi : W_2 \to W_1, \qquad H \circ \phi = \mathrm{Id}_{W_2}.$$

Write $\phi(x,z) = (\phi_1(x,z), \phi_2(x,z))$ with $\phi_1 : W_2 \to X$ and $\phi_2 : W_2 \to Y$. Then, for every $(x,z) \in W_2$,

$$(x,z) = H(\phi(x,z)) = \big(\phi_1(x,z),\ F(\phi_1(x,z), \phi_2(x,z))\big),$$

hence

$$\phi_1(x,z) = x, \qquad F\big(x, \phi_2(x,z)\big) = z.$$

Now restrict to $z = 0$: define

$$G(x) := \phi_2(x,0).$$

For U a small open neighborhood of u_0 with $(x,0) \in W_2$ for all $x \in U$, set $V := G(U)$. Then $G \in C^1(U,Y)$ and

$$F(x, G(x)) = 0 \qquad \forall x \in U.$$

For the uniqueness statement, if $(x,y) \in U \times V \subset W_1$ and $F(x,y) = 0$, then

$$H(x,y) = (x,0) = H(x, G(x)),$$

and since H is injective on W_1 we get $y = G(x)$.

Finally, differentiating $F(u, G(u)) = 0$ at u in direction $h \in X$ gives

$$0 = D_X F(u, G(u))[h] + D_Y F(u, G(u))[dG(u)[h]],$$

and since $D_Y F(u, G(u))$ is invertible for u close enough to u_0, we obtain

$$dG(u)[h] = -\big(D_Y F(u, G(u))\big)^{-1}\big(D_X F(u, G(u))[h]\big).$$

□

5.1 Smooth dependence w.r.t. initial data

Theorem 5.4 (Smooth dependence w.r.t. initial condition)**.** *Assume $f \in C^1(\mathbb{R}_t \times \mathbb{R}^n_x, \mathbb{R}^n)$ and consider the ODE system*

$$\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0. \tag{5.4}$$

Then there exist $\delta > 0$ and a neighborhood $\mathcal{U}$ of x_0 such that, for every $\xi \in \mathcal{U}$, the problem (5.4) (with $x_0 = \xi$) has a unique solution $x(\,\cdot\,, \xi) \in C^1([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$. Moreover, the map

$$\mathcal{U} \ni \xi \longmapsto x(\,\cdot\,, \xi) \in C^1([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$$

is of class C^1. In particular, for every fixed $t \in [t_0 - \delta, t_0 + \delta]$, the map $\xi \mapsto x(t,\xi)$ is C^1.

If $f \in C^k$ (with $k \geq 1$), then $\xi \mapsto x(\,\cdot\,, \xi)$ is C^k (hence $\xi \mapsto x(t,\xi)$ is C^k for each fixed t).

Proof. Fix $\delta > 0$ so small that the (unique) solution $\phi_0(\cdot) := x(\cdot, x_0)$ of (5.4) exists on $I := [t_0 - \delta, t_0 + \delta]$ (local existence/uniqueness follows from Picard–Lindelöf since f is C^1).

Set the Banach space

$$X := C^1(I, \mathbb{R}^n), \qquad \|\varphi\|_X := \|\varphi\|_\infty + \|\dot\varphi\|_\infty .$$

Define the map

$$F : X \times \mathbb{R}^n \longrightarrow X, \qquad F(\varphi, \xi)(t) := \varphi(t) - \xi - \int_{t_0}^{t} f(\tau, \varphi(\tau))\, \mathrm{d}\tau.$$

Then $F(\phi_0, x_0) = 0$, and one checks that $F \in C^1(X \times \mathbb{R}^n, X)$ (in fact C^k if $f \in C^k$).

We apply the implicit function theorem to $F(\varphi, \xi) = 0$ solving for φ as a function of ξ. Compute the partial differential with respect to φ: for $h \in X$,

$$d_\varphi F(\varphi, \xi)[h](t) = h(t) - \int_{t_0}^{t} D_x f(\tau, \varphi(\tau))\, h(\tau)\, \mathrm{d}\tau.$$

Fix $(\varphi, \xi) = (\phi_0, x_0)$ and abbreviate

$$A(t) := D_x f(t, \phi_0(t)) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \qquad t \in I.$$

We claim that the linear operator $L := d_\varphi F(\phi_0, x_0) \in \mathcal{L}(X, X)$ is invertible. Indeed, given $g \in X$, we want $h \in X$ such that

$$h(t) - \int_{t_0}^{t} A(\tau) h(\tau)\, \mathrm{d}\tau = g(t), \qquad t \in I.$$

Differentiating in t yields the linear non-homogeneous ODE

$$\dot h(t) = A(t) h(t) + \dot g(t), \qquad h(t_0) = g(t_0).$$

Since A is continuous on the compact interval I, this ODE has a unique solution $h \in C^1(I, \mathbb{R}^n)$. Equivalently, using the resolvent matrix $M(t)$ solving

$$\dot M(t) = A(t) M(t), \qquad M(t_0) = \mathrm{Id}_{\mathbb{R}^n},$$

we have the explicit formula

$$h(t) = M(t) g(t_0) + M(t) \int_{t_0}^{t} M(\tau)^{-1} \dot g(\tau)\, \mathrm{d}\tau,$$

which shows linear dependence on g. Standard estimates (e.g. Grönwall applied to M and M^{-1}) imply that the map $g \mapsto h$ is bounded $X \to X$, hence L is invertible in $\mathcal{L}(X, X)$.

Therefore, by the implicit function theorem there exist a neighborhood $\mathcal{U}$ of x_0 in $\mathbb{R}^n$ and a unique map $\Phi : \mathcal{U} \to X$, $\Phi \in C^1$, such that

$$F(\Phi(\xi), \xi) = 0 \qquad \forall \xi \in \mathcal{U},$$

i.e. $\Phi(\xi)$ is the (unique) solution of (5.4) with initial condition $x(t_0) = \xi$. Thus $\xi \mapsto x(\cdot, \xi) = \Phi(\xi)$ is C^1 as a map into $C^1(I, \mathbb{R}^n)$. If $f \in C^k$, then the same argument (with $F \in C^k$) gives $\Phi \in C^k$. □

6 Nemitski operators – Nis-Erik Bohne - 09/03/2026

In this lecture we define vector valued Nemitski operators and study their properties. In a first part we derive conditions on the function f such that the associated Nemitski operator N_f is continuous form L^p to L^q for $p, q \geq 1$. In a second part we analyze the differentiability of Nemitski operators. We derive conditions on f for the Fréchet differentiability of N_f for $p > 2$ and we prove that N_f is only Gateaux differentiable for $p = 2$ and we present an example that fails to be Fréchet differentiable. Finally, provide a characterization of the function f in case N_f is Fréchet differentiable for some $u^* \in L^2$.

These lecture notes are adapted from [2, Sec.], which only deals with real valued functions f.

6.1 General setting

From the lecture on 02.03.2026, recall the ODE

$$\begin{aligned} \partial_t x(t) &= f(t, x(t)) \qquad \forall t \in \mathbb{R}, \\ x(t_0) &= x_0 \in \mathbb{R}^d, \end{aligned}$$

where $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is C^1. The goal of this lecture is to analyze the properties of operators of the form $u \mapsto f(\cdot, u(\cdot))$. In order to cover a wide range of applications, we try to be as general as possible.

Let $d, n, m \in \mathbb{N}$ be given and let $\Omega \subseteq \mathbb{R}^d$ be open and bounded. We denote the space of vector valued measurable functions by

$$M(\Omega, \mathbb{R}^n) := \{u : \Omega \to \mathbb{R}^n \mid u \text{ Lebesgue measurable}\},$$

and we denote the Lebesgue measure by μ. For $p \geq 1$ recall the space of p integrable vector valued functions functions

$$L^p(\Omega, \mathbb{R}^n) := \{u \in M(\Omega, \mathbb{R}^n) \mid \int_\Omega \|u(x)\|^p \,\mathrm{d}\,\mu(x) < \infty\},$$

where $\|\cdot\|$ is some norm on $\mathbb{R}^n$.

Let $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be given. The following assumption holds for the function f throughout this lecture.

Assumption 6.1. The function f is a Carathéodory function, i.e.,

$$s \mapsto f(x, s) \in C^0(\mathbb{R}^n, \mathbb{R}^m) \qquad \text{for a.e. } x \in \Omega, \tag{6.1a}$$

$$x \mapsto f(x, s) \in M(\Omega, \mathbb{R}^m) \qquad \forall s \in \mathbb{R}^n. \tag{6.1b}$$

Definition 6.2 (Nemitski operator)**.** For any given function f satisfying Assumption 6.1, the *Nemitski* operator $N_f : M(\Omega, \mathbb{R}^n) \to M(\Omega, \mathbb{R}^m)$ is given by

$$(N_f u)(x) := f(x, u(x)) \qquad \forall x \in \Omega. \tag{6.2}$$

Let $\chi_n \in M(\Omega, \mathbb{R}^n)$ be a sequence of component wise simple functions such that $\chi_n \overset{n\to\infty}{\longrightarrow} u$ almost everywhere. Then (6.1) guarantees that $N_f u$ is indeed measurable.

6.2 Continuity of Nemitski operators

Let $p, q \geq 1$ be given and assume that there exist constants $a, b \in \mathbb{R}_{\geq 0}$ with $b > 0$ such that

$$\|f(x, s)\| \leq a + b\|s\|^{p/q}. \tag{6.3}$$

Theorem 6.3. *If f satisfies Assumption 6.1 and (6.3), then the operator $N_f : L^p(\Omega, \mathbb{R}^n) \to L^q(\Omega, \mathbb{R}^m)$ is continuous.*

For the proof of Theorem 6.3, we need the following auxiliary result.

Proposition 6.4. *If $\mu(\Omega) < \infty$, then for every converging sequence $(u_k)_{k\in\mathbb{N}} \subset L^p(\Omega,\mathbb{R}^n)$ there exists a subsequence $(u_{k_j})_{j\in\mathbb{N}}$ such that*

$$u_{k_j} \overset{j\to\infty}{\longrightarrow} u \text{ a.e. in } \Omega, \tag{6.4a}$$

$$\|u_{k_j}\| \leq \|h\| \text{ a.e. in } \Omega, \tag{6.4b}$$

for some function $h \in L^p(\Omega,\mathbb{R}^n)$.

Idea of proof. We write $u_k = (u_k^1, u_k^2, \dots, u_k^n)^T$ and $u = (u^1, u^2, \dots, u^n)^T$. Using the 1D version of this result, see e.g., [1, Thm. IV.9], there exists a subsequence $(u_{k_{j_1}})_{j_1\in\mathbb{N}}$ such that $u_{k_{j_1}} \overset{j_1\to\infty}{\longrightarrow} u^1$ almost everywhere in Ω and $|u_{k_{j_1}}| \leq |h_1|$. Iterate the same argument for each component of u by choosing subsequences of subsequences and set $h := (h_1, h_2, \dots, h_n)^T$. □

Proof of Theorem 6.3. Since $q \geq 1$ the function $x \to x^q$ is convex and monotonically non-decreasing for $x \geq 0$. Therefore we observe that

$$\|f(x,u(x))\|^q \leq (a + b\|u(x)\|^{p/q})^q \leq 2^{q-1}(a^q + b^q\|u(x)\|^p). \tag{6.5}$$

Since Ω is bounded, $\mu(\Omega) < \infty$ and therefore $N_f u \in L^q(\Omega,\mathbb{R}^m)$ for every $u \in L^p(\Omega,\mathbb{R}^n)$.

Let $(u_k)_{k\in\mathbb{N}} \subset L^p(\Omega,\mathbb{R}^n)$ be any converging subsequence with limit $u \in L^p(\Omega,\mathbb{R}^n)$. Then there exists a subsequence $(u_{k_j})_{j\in\mathbb{N}}$ such that $u_{k_j} \overset{j\to\infty}{\longrightarrow} u$ almost everywhere and $\|u_{k_j}\| \leq \|h\|$ for some $h \in L^p(\Omega,\mathbb{R}^n)$. Since f satisfies (6.1), we have that $N_f u_{k_j} \overset{j\to\infty}{\longrightarrow} N_f u$ almost everywhere and that

$$\|N_f u_{k_j}\| \leq a + b\|h\|^{p/q}.$$

Then the dominated convergence theorem implies that $N_f u_{k_j} \overset{j\to\infty}{\longrightarrow} N_f u$ in $L^q(\Omega,\mathbb{R}^m)$. Applying Lemma 6.10 to $N_f u_k$ concludes the proof. □

Observe that Theorem 6.3 excludes the choice of $q, p = \infty$. We now provide a counter example. Let $I \subseteq \mathbb{R}$ be some bounded interval. Take any $p \geq 1$ be arbitrary but fixed and choose $q = \infty$. Let $f : I \times \mathbb{R} \to \mathbb{R}$ be given by $f(x,s) = \sin(s)$ and consider the Nemitski operator $N_f : L^p(I,\mathbb{R}) \to L^\infty(I,\mathbb{R})$ given by $(N_f u)(x) = \sin(u(x))$ for all $x \in I$. Set $u_n \in L^p(I,\mathbb{R})$ as

$$u_n := \frac{\pi}{2}\chi_{]0,n^{-p}[},$$

where again χ denotes the characteristic function. One easily computes that $\|u_n\|_{L^p} = \pi/(2n)$ and therefore u_n converges to 0 in L^p. Further we observe that $N_f(0) = 0$. By construction, we have that $N_f u_n = \chi_{]0,n^{-p}[}$. Therefore we compute that

$$\|N_f u_n - N_f(0)\|_{L^\infty} = \|N_f u_n\|_{L^\infty} = 1 \neq 0$$

and hence we have shown that $N_f : L^p(I,\mathbb{R}) \to L^\infty(I,\mathbb{R})$ is not continuous at 0.

6.3 Differentiability of Nemitski operators

For this section we fix that for a vector $x \in \mathbb{R}^n$, $\|v\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$ is the euclidean norm and for a matrix $Q \in \mathbb{R}^{m\times n}$, $\|Q\| = \sup_{v\in\mathbb{R}^n\setminus\{0\}} \|Qv\|/\|v\|$ is the matrix norm induced by the euclidean norm.

Let $p \geq 2$ be given and assume that $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ has a partial derivative $\partial_2 f(x,s) \in \mathbb{R}^{m\times n}$ almost everywhere in Ω. We assume that $\partial_2 f$ satisfies (6.1) as a matrix valued function and that

$$\|\partial_2 f(x,s)\| \leq a + b\|s\|^{p-2}, \tag{6.6}$$

for some constants $a, b \in \mathbb{R}_{\geq 0}$ and $b > 0$.

Remark 6.5. Note that Theorem 6.3 implies that $N_{\partial_2 f} : L^p(\Omega, \mathbb{R}^n) \to L^r(\Omega, \mathbb{R}^{m\times n})$ is continuous if $p > 2$ with $r = p/(p-2)$ since $\mathbb{R}^{m\times n}$ is equivalent to $\mathbb{R}^{mn}$. The case $p = 2$ implies that $r = \infty$ and this case is not covered by Theorem 6.3 and as we have seen, there are examples of non-continuous Nemitski operators from L^2 to L^∞.

Lemma 6.6. *If $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ satisfies (6.6), then for $(N_{\partial_2 f}u)v \in L^{p'}(\Omega, \mathbb{R}^m)$ for all $u, v \in L^p(\Omega, \mathbb{R}^n)$ where $p' = p/(p-1)$.*

Proof. Set $\nu := (p-1)/(p-2)$ and $\rho := p - 1$. Then by construction $1/\nu + 1/\rho = 1$. Thus an application of the Young's inequality (cf., Lemma 6.11) yields

$$\|((N_{\partial_2 f}u)v)(x)\| \le \|\partial_2 f(x, u(x))\|\|v(x)\| \le \frac{\|\partial_2 f(x, u(x))\|^\nu}{\nu} + \frac{\|v(x)\|^\rho}{\rho} \overset{(6.6)}{\le} \frac{(a + b\|u(x)\|^{p-2})^\nu}{\nu} + \frac{\|v(x)\|^\rho}{\rho}.$$

Since $(p-2)\nu p' = \rho p' = p$, using a convexity argument reveals that there exists a constant $C(p) > 0$ depending on p such that

$$\|((N_{\partial_2 f}u)v)(x)\|^{p'} \le C(p)(a^\nu + b^\nu\|u(x)\|^p + \|v(x)\|^p). \tag{6.7}$$

Integrating over Ω implies the statement, since $u, v \in L^p(\Omega, \mathbb{R}^n)$ by assumption. □

Lemma 6.6 allows us to prove the second theorem of this lecture.

Theorem 6.7. *For $p > 2$ and $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ given, assume that f and $\partial_2 f$ are both (6.1) and that $\partial_2 f$ satisfies (6.6) Additionally assume that $x \mapsto f(x, 0)$ is bounded by some constant $c > 0$ in Ω. Then the operator $N_f : L^p(\Omega, \mathbb{R}^n) \to L^{p'}(\Omega, \mathbb{R}^m)$ is Fréchet differentiable with $d(N_f u)v := (N_{\partial_2 f}u)v$ for all $u, v \in L^p(\Omega, \mathbb{R}^2)$.*

Proof. Using the fundamental theorem of calculus, we compute that

$$\|f(x, u(x))\| \le \|f(x, u(x)) - f(x, 0)\| + \|f(x, 0)\| \le \|\int_0^1 \partial_2 f(x, tu(x))u(x)\,\mathrm{d}\,t\| + c,$$

where the integral in the first summand is component wise. Then (6.6) implies that

$$\|\int_0^1 \partial_2 f(x, tu(x))u(x)\,\mathrm{d}\,t\| \le \|u(x)\| \int_0^1 (a + bt^{p-2}\|u(x)\|^{p-2})\,\mathrm{d}\,t = a\|u(x)\| + \frac{b}{p-1}\|u(x)\|^{p-1}.$$

Applying a properly scaled Young's inequality for the term $a\|u(x)\|$ implies the existence of constants $A, B > 0$ such that $\|f(x, u(x))\| \le A + B\|u(x)\|^{p-1}$. Thus N_f satisfies Theorem 6.3 and is therefore continuous from $L^p(\Omega, \mathbb{R}^n)$ to $L^{p'}(\Omega, \mathbb{R}^m)$.

For $u, v \in L^p(\Omega, \mathbb{R}^n)$, set

$$\begin{aligned}\omega(u, v) &:= \|N_f(u + v) - N_f u - (N_{\partial_2 f}u)v\|_{L^{p'}} \\ &= \left(\int_\Omega \|f(x, u(x) + v(x)) - f(x, u(x)) - \partial_2 f(x, u(x))v(x)\|^{p'}\,\mathrm{d}\,\mu(x)\right)^{1/p'}.\end{aligned}$$

Again, using the fundamental theorem of calculus, we observe that

$$\begin{aligned}&\|f(x, u(x) + v(x)) - f(x, u(x)) - \partial_2 f(x, u(x))v(x)\| \\ &\qquad = \|\int_0^1 \partial_2 f(x, u(x) + tv(x))v(x) - \partial_2 f(x, u(x))v(x)\,\mathrm{d}\,t\| \\ &\qquad = \|w(x)v(x)\|,\end{aligned}$$

where $w(x) := \int_0^1 \partial_2 f(x, u(x) + tv(x)) - \partial_2 f(x, u(x))\,\mathrm{d}\,t \in \mathbb{R}^{m\times n}$. Applying the Hölder inequality with ρ, ν as in the proof of Lemma 6.6 we obtain that

$$\begin{aligned}\omega(u, v) &\le \left(\int_\Omega \|wv\|^{p'}\,\mathrm{d}\,\mu\right)^{1/p'} \le \left(\int_\Omega \|w\|^{p'}\|v\|^{p'}\,\mathrm{d}\,\mu\right)^{1/p'} = \left\|\|w\|^{p'}\|v\|^{p'}\right\|_{L^1}^{1/p'} \\ &\le \left\|\|w\|^{p'}\right\|_{L^\nu}^{1/p'} \left\|\|v\|^{p'}\right\|_{L^\rho}^{1/p'} = \|v\|_{L^p}\|w\|_{L^r}\end{aligned}$$

with $r := p/(p-2)$. The term $\|w\|_{L^r}$ is estimated by

$$\|w\|_{L^r}^r \leq \int_\Omega \int_0^1 \|\partial_2 f(x, u(x) + tv(x)) - \partial_2 f(x, u(x))\|^r \,\mathrm{d}\, t \,\mathrm{d}\, \mu(x) = \int_0^1 \|N_{\partial_2 f}(u + tv) - N_{\partial_2 f} u\|_{L^r}^r \,\mathrm{d}\, t.$$

Since $N_{\partial_2 f}$ is continuous from $L^p(\Omega, \mathbb{R}^n) \to L^r(\Omega, \mathbb{R}^{m\times n})$ as mentioned in Remark 6.5, we observe that $\lim_{\|v\|_{L^p}\to 0} \|N_{\partial_2 f}(u + tv) - N_{\partial_2 f} u\|_{L^r} = 0$ and conclude that $\lim_{\|v\|_{L^p}\to 0} \|w\|_{L^r} = 0$ which, in turn, implies that

$$\lim_{\|v\|_{L^p}\to 0} \frac{\omega(u, v)}{\|v\|_{L^p}} = 0 \qquad \forall u \in L^p(\Omega, \mathbb{R}^n). \qquad \square$$

The case $p = 2$ is somewhat peculiar. We show that a Nemitski operator associated with $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ where f and $\partial_2 f$ satisfy (6.1) and $|\partial_2 f(x, s)| \leq c$ for some constant $c > 0$ is Gateaux differentiable possibly not Fréchet differentiable.

Theorem 6.8. *For $p = 2$ assume that $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ and $\partial_2 f$ satisfy (6.1) and (6.6). Then $N_f : L^2(\Omega, \mathbb{R}^n) \to L^2(\Omega, \mathbb{R}^m)$ is Gateaux differentiable with $\mathrm{d}^G(N_f u)v = (N_{\partial_2 f} u)v$.*

Proof. Using the fundamental theorem of calculus, we observe that

$$\left(\frac{N_f(u + tv) - N_f u}{t} - (N_{\partial_2 f} u)v\right)(x) = \left(\int_0^1 \partial_2 f(x, u(x) + \xi tv(x)) - \partial_2 f(x, u(x)) \,\mathrm{d}\, \xi\right) v(x) =: \omega_t(u, v)(x).$$

We therefore observe that

$$\begin{aligned}\left\|\frac{N_f(u + tv) - N_f u}{t} - (N_{\partial_2 f} u)v\right\|_{L^2}^2 &= \int_\Omega v^2 \omega_t(u, v)^2 \,\mathrm{d}\, \mu \\ &\leq \int_\Omega v^2 \int_0^1 \|\partial_2 f(x, u(x) + \xi tv(x)) - \partial_2 f(x, u(x))\|^2 \,\mathrm{d}\, \xi \,\mathrm{d}\, \mu(x).\end{aligned}$$

Since $\xi tv \overset{t\to 0}{\longrightarrow} 0$ almost everywhere in Ω, it follows that $\partial_2 f(x, u(x) + \xi tv(x)) - \partial_2 f(x, u(x)) \overset{t\to 0}{\longrightarrow} 0$ for almost every $x \in \Omega$. Additionally, it holds by assumption, that $\|\partial_2 f(x, u(x) + \xi tv(x)) - \partial_2 f(x, u(x))\| \leq c$ for some constant $c > 0$, we conclude by the dominated convergence theorem that

$$\lim_{t\to 0} \int_0^1 \|\partial_2 f(x, u(x) + \xi tv(x)) - \partial_2 f(x, u(x))\|^2 \,\mathrm{d}\, \xi = 0,$$

which concludes the proof. $\square$

We now give an example of a Nemitski operator from L^2 to L^2 that is Gateaux differentiable in but not Fréchet differentiable in 0. Let $I :=]0, 1[$ and let $f : I \times \mathbb{R} \to \mathbb{R}$ be given by $f(x, s) = s^2/(1 + |s|)$. f clearly satisfies (6.1) and

$$\partial_2 f(x, s) = \begin{cases} \frac{2s}{1+s} - \frac{s^2}{(1+s)^2} & s > 0, \\ 0 & s = 0, \\ \frac{2s}{1-s} + \frac{s^2}{(1-s)^2} & s < 0. \end{cases}$$

Thus f is C^1 with respect to the second variable and it is clear that $|f(x, s)| \leq 1$. First, we prove that $\mathrm{d}^G(N_f)(0) = 0$. Given any $v \in L^2(I, \mathbb{R})$ we compute that

$$\left|\frac{N_f tv}{t}\right| \leq |v|.$$

It is clear that $N_f(tv)/t \overset{t\to 0}{\longrightarrow} 0$ almost everywhere in I via a similar estimate as above. Thus by dominated convergence, it follows that $\lim_{t\to 0} \|N_f tv\|_{L^2}/t = 0$ and therefore N_f is Gateaux differentiable at 0 with $d^G(N_f 0)v = 0$ for all $v \in L^2(I, \mathbb{R})$.

Let us now assume by contradiction that N_f is Fréchet differentiable. Since Fréchet differentiability implies Gateaux differentiability and the Fréchet and Gateaux differential coincide we know that $dN_f 0 = 0$ from the previous computation. Let $v_n := n\chi_{]0,n^{-4}[}$, where χ is the characteristic function. A simple computation reveals that $\|v_n\|_{L^2} = 1/n$. Further more we observe that $N_f v_n = n^2/(1+n)\chi_{]0,n^{-4}[}$. Consequently, it follows that $\|N_f v_n\|_{L^2} = 1/(1+n)$. Thus

$$\lim_{n\to\infty} \frac{\|N_f v_n\|_{L^2}}{\|v_n\|} = \lim_{n\to\infty} \frac{1/(1+n)}{1/n} = \lim_{n\to\infty} \frac{n}{1+n} = 1 \neq 0.$$

This contradicts the assumption that N_f is Fréchet differentiable in 0.

We complete the picture for the case of $p = 2$ with a converse statement to Theorem 6.8.

Lemma 6.9. *For $p = 2$ and $f : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$, we assume that f and $\partial_2 f$ satisfy (6.1) and $\partial_2 f$ additionally satisfies (6.6). If N_f is Fréchet differentiable at some $u^* \in L^2(\Omega, \mathbb{R}^n)$, then for almost every $x \in \Omega$, there exist $a(x) \in M(\Omega, \mathbb{R}^{m\times n})$ and $b \in M(\Omega, \mathbb{R}^m)$ such that*

$$f(x,u) = a(x)u + b(x) \qquad \forall u \in L^2(\Omega, \mathbb{R}^n).$$

Proof. First, we assume that $u^* = 0$ and $f(x,0) = 0$. Let $D(y,\delta) \subseteq \Omega$ be a ball centered at $y \in \Omega$ with $\mu(D(y,\delta)) = \delta$. Given any $x^* \in \Omega$ and $\lambda \in \mathbb{R}^n$ consider $v_\delta \in L^2(\Omega, \mathbb{R}^n)$ given by

$$v_\delta(x) = \begin{cases} \lambda & x \in D(x^*, \delta), \\ 0 & x \in \Omega \setminus D(x^*, \delta). \end{cases}$$

Since N_f is Gateaux differentiable in L^2 by Theorem 6.8 and Fréchet differentiable at $u^* = 0$, we know that $\mathrm{d}(N_f 0) = \mathrm{d}^G(N_f 0) = N_{\partial_2 f} 0 \in L^2(\Omega, \mathbb{R}^{m\times n})$. As $f(x,0) = 0$ almost everywhere, we observe that

$$\lim_{\delta\to 0} \frac{\|N_f v_\delta - (N_{\partial_2 f}0)v_\delta\|_{L^2}}{\|v_\delta\|_{L^2}} = 0. \tag{6.8}$$

A direct computation reveals

$$\frac{\|N_f v_\delta - (N_{\partial_2 f}0)v_\delta\|_{L^2}}{\|v_\delta\|_{L^2}} = \frac{1}{|\lambda|\sqrt{\delta}} \left(\int_{D(x^*,\delta)} \|f(x,\lambda) - \partial_2 f(x,0)\lambda\|^2 \,\mathrm{d}\,\mu(x) \right)^{1/2}$$
$$\frac{1}{\sqrt{\delta}} \left(\int_{D(x^*,\delta)} \psi_\lambda(x) \,\mathrm{d}\,\mu(x) \right)^{1/2},$$

with $\psi_\lambda(x) := \|(f(x,\lambda) - \partial_2 f(x,0)\lambda)/\lambda\|^2$. Then from (6.8) follows that

$$\lim_{\delta\to 0} \frac{1}{\delta} \int_{D(x^*,\delta)} \psi_\lambda \,\mathrm{d}\,\mu = 0.$$

Since for all $\psi \in L^1(\Omega, \mathbb{R})$, $\lim_{\delta\to 0}(1/\delta)\int_{D(y,\delta)} \psi \,\mathrm{d}\,\mu = \psi(y)$ almost every where in Ω, we conclude that for all $\lambda \in \mathbb{R}$ there exists a set $N_\lambda \subseteq \Omega$ of measure 0, such that for all $y \in N_\lambda$, $\psi_\lambda(y) \neq 0$. Since $\mathbb{Q}^n \subseteq \mathbb{R}^n$ is a countable and dense subset, the set $N := \bigcup_{\lambda\in\mathbb{Q}} N_\lambda$ has measure 0 and by construction, we have that $\psi_\lambda(y) = 0$ for all $y \in \Omega \setminus N$ and $\lambda \in \mathbb{Q}^n$. This implies that

$$f(x,\lambda) = \partial_2 f(x,0)\lambda \qquad \forall \lambda \in \mathbb{Q}^n, \text{ and almost all } x \in \Omega. \tag{6.9}$$

The continuity of f in the second variable almost everywhere (cf., (6.1a)) implies that (6.9) holds for all $\lambda \in \mathbb{R}^n$.

Second, set $N_g(u)(x) := f(x, u(x) + u^*(x)) - f(x, u^*(x))$. Then by assumption N_g is Fréchet differentiable at 0 and $g(x,0) = 0$ (by some slight abuse of notation). Following the preceding arguments for N_g, yields that

$$f(x, u + u^*) - f(u^*) = \partial_2 f(x, u^*)u. \qquad \square$$

6.4 Technical results

Lemma 6.10. *Let $(X, \|\cdot\|)$ be any normed vector space and consider a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$. If there exists an $x \in X$ such that every subsequence $(x_{n_k})_{n_k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ has a converging subsequence $(x_{n_{k_j}})_{j\in\mathbb{N}} \subseteq (x_{n_k})_{n_k\in\mathbb{N}}$ with $\lim_{j\to\infty} x_{n_{k_j}} = x$, then $(x_n)_{n\in\mathbb{N}}$ converges with $\lim_{n\to\infty} x_n = x$.*

The proof of this Lemma only needs arguments from Analysis I and is therefore left as an exercise for the reader.

Lemma 6.11 (Young)**.** *Let $p, q \geq 1$ such that $1/p + 1/p = 1$ Then for any $A, B \geq 0$ it holds that*

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

We refer to your favourite Analysis I lecture notes for a proof.

Bibliography

[1] Haïm Brezis (1983) *Analyse fonctionnelle*, Collection Mathématiques Appliquées pour la Maîtrise, Masson, Paris.

[2] Antonio Ambrosetti and Giovanni Prodi (1993) *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, vol. 34, Cambridge University Press, Cambridge.

7 Bifurcation Theory 1 - 10/03/2026

In this section we will always assume $F \in C^2$ if not stated differently.

7.1 Bifurcation points

Example 7.1 (Motivating example)**.** Consider the ODE

$$-u'' + \lambda u = 0, \qquad u(0) = u(1) = 0. \tag{7.1}$$

This problem always has the trivial solution $u = 0$. Moreover, if $\lambda = -k^2\pi^2, k \in \mathbb{N}$, then there are nontrivial solutions, namely

$$u(x) = c\sin(k\pi x), \qquad c \in \mathbb{R}.$$

Hence at the values $\lambda = -k^2\pi^2$ we can bifurcate the "solution graph".

Definition 7.2 (Bifurcation point)**.** Let $F : \mathbb{R} \times X \to Y$ and assume that

$$F(\lambda, 0) = 0 \qquad \forall \lambda \in \mathbb{R}.$$

A number $\lambda_* \in \mathbb{R}$ is called a *bifurcation point* for F if, for every neighbourhood $\mathcal{U}$ of $(\lambda_*, 0)$ in $\mathbb{R} \times X$, there exists $(\lambda, u) \in \mathcal{U}$ such that

$$F(\lambda, u) = 0, \qquad u \neq 0.$$

Equivalently, if

$$S := \{(\lambda, u) \in \mathbb{R} \times X : F(\lambda, u) = 0,\ u \neq 0\},$$

then λ_* is a bifurcation point if and only if $(\lambda_*, 0) \in \overline{S}$.

Lemma 7.3. *If λ_* is a bifurcation point, then $\partial_u F(\lambda_*, 0) \in \mathcal{L}(X, Y)$ is not invertible.*

Proof. Assume by contradiction that $\partial_u F(\lambda_*, 0)$ is invertible. Since F is C^1, we can apply the Implicit Function Theorem 5.3 to the equation

$$F(\lambda, u) = 0$$

with respect to the variable u at the point $(\lambda_*, 0)$. Hence there exist a neighbourhood I of λ_* in $\mathbb{R}$, a neighbourhood U of 0 in X, and a unique C^1 map

$$\varphi : I \to U$$

such that, for $(\lambda, u) \in I \times U$,

$$F(\lambda, u) = 0 \quad \iff \quad u = \varphi(\lambda).$$

But $F(\lambda, 0) = 0$ for every λ, hence by uniqueness necessarily

$$\varphi(\lambda) = 0 \qquad \forall \lambda \in I.$$

Therefore, in a neighbourhood of $(\lambda_*, 0)$, the only solutions of $F(\lambda, u) = 0$ are the trivial ones $(\lambda, 0)$. This contradicts the fact that λ_* is a bifurcation point. □

Example 7.4.

1. Let $T \in \mathcal{L}(X, X)$ and $F(\lambda, u) = \lambda u - Tu$. Then

   $$\partial_u F(\lambda, 0) = \lambda I - T.$$

   Hence every bifurcation point λ_* must satisfy $\lambda_* \in \sigma(T)$.

2. Let $G \in C^1(X, X)$ with $G(0) = 0$, and let

   $$F(\lambda, u) = \lambda u - G(u).$$

   Then

   $$\partial_u F(\lambda, 0) = \lambda I - dG(0).$$

   Hence every bifurcation point λ_* must satisfy $\lambda_* \in \sigma(dG(0))$.

3. Consider
$$F(\lambda,(x,y)) = \lambda(x,y) - (x+y^3, y-x^3).$$
Then
$$\partial_u F(\lambda,(0,0)) = (\lambda-1)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$
Therefore the only possible bifurcation point is $\lambda = 1$.
However,
$$F(1,(x,y)) = (-y^3, x^3),$$
so that $F(1,(x,y)) = 0$ implies $x = y = 0$. Hence $\lambda = 1$ is *not* a bifurcation point.

Remark 7.5. The previous examples show that the spectral condition is only necessary. Indeed, consider for example the Volterra operator
$$T : C([0,1]) \to C([0,1]), \qquad (Tu)(x) := \int_0^x u(s)\,\mathrm{d}s.$$
It is well known that $\sigma(T) = \{0\}$ and if one sets $F(\lambda,u) = \lambda u - Tu$, then $\lambda = 0$ is a spectral value, but the equation $F(0,u) = 0$ reduces to $Tu = 0$, which implies $u = 0$. Hence $\lambda = 0$ is not a bifurcation point.

Exercise 7.6. Determine the bifurcation points for
$$u'' + \lambda(u - u^3) = 0, \qquad u(0) = u(\pi) = 0, \tag{7.2}$$
and the number of solutions for $\lambda \geq 0$.

7.2 Bifurcation from simple eigenvalues

Let $F \in C^2(\mathbb{R}\times X, Y)$ and assume that $F(\lambda,0) = 0\ \forall\lambda\in\mathbb{R}$. Fix $\lambda_* \in \mathbb{R}$ and set
$$L := \partial_u F(\lambda_*, 0), \qquad V := \ker L, \qquad R := \mathrm{Ran}(L).$$
We assume:

(H1) V is one-dimensional, namely $V = \mathrm{span}\{u_*\}$ for some $u_* \neq 0$;

(H2) R is closed and has codimension 1 in Y.

Choose a closed complement W of V in X and a one-dimensional closed complement Z of R in Y, so that
$$X = V \oplus W, \qquad Y = R \oplus Z,$$
and let $P : Y \to R,\ Q : Y \to Z$ be the corresponding projections.

Example 7.7. A standard situation in which assumptions *(H1)* and *(H2)* hold if for Fredholm operators of index 0 with one dimensional kernel. Moreover, it holds for operators of the form
$$L = I + K, \qquad K : X \to X \text{ compact},$$
provided one also knows that $\dim\ker(I+K) = 1$. Indeed, compact perturbations of the identity are Fredholm operators of index 0, and the additional one-dimensional kernel condition then yields *(H1)* and *(H2)*.

Following the structure given by the operator L, we proceed with a simplification of the problem which is know as the Lyapunov-Schmidt reduction. Since $V = \mathrm{span}\{u_*\}$, every $u \in X$ can be written uniquely as
$$u = tu_* + w, \qquad t \in \mathbb{R},\ w \in W.$$
Hence the equation $F(\lambda,u) = 0$ is equivalent to
$$\begin{cases} PF(\lambda, tu_* + w) = 0, & \text{range equation}\\ QF(\lambda, tu_* + w) = 0, & \text{bifurcation equation.}\end{cases} \tag{7.3}$$
The second equation is finite-dimensional, since Z is one-dimensional, while for the first equation the following Lemma applies.

Lemma 7.8. *Under the assumptions $(H1)$ and $(H2)$, there exist neighbourhoods $I_{\lambda_*} \subset \mathbb{R},\ J \subset \mathbb{R},\ U \subset W$ of λ_*, 0 (i.e. small t), and 0 (i.e. small u), respectively, and a unique map*

$$w : I_{\lambda_*} \times J \to U$$

of class C^1 such that, for $(\lambda, t, w) \in I_{\lambda_*} \times J \times U$,

$$PF(\lambda, tu_* + w) = 0 \quad \Longleftrightarrow \quad w = w(\lambda, t).$$

Moreover, $w(\lambda, 0) = 0$ for every $\lambda \in I_{\lambda_*}$, and $\partial_t w(\lambda_*, 0) = 0$.

Proof. Define

$$G(\lambda, t, w) := PF(\lambda, tu_* + w), \qquad (\lambda, t, w) \in \mathbb{R} \times \mathbb{R} \times W.$$

Then G is of class C^1 and

$$G(\lambda_*, 0, 0) = PF(\lambda_*, 0) = 0.$$

We compute the derivative with respect to w at $(\lambda_*, 0, 0)$:

$$\partial_w G(\lambda_*, 0, 0)[\eta] = P\,\partial_u F(\lambda_*, 0)[\eta] = P(L\eta), \qquad \eta \in W.$$

Since $L(W) = R$ and $L|_W : W \to R$ is injective, the map

$$L|_W : W \to R$$

is invertible as a map between Banach spaces. Moreover P acts as the identity on R, hence

$$\partial_w G(\lambda_*, 0, 0) = L|_W \in \mathrm{Inv}(W, R).$$

Therefore, by the implicit function theorem, there exist neighbourhoods $I_{\lambda_*},\ J,\ U$ and a unique C^1 map $w : I_{\lambda_*} \times J \to U$ such that

$$G(\lambda, t, w) = 0 \quad \Longleftrightarrow \quad w = w(\lambda, t).$$

Next, since $F(\lambda, 0) = 0$ for all λ, we have $PF(\lambda, 0) = 0$, and by uniqueness in the implicit function theorem, it follows that $w(\lambda, 0) = 0$ for all $\lambda \in I_{\lambda_*}$. Finally, differentiate the identity

$$PF(\lambda, tu_* + w(\lambda, t)) = 0$$

with respect to t at $(\lambda, t) = (\lambda_*, 0)$ and obtain

$$P\,\partial_u F(\lambda_*, 0)\big[u_* + \partial_t w(\lambda_*, 0)\big] = L\big(\partial_t w(\lambda_*, 0)\big) = 0.$$

since $Lu_* = 0$. But $\partial_t w(\lambda_*, 0) \in W$, whereas $\ker L = V,\ V \cap W = \{0\}$. Hence $\partial_t w(\lambda_*, 0) = 0$. □

Remark 7.9. The Lyapunov–Schmidt reduction separates the infinite-dimensional part from the genuinely bifurcating direction. Indeed, the first equation

$$PF(\lambda, tu_* + w) = 0$$

can be solved uniquely for $w = w(\lambda, t)$, so all nearby solutions must lie on the graph

$$u = tu_* + w(\lambda, t).$$

Substituting this into the second equation

$$QF(\lambda, tu_* + w(\lambda, t)) = 0$$

reduces the original problem to a single scalar equation in the two variables (λ, t).

Exercise 7.10. Since R is closed and has codimension 1, prove that there exists $\psi \in Y^* \setminus \{0\}$ such that $R = \ker \psi$.

Remark 7.11. This observation (exercise) allows one to replace the projected equation $QF = 0$ by a scalar equation, i.e. for $v \in Y$,

$$Qv = 0 \quad \Longleftrightarrow \quad \langle \psi, v \rangle = 0.$$

If we fix $\psi \in Y^* \setminus \{0\}$ such that $R = \ker \psi$, then the reduced scalar equation can be written as

$$\big\langle \psi, F(\lambda, tu_* + w(\lambda, t)) \big\rangle = 0.$$

Theorem 7.12 (Crandall–Rabinowitz)**.** *Let $F \in C^2(\mathbb{R} \times X, Y)$ satisfy*

$$F(\lambda, 0) = 0 \qquad \forall \lambda \in \mathbb{R},$$

and assume the hypotheses $(H1)$ and $(H2)$ hold. Let

$$V = \ker \partial_u F(\lambda_*, 0) = \operatorname{span}\{u_*\}, \qquad R = \operatorname{Ran}(\partial_u F(\lambda_*, 0)).$$

Assume moreover the following transversality condition

$$\partial^2_{u\lambda} F(\lambda_*, 0)[u_*] \notin R.$$

Then λ_* is a bifurcation point for F, and the set of nontrivial solutions of $F(\lambda, u) = 0$ is a unique C^1 curve in a neighbourhood of $(\lambda_*, 0)$. More precisely, there exist $\varepsilon > 0$, a map $\mu \in C^1((-\varepsilon, \varepsilon), \mathbb{R})$, and a map $\gamma \in C^1((-\varepsilon, \varepsilon), W)$ such that

$$\mu(0) = 0, \qquad \gamma(0) = 0, \qquad \gamma'(0) = 0,$$

and for every $t \in (-\varepsilon, \varepsilon)$,

$$F\big(\lambda_* + \mu(t),\, tu_* + \gamma(t)\big) = 0.$$

Moreover, every nontrivial solution sufficiently close to $(\lambda_*, 0)$ is of this form, and for $t \neq 0$ one has

$$tu_* + \gamma(t) \neq 0.$$

8 Bifurcation Theory 2 - 16/03/2026

Prof of Theorem 7.12. Let $w(\lambda, t)$ be the map provided by the Lyapunov–Schmidt Lemma 7.8. Then a nearby solution of $F(\lambda, u) = 0$ must satisfy

$$u = tu_* + w(\lambda, t),$$

and the remaining equation is

$$QF(\lambda, tu_* + w(\lambda, t)) = 0.$$

Choose $\psi \in Y^* \setminus \{0\}$ such that $R = \ker \psi$, since Z is one-dimensional, the condition $Qy = 0$ is equivalent to $\psi(y) = 0$. Hence we define

$$\beta(\mu, t) := \big\langle \psi, F(\lambda_* + \mu,\ tu_* + w(\lambda_* + \mu, t)) \big\rangle.$$

Then β is of class C^2, and solving $F(\lambda, u) = 0$ near $(\lambda_*, 0)$ is equivalent to solving $\beta(\mu, t) = 0$ near $(0, 0)$. We claim the following:

> Near $(0, 0)$ we have $\beta(\mu, t) \simeq at^2 + bt\mu$ and $b \neq 0$.

Assuming the claim it is indeed possible to write $\mu = \mu(t)$ from $\beta(\mu, t) = 0$ near $(0, 0)$. We proceed by expanding in Taylor the function β up to second order terms obtaining

$$\beta(\mu, t) = \beta(0, 0) + \partial_t\beta(0, 0)t + \partial_\mu\beta(0, 0)\mu \tag{8.1}$$

$$+ \frac{1}{2}\partial_{tt}\beta(0, 0)t^2 + \partial_{t\mu}\beta(0, 0)t\mu + \frac{1}{2}\partial_{\mu\mu}\beta(0, 0)\mu^2 + R(t, \mu), \tag{8.2}$$

where $R(t, \mu) = o(t^2 + \mu^2)$.

- $\beta(0, 0) = 0$, since $w(\lambda_*, 0) = 0$, $F(\lambda_*, 0) = 0$, and

$$\beta(0, 0) = \langle \psi, F(\lambda_*, 0)\rangle = 0; \tag{8.3}$$

- Because $F(\lambda, 0) = 0$ and $w(\lambda, 0) = 0$ for all λ near 0, we have $\beta(\mu, 0) = 0$, and hence we deduce

$$\partial_\mu\beta(0, 0) = 0, \quad \partial_{\mu\mu}\beta(0, 0) = 0; \tag{8.4}$$

- For the time derivatives we have

$$\partial_t\beta(0, 0) = \langle \psi, \partial_u F(\lambda_*, 0)[u_* + \partial_t w(\lambda_*, 0)]\rangle = 0, \tag{8.5}$$

  since $u_* \in \ker(\partial_u F(\lambda_*, 0))$ and $\partial_t w(\lambda_*, 0) = 0$;

- For the mixed derivative we compute

$$\partial_t\beta(\mu, t) = \big\langle \psi, \partial_u F(\lambda_* + \mu, tu_* + w(\lambda_* + \mu, t))\big[u_* + \partial_t w(\lambda_* + \mu, t)\big]\big\rangle.$$

  Differentiating with respect to μ, we obtain

$$\begin{aligned}\partial^2_{\mu t}\beta(\mu, t) = &\big\langle \psi, \partial_{\lambda u} F(\lambda_* + \mu, tu_* + w(\lambda_* + \mu, t))\big[u_* + \partial_t w(\lambda_* + \mu, t)\big]\big\rangle \\ &+ \big\langle \psi, \partial_{uu} F(\lambda_* + \mu, tu_* + w(\lambda_* + \mu, t))\big[\partial_\lambda w(\lambda_* + \mu, t),\ u_* + \partial_t w(\lambda_* + \mu, t)\big]\big\rangle \\ &+ \big\langle \psi, \partial_u F(\lambda_* + \mu, tu_* + w(\lambda_* + \mu, t))\big[\partial^2_{\lambda t} w(\lambda_* + \mu, t)\big]\big\rangle.\end{aligned}$$

  In particular, at $(\mu, t) = (0, 0)$ one gets

$$\begin{aligned}\partial^2_{\mu t}\beta(0, 0) = &\big\langle \psi, \partial_{\lambda u} F(\lambda_*, w(\lambda_*, 0))\big[u_* + \partial_t w(\lambda_*, 0)\big]\big\rangle \\ &+ \big\langle \psi, \partial_{uu} F(\lambda_*, w(\lambda_*, 0))\big[\partial_\lambda w(\lambda_*, 0),\ u_* + \partial_t w(\lambda_*, 0)\big]\big\rangle \\ &+ \big\langle \psi, \partial_u F(\lambda_*, w(\lambda_*, 0))\big[\partial^2_{\lambda t} w(\lambda_*, 0)\big]\big\rangle.\end{aligned}$$

Since

$$\partial_t w(\lambda_*, 0) = 0, \qquad \partial_\lambda w(\lambda_*, 0) = 0, \qquad \partial^2_{\lambda t} w(\lambda_*, 0) \in W$$

the formula simplifies to

$$\partial^2_{\mu t}\beta(0,0) = \big\langle \psi, \partial_{\lambda u} F(\lambda_*, 0)[u_*] \big\rangle \neq 0,$$

by transversality assumption.

To conclude, define

$$h(\mu, t) := \begin{cases} \dfrac{\beta(\mu, t)}{t}, & t \neq 0, \\ \partial_t \beta(\mu, 0), & t = 0. \end{cases}$$

Then $h \in C^1$ and for $t \neq 0$

$$\beta(\mu, t) = 0 \quad \iff \quad h(\mu, t) = 0.$$

Moreover, $h(0,0) = 0$ and

$$\partial_\mu h(0,0) = \partial^2_{\mu t}\beta(0,0) \neq 0.$$

Therefore the implicit function theorem applies to the equation $h(\mu, t) = 0$ at $(0,0)$. Hence there exist $\varepsilon > 0$ and a unique map $\mu \in C^1((-\varepsilon, \varepsilon), \mathbb{R})$ such that

$$\mu(0) = 0, \qquad h(\mu(t), t) = 0.$$

For $t \neq 0$ this implies $\beta(\mu(t), t) = 0$, thus

$$F\big(\lambda_* + \mu(t),\, tu_* + w(\lambda_* + \mu(t), t)\big) = 0.$$

Set

$$\gamma(t) := w(\lambda_* + \mu(t), t) \in W.$$

Then

$$\gamma(0) = w(\lambda_*, 0) = 0$$

and

$$\gamma'(0) = \partial_\lambda w(\lambda_*, 0)\mu'(0) + \partial_t w(\lambda_*, 0) = 0,$$

because $w(\lambda, 0) \equiv 0$ and $\partial_t w(\lambda_*, 0) = 0$. Finally, for $t \neq 0$,

$$tu_* + \gamma(t) \neq 0.$$

Indeed, $tu_* \in V$ and $\gamma(t) \in W$, while $X = V \oplus W$ is a direct sum. Hence

$$tu_* + \gamma(t) = 0 \quad \Longrightarrow \quad tu_* = 0 \quad \Longrightarrow \quad t = 0,$$

a contradiction.

The uniqueness of the local curve follows from the uniqueness in the two applications of the implicit function theorem. □

Exercise 8.1. Prove the above mentioned properties of the function $h(\mu, t)$.

Remark 8.2. The proof gives more than existence of a branch: it also shows that every nearby nontrivial solution lies on that branch. Thus the local set of nontrivial solutions is exhausted by a single C^1 curve.

Remark 8.3. The mixed second derivative

$$\partial^2_{u\lambda} F(\lambda_*, 0) \in \mathcal{L}(X, Y)$$

can be interpreted as the derivative, with respect to λ, of the family

$$\lambda \mapsto \partial_u F(\lambda, 0) \in \mathcal{L}(X, Y).$$

More precisely,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\partial_u F(\lambda, 0) = \partial^2_{u\lambda} F(\lambda, 0).$$

This is useful in applications where the linearized operator depends explicitly on the parameter.

Remark 8.4. Differentiating the relation $h(\mu(t), t) = 0$ at $t = 0$ gives

$$\mu'(0) = -\frac{\partial_t h(0,0)}{\partial_\mu h(0,0)} = -\frac{1}{2}\frac{\langle \psi, \partial_{uu} F(\lambda_*, 0)[u_*, u_*]\rangle}{\langle \psi, \partial^2_{u\lambda} F(\lambda_*, 0)[u_*]\rangle}.$$

In particular, if

$$\langle \psi, \partial_{uu} F(\lambda_*, 0)[u_*, u_*]\rangle \neq 0,$$

then the bifurcating branch is not tangent to the vertical line $\lambda = \lambda_*$.

Exercise 8.5. Assume in addition that $F \in C^3$ and that

$$\langle \psi, \partial_{uu} F(\lambda_*, 0)[u_*, u_*]\rangle = 0.$$

Show that $\mu'(0) = 0$ and

$$\mu''(0) = -\frac{1}{3a}\left\langle \psi, d^3_{uuu} F(\lambda_*, 0)[u_*, u_*, u_*]\right\rangle, \qquad a = \left\langle \psi, \partial^2_{u\lambda} F(\lambda_*, 0)[u_*]\right\rangle. \tag{8.6}$$

Example 8.6. Consider $F : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ defined as

$$F(\lambda, (x, y)) = (\lambda x - y - x^3, \lambda y + x^2).$$

Then $F(\lambda, (0, 0)) = (0, 0)$ for all $\lambda \in \mathbb{R}$, and

$$\partial_u F(\lambda, (0,0))[(h,k)] = (\lambda h - k, \lambda k) \Rightarrow \partial_u F(\lambda, (0,0)) = \begin{pmatrix} \lambda & -1 \\ 0 & \lambda \end{pmatrix}.$$

Therefore the only possible bifurcation point is $\lambda = 0$.

At $\lambda = 0$ one has

$$\partial_u F(0, (0,0)) = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix},$$

so

$$\ker \partial_u F(0, (0,0)) = \operatorname{span}\{(1,0)\}, \qquad \operatorname{Ran} \partial_u F(0, (0,0)) = \operatorname{span}\{(1,0)\}.$$

Thus assumptions (1) and (2) are satisfied with $u_* = (1, 0)$. However,

$$\partial^2_{u\lambda} F(0, (0,0))[u_*] = (1, 0) \in \operatorname{Ran} \partial_u F(0, (0,0)),$$

so the transversality condition fails and the Crandall-Rabinowitz theorem cannot be applied.

In fact, $\lambda = 0$ is not a bifurcation point. Indeed,

$$F(0, (x, y)) = (-y - x^3, x^2),$$

and therefore

$$F(0, (x, y)) = 0 \quad \Longrightarrow \quad x = 0,\ y = 0.$$

Hence there are no nontrivial solutions near $(0, (0, 0))$.

Corollary 8.7. *Let $G \in C^2(X, X)$ with $G(0) = 0$, and consider*

$$F(\lambda, u) = \lambda u - G(u).$$

Assume that for some $\lambda_* \in \mathbb{R}$:

1. *The differential with respect to u of F at $(\lambda_*, 0)$ has a one dimensional kernel, namely*

$$\dim \ker(\lambda_* I - dG(0)) = 1;$$

2. *The differential with respect to u of F at $(\lambda_*, 0)$ has a closed and codimension one range, namely*

$$\operatorname{Ran}(\lambda_* I - dG(0))$$

3. *The differential with respect to u of F at $(\lambda_*, 0)$ has transversal range and kernel, namely*

$$\ker(\lambda_* I - dG(0)) \cap \operatorname{Ran}(\lambda_* I - dG(0)) = \{0\}.$$

Then λ_* is a bifurcation point.

Proof. We have

$$\partial_u F(\lambda, 0) = \lambda I - dG(0),$$

so the first two assumptions are exactly the hypotheses of the Crandall–Rabinowitz theorem.

Moreover,

$$\partial^2_{u\lambda} F(\lambda, 0)[v] = v \qquad \forall v \in X.$$

Hence, if u_* generates

$$\ker(\lambda_* I - dG(0)),$$

the transversality condition becomes

$$u_* \notin \operatorname{Ran}(\lambda_* I - dG(0)).$$

But this is equivalent to

$$\ker(\lambda_* I - dG(0)) \cap \operatorname{Ran}(\lambda_* I - dG(0)) = \{0\},$$

because the kernel is one-dimensional and generated by $u_* \neq 0$. □

Exercise 8.8. Consider the semilinear Dirichlet problem

$$-u'' + \lambda u = g(u), \qquad u(0) = u(1) = 0,$$

where $g \in C^2(\mathbb{R}, \mathbb{R})$ satisfies

$$g(0) = g'(0) = 0.$$

Show that the trivial branch $u \equiv 0$ exists for every $\lambda \in \mathbb{R}$, and determine all values of λ at which bifurcation from the trivial branch occurs.

Guided hints. Define

$$X := C^2_0([0,1]), \qquad Y := C([0,1]), \qquad F(\lambda, u) := -u'' + \lambda u - g(u).$$

1. Show that $F(\lambda, 0) = 0$ for every $\lambda \in \mathbb{R}$.
2. Compute the linearized operator $\partial_u F(\lambda, 0)$.
3. Determine the values of λ for which $\ker \partial_u F(\lambda, 0) \neq \{0\}$.
4. Show that for such values the kernel is one-dimensional and the range has codimension 1.
5. Verify the transversality condition.
6. Conclude that the corresponding values of λ are bifurcation points from the trivial branch.

□

8.1 Time-periodic solutions for second-order systems

Consider the autonomous system

$$\partial_t^2 u = f(u), \qquad u(t) \in \mathbb{R}^n, \qquad f \in C^2(\mathbb{R}^n, \mathbb{R}^n), \qquad f(0) = 0.$$

Set

$$A := df(0) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n),$$

and assume that A has negative eigenvalues

$$-\omega_1^2, \dots, -\omega_n^2, \qquad 0 < \omega_1 < \cdots < \omega_n.$$

Fix an index $j \in \{1, \dots, n\}$ and assume the nonresonance condition

$$\frac{\omega_s}{\omega_j} \notin \mathbb{N} \qquad \text{for every } s \neq j.$$

Remark 8.9. Since the numbers $-\omega_1^2, \dots, -\omega_n^2$ are n distinct real eigenvalues of $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, the operator A is diagonalizable over $\mathbb{R}$. Hence there exists an invertible matrix P such that

$$P^{-1}AP = \operatorname{diag}(-\omega_1^2, \dots, -\omega_n^2).$$

If we perform the linear change of variables

$$u = Pv,$$

then u solves

$$\partial_t^2 u = f(u)$$

if and only if v solves

$$\partial_t^2 v = g(v), \qquad g(v) := P^{-1} f(Pv).$$

Moreover,

$$g \in C^2(\mathbb{R}^n, \mathbb{R}^n), \qquad g(0) = 0, \qquad dg(0) = P^{-1}AP = \operatorname{diag}(-\omega_1^2, \dots, -\omega_n^2).$$

Therefore the existence of small periodic solutions is invariant under this linear change of coordinates. Hence, without loss of generality, we may assume from the start that

$$A = \operatorname{diag}(-\omega_1^2, \dots, -\omega_n^2).$$

Theorem 8.10. *Under the assumptions above, there exists a branch of nontrivial periodic solutions of*

$$\partial_t^2 u = f(u)$$

whose amplitudes tend to 0 *and whose frequencies tend to* ω_j.

More precisely, there exist $\varepsilon > 0$, *a* C^1 *function*

$$\omega : (-\varepsilon, \varepsilon) \to \mathbb{R}$$

and a C^1 *map*

$$u : (-\varepsilon, \varepsilon) \to C^2_{2\pi}(\mathbb{R}, \mathbb{R}^n)$$

such that

$$u(0) = 0, \qquad u(s) \neq 0 \text{ for } s \neq 0, \qquad \omega(0) = \omega_j,$$

and

$$\omega(s)^2\, u(s)''(\tau) = f(u(s)(\tau)) \qquad \forall \tau \in \mathbb{R}.$$

Consequently,

$$\tilde{u}_s(t) := u(s)(\omega(s)t)$$

is a nontrivial periodic solution of $\partial_t^2 u = f(u)$, *with frequency* $\omega(s) \to \omega_j$ *as* $s \to 0$.

Proof. We apply the Crandall–Rabinowitz theorem to the rescaled equation. Set

$$X := \{u \in C^2_{2\pi}(\mathbb{R}, \mathbb{R}^n) : u(-\tau) = u(\tau)\}, \qquad Y := \{v \in C^0_{2\pi}(\mathbb{R}, \mathbb{R}^n) : v(-\tau) = v(\tau)\}.$$

We work in the even subspaces in order to remove the time-translation degeneracy: on the full space of 2π-periodic functions the kernel would contain both $\cos\tau$ and $\sin\tau$ modes. Define

$$F : \mathbb{R} \times X \to Y, \qquad F(\omega, u) := \omega^2 u'' - f(u).$$

Since $f \in C^2(\mathbb{R}^n, \mathbb{R}^n)$, we have

$$F \in C^2(\mathbb{R} \times X, Y), \qquad F(\omega, 0) = 0 \quad \forall \omega \in \mathbb{R}.$$

We study the linearization at $(\omega_j, 0)$:

$$L := \partial_u F(\omega_j, 0) : X \to Y, \qquad Lh = \omega_j^2 h'' - Ah.$$

Since $A = \operatorname{diag}(-\omega_1^2, \dots, -\omega_n^2)$, writing $h = (h_1, \dots, h_n)$, we obtain componentwise

$$(Lh)_r = \omega_j^2 h_r'' + \omega_r^2 h_r, \qquad r = 1, \dots, n.$$

We now check the three hypotheses of the Crandall-Rabinowitz Theorem 7.12

1. *Computation of the kernel*: Each component h_r is an even 2π-periodic function, hence it has a cosine expansion

   $$h_r(\tau) = \sum_{k=0}^{\infty} h_{r,k} \cos(k\tau), \qquad h_{r,k} \in \mathbb{R}.$$

   Therefore

   $$(Lh)_r(\tau) = \sum_{k=0}^{\infty} (\omega_r^2 - k^2\omega_j^2)\, h_{r,k} \cos(k\tau).$$

   Thus $Lh = 0$ is equivalent to

   $$(\omega_r^2 - k^2\omega_j^2)\, h_{r,k} = 0 \qquad \forall r = 1, \dots, n,\ \forall k \geq 0.$$

   We now determine when the coefficient $\omega_r^2 - k^2\omega_j^2$ can vanish.

   (a) If $k = 0$, then

   $$\omega_r^2 - k^2\omega_j^2 = \omega_r^2 > 0,$$

   so all constant modes vanish.

   (b) If $k = 1$, then

   $$\omega_r^2 - \omega_j^2 = 0$$

   if and only if $r = j$, because

   $$0 < \omega_1 < \cdots < \omega_n.$$

   (c) If $k \geq 2$, then

   $$\omega_r^2 - k^2\omega_j^2 = 0 \quad \Longleftrightarrow \quad \omega_r = k\omega_j.$$

   For $r \neq j$ this is excluded by the nonresonance assumption, and for $r = j$ it is impossible since $k \geq 2$.

   Hence the only vanishing coefficient occurs for $(r, k) = (j, 1)$, therefore

   $$\ker L = \operatorname{span}\{e_j \cos\tau\},$$

   where e_j is the j-th vector of the canonical basis of $\mathbb{R}^n$. In particular, $\dim\ker L = 1$.

2. *Description of the range*:

   Let $v = (v_1, \dots, v_n) \in Y$, with cosine expansion

$$v_r(\tau) = \sum_{k=0}^{\infty} v_{r,k} \cos(k\tau).$$

   The equation $Lh = v$ is equivalent to the family of scalar equations

$$(\omega_r^2 - k^2\omega_j^2)\, h_{r,k} = v_{r,k} \qquad \forall r, \forall k.$$

   Since the coefficient is nonzero for every pair (r, k) except $(j, 1)$, this system is solvable if and only if $v_{j,1} = 0$. Thus

$$\operatorname{Ran} L = \{v \in Y : \ v_{j,1} = 0\}.$$

   This is a closed subspace of codimension 1.

3. *Transversality*: Now set

$$u_* := e_j \cos\tau.$$

   Note that $u_* \in \ker L$, but $u_* \notin \operatorname{Ran} L$ because its $(j, 1)$-Fourier coefficient is equal to 1. Hence

$$\ker L \cap \operatorname{Ran} L = \{0\}.$$

   We compute

$$\partial_u F(\omega, u)[h] = \omega^2 h'' - df(u)[h],$$

   hence

$$d^2_{u\omega} F(\omega_j, 0)[h] = 2\omega_j h''.$$

   Applying this to the generator $u_* = e_j \cos\tau$ of $\ker L$, we get

$$d^2_{u\omega} F(\omega_j, 0)[u_*] = 2\omega_j (e_j \cos\tau)'' = -2\omega_j e_j \cos\tau = -2\omega_j u_*.$$

   Since $u_* \notin \operatorname{Ran} L$, we conclude that

$$d^2_{u\omega} F(\omega_j, 0)[u_*] \notin \operatorname{Ran} L.$$

   Thus the Crandall–Rabinowitz transversality condition holds.

All hypotheses of the Crandall–Rabinowitz theorem are satisfied at $(\omega_j, 0)$. Hence there exist $\varepsilon > 0$, a C^1 function

$$\omega : (-\varepsilon, \varepsilon) \to \mathbb{R}, \qquad \omega(0) = \omega_j,$$

and a C^1 map

$$u : (-\varepsilon, \varepsilon) \to X$$

such that

$$u(0) = 0, \qquad u(s) \neq 0 \text{ for } s \neq 0, \qquad F(\omega(s), u(s)) = 0$$

for every $s \in (-\varepsilon, \varepsilon)$. That is,

$$\omega(s)^2\, u(s)''(\tau) = f(u(s)(\tau)) \qquad \forall \tau \in \mathbb{R}.$$

Finally, define

$$\tilde{u}_s(t) := u(s)(\omega(s)t).$$

Then, by the chain rule,

$$\partial_t^2 \tilde{u}_s(t) = \omega(s)^2 u(s)''(\omega(s)t) = f(u(s)(\omega(s)t)) = f(\tilde{u}_s(t)).$$

Hence $\tilde{u}_s$ is a solution of

$$\partial_t^2 u = f(u).$$

Since $u(s)$ is 2π-periodic in τ, the function $\tilde{u}_s$ is periodic in t with period

$$T(s) = \frac{2\pi}{\omega(s)}.$$

Therefore its frequency is $\omega(s)$, and

$$\omega(s) \to \omega_j, \qquad u(s) \to 0 \text{ in } C^2_{2\pi}(\mathbb{R}, \mathbb{R}^n)$$

as $s \to 0$. This proves the theorem. □

9 Water waves 1 - 17/03/2026

We start discussing a very important application of the Crandall–Rabinowitz Theorem to the water wave equations. The main goal of this section is to mathematically prove that waves that look like the ones in the (beautiful) picture below exists.

Figure 1: Periodic travelling waves

To do so, we introduce the key mathematical concepts useful to formulate the problem, and we build up the setting toward the main theorem of this section.

9.1 The water wave system

We consider a homogeneous inviscid fluid of constant density $\rho \equiv 1$, and denote by $u = u(t, x)$ the velocity of the fluid, by $p = p(t, x)$ the pressure, and by W the external potential.

Euler equation and incompressibility. For an inviscid fluid of constant density $\rho \equiv 1$, moving in a time-dependent domain $\Omega(t) \subset \mathbb{R}^d$ and subject to a conservative body force $-\nabla W$, the Euler equations read

$$\partial_t u + (u \cdot \nabla)u = -\nabla p - \nabla W \qquad \text{in } \Omega(t), \tag{9.1}$$

where $u = u(t, x)$ is the velocity field and $p = p(t, x)$ is the pressure. The incompressibility constraint comes from conservation of mass:

$$\partial_t \rho + \nabla \cdot (\rho u) = 0. \tag{9.2}$$

Since here $\rho \equiv 1$, this reduces to

$$\nabla \cdot u = 0. \tag{9.3}$$

Equivalently, the flow preserves volume. Indeed, if Φ_t denotes the flow map generated by u, then for every material set $E \subset \Omega(0)$ one has

$$\frac{d}{dt}|\Phi_t(E)| = \int_{\Phi_t(E)} \nabla \cdot u(t, x)\, \mathrm{d}x, \tag{9.4}$$

so $\nabla \cdot u = 0$ is precisely the infinitesimal form of volume preservation.

Vorticity formulation. We introduce the vorticity

$$\omega := \nabla \times u. \tag{9.5}$$

Taking the curl of the Euler equation and using $\nabla \times \nabla p = \nabla \times \nabla W = 0$, we obtain

$$\partial_t \omega + \nabla \times \big((u \cdot \nabla)u\big) = 0. \tag{9.6}$$

Using the vector identity

$$(u \cdot \nabla)u = \nabla\left(\frac{1}{2}|u|^2\right) - u \times \omega, \tag{9.7}$$

it follows that

$$\partial_t \omega - \nabla \times (u \times \omega) = 0. \tag{9.8}$$

Since $\nabla \cdot u = 0$ and $\nabla \cdot \omega = 0$, this becomes

$$\partial_t \omega + (u \cdot \nabla)\omega - (\omega \cdot \nabla)u = 0. \tag{9.9}$$

This is the vorticity equation for Euler.

Exercise 9.1. Assume the Euler solution is sufficiently smooth. Prove that, if $\omega(0, \cdot) = 0$, then $\omega(t, \cdot) = 0$ for all later times.

Hints. Let $X(t; a)$ be the flow map associated with u, that is,

$$\begin{cases} \dot{X}(t; a) = u(t, X(t; a)), \\ X(0; a) = a, \end{cases} \tag{9.10}$$

and compute

$$\frac{d}{dt}\omega(t, X(t; a)). \tag{9.11}$$

□

Irrotational flows and velocity potential. An irrotational flow is characterized by

$$\omega = \nabla \times u = 0. \tag{9.12}$$

If the fluid domain is simply connected [1] there is a scalar potential ϕ such that

$$u = \nabla \phi. \tag{9.13}$$

Combining this with incompressibility gives

$$\Delta \phi = \nabla \cdot u = 0 \qquad \text{in } \Omega(t). \tag{9.14}$$

Substituting $u = \nabla \phi$ into Euler yields

$$\partial_t \nabla \phi + (\nabla \phi \cdot \nabla)\nabla \phi = -\nabla(p + W). \tag{9.15}$$

Since

$$(\nabla \phi \cdot \nabla)\nabla \phi = \nabla\left(\frac{1}{2}|\nabla \phi|^2\right), \tag{9.16}$$

we obtain

$$\nabla\left(\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + p + W\right) = 0. \tag{9.17}$$

Hence

$$\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + p + W = c(t) \tag{9.18}$$

for some function $c(t)$ depending only on time. Since adding a function of time to ϕ does not change $\nabla \phi$, one can absorb $c(t)$ into the definition of ϕ and write the Bernoulli equation in the form

$$\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + p + W = 0. \tag{9.19}$$

For gravity waves, with the vertical variable y pointing upward, one typically takes

$$W(x, y) = gy. \tag{9.20}$$

[1] or, in the periodic water-wave setting, after imposing zero circulation so that a global potential exists

The water-wave geometry. We now restrict to a two-dimensional periodic setting, with horizontal variable $x \in \mathbb{T}$ and vertical variable $y \in \mathbb{R}$. We denote by $h > 0$ the mean depth, by $\beta = \beta(x)$ the bottom topography, and by $\eta = \eta(t, x)$ the free surface. The fluid occupies the time-dependent domain

$$\Omega(t) = \big\{(x, y) \in \mathbb{T} \times \mathbb{R} : \, -h + \beta(x) < y < \eta(t, x)\big\}. \tag{9.21}$$

Thus the lower boundary is

$$\Gamma_b = \big\{(x, y) : y = -h + \beta(x)\big\}, \tag{9.22}$$

while the free surface is

$$\Sigma(t) = \big\{(x, y) : y = \eta(t, x)\big\}. \tag{9.23}$$

At the bottom one imposes impermeability, namely

$$u \cdot n_b = 0 \qquad \text{on } \Gamma_b, \tag{9.24}$$

where n_b is any normal vector to the bottom.

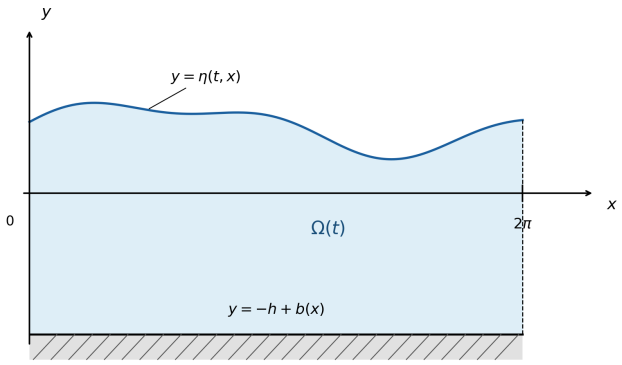

Figure 2: Approximate geometry of a section of the sea

Boundary conditions and the water-wave system. From now on we assume a flat bottom, that is $\beta \equiv 0$, so that the fluid domain is

$$\Omega(t) = \{(x, y) \in \mathbb{T} \times \mathbb{R} : \, -h < y < \eta(t, x)\}, \tag{9.25}$$

where $\eta = \eta(t, x)$ describes the free surface. Since the flow is assumed irrotational and incompressible, there exists a velocity potential ϕ such that

$$u = \nabla \phi,$$

and ϕ is harmonic in the fluid region:

$$\Delta \phi = 0 \qquad \text{in } \Omega(t). \tag{9.26}$$

At the flat bottom $y = -h$ impermeability reads

$$\partial_y \phi(t, x, -h) = 0. \tag{9.27}$$

On the free surface $\Sigma(t) = \{y = \eta(t, x)\}$ one has two boundary conditions. The first one is the *kinematic boundary condition*: the free surface moves with the fluid. Indeed, if a particle on the surface is written as

$$\big(x(t), y(t)\big) = \big(x(t), \eta(t, x(t))\big),$$

then, since its velocity is $\nabla \phi = (\phi_x, \phi_y)$, we obtain

$$\dot{y}(t) = \partial_t \eta(t, x(t)) + \partial_x \eta(t, x(t))\, \dot{x}(t) = \partial_t \eta + \eta_x \phi_x,$$

while also $\dot{y}(t) = \phi_y$. Therefore

$$\partial_t \eta = \phi_y - \eta_x \phi_x \qquad \text{on } y = \eta(t, x). \tag{9.28}$$

The second one is the *dynamic boundary condition*. If the pressure on the free surface is constant and normalized to zero, then Bernoulli's law with gravity gives

$$\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + g\eta = 0 \qquad \text{on } y = \eta(t,x). \tag{9.29}$$

If surface tension with coefficient $\kappa \geq 0$ is included, this becomes

$$\partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + g\eta = \kappa H(\eta) \qquad \text{on } y = \eta(t,x), \tag{9.30}$$

where, with the present sign convention,

$$H(\eta) = \partial_x \left(\frac{\eta_x}{\sqrt{1+\eta_x^2}} \right) = \frac{\eta_{xx}}{(1+\eta_x^2)^{3/2}} \tag{9.31}$$

is the curvature of the graph $y = \eta(x)$.

Collecting all the equations, the water-wave system in the flat-bottom irrotational setting reads

$$\begin{cases} \Delta \phi = 0 & \text{in } \Omega(t), \\ \partial_y \phi = 0 & \text{on } y = -h, \\ \partial_t \eta = \phi_y - \eta_x \phi_x & \text{on } y = \eta(t,x), \\ \partial_t \phi + \frac{1}{2}|\nabla \phi|^2 + g\eta = \kappa H(\eta) & \text{on } y = \eta(t,x). \end{cases} \tag{9.32}$$

9.2 Zakharov variables and Dirichlet–Neumann operator

We introduce the surface trace of the velocity potential:

$$\psi(t,x) := \phi\big(t,x,\eta(t,x)\big). \tag{9.33}$$

For a given surface profile η and boundary datum ψ, let

$$\phi = \phi[\eta,\psi]$$

denote the unique solution of the elliptic boundary-value problem

$$\begin{cases} \Delta \phi = 0 & \text{in } \Omega_\eta, \\ \phi = \psi & \text{on } y = \eta(x), \\ \phi_y = 0 & \text{on } y = -h, \end{cases} \tag{9.34}$$

where

$$\Omega_\eta := \{(x,y) \in \mathbb{T} \times \mathbb{R} : \, -h < y < \eta(x)\}. \tag{9.35}$$

Definition 9.2 (Dirichlet–Neumann operator)**.** For η fixed, the *Dirichlet–Neumann operator* is defined by

$$G(\eta)[\psi] := \sqrt{1+\eta_x^2}\, \partial_n \phi\big|_{y=\eta(x)} = \big(\phi_y - \eta_x \phi_x\big)\big|_{y=\eta(x)}, \tag{9.36}$$

where $\phi = \phi[\eta,\psi]$ is the solution of (9.34) and

$$n = \frac{(-\eta_x, 1)}{\sqrt{1+\eta_x^2}} \tag{9.37}$$

is the upward unit normal to the free surface.

Remark 9.3. The function ψ is the Dirichlet datum of the potential on the free boundary. Once the harmonic extension problem (9.34) is solved, the operator $G(\eta)$ returns the corresponding normal derivative on the free surface, that is, the Neumann datum.

Thus $G(\eta)$ encodes the bulk Laplace problem into an operator acting only on the boundary variables (η,ψ).

Proposition 9.4 (Computation of $G(0)$)**.** *For the flat surface $\eta = 0$, the Dirichlet–Neumann operator is the Fourier multiplier*

$$G(0) = |D| \tanh(h|D|). \tag{9.38}$$

Equivalently, if

$$\psi(x) = \sum_{j\in\mathbb{Z}} \hat{\psi}_j e^{ijx},$$

then

$$G(0)[\psi] = \sum_{j\in\mathbb{Z}} |j| \tanh(h|j|)\, \hat{\psi}_j e^{ijx}. \tag{9.39}$$

Proof. When $\eta = 0$, the fluid occupies the strip $\Omega_0 = \{(x, y) \in \mathbb{T} \times \mathbb{R} : -h < y < 0\}$. Let ϕ solve

$$\begin{cases} \Delta\phi = 0 & \text{in } \Omega_0, \\ \phi(x, 0) = \psi(x), \\ \phi_y(x, -h) = 0. \end{cases}$$

Expand

$$\psi(x) = \sum_{j\in\mathbb{Z}} \hat{\psi}_j e^{ijx}, \qquad \phi(x, y) = \sum_{j\in\mathbb{Z}} \phi_j(y) e^{ijx}.$$

Then each Fourier mode satisfies

$$\phi_j'' - j^2\phi_j = 0.$$

If $j = 0$, then $\phi_0(y) = a_0 + b_0 y$. The bottom condition $\phi_y(-h) = 0$ gives $b_0 = 0$, while $\phi(x, 0) = \psi(x)$ yields $a_0 = \hat{\psi}_0$. If $j \neq 0$, then

$$\phi_j(y) = a_j \sinh(|j|(y + h)) + b_j \cosh(|j|(y + h)).$$

Differentiating,

$$\phi_j'(y) = |j|a_j \cosh(|j|(y + h)) + |j|b_j \sinh(|j|(y + h)).$$

The bottom condition gives

$$\phi_j'(-h) = |j|a_j = 0,$$

hence $a_j = 0$. Therefore

$$\phi_j(y) = b_j \cosh(|j|(y + h)).$$

Since $\phi_j(0) = \hat{\psi}_j$,

$$b_j \cosh(|j|h) = \hat{\psi}_j, \qquad b_j = \frac{\hat{\psi}_j}{\cosh(|j|h)}.$$

Thus

$$\phi_j'(0) = |j| \frac{\sinh(|j|h)}{\cosh(|j|h)} \hat{\psi}_j = |j| \tanh(|j|h)\hat{\psi}_j.$$

Since $G(0)[\psi] = \phi_y(\cdot, 0)$, we obtain (9.39), and therefore (9.38). □

We now rewrite the water-wave system in terms of the boundary variables (η, ψ). In the flat-bottom case, the bulk formulation is

$$\Delta\phi = 0 \qquad \text{in } \Omega(t), \tag{9.40a}$$

$$\phi_y = 0 \qquad \text{on } y = -h, \tag{9.40b}$$

$$\partial_t\eta = \phi_y - \eta_x\phi_x \qquad \text{on } y = \eta(t, x), \tag{9.40c}$$

$$\partial_t\phi + \frac{1}{2}|\nabla\phi|^2 + g\eta = \kappa H(\eta) \qquad \text{on } y = \eta(t, x). \tag{9.40d}$$

Here

$$H(\eta) = \partial_x \left(\frac{\eta_x}{\sqrt{1 + \eta_x^2}} \right)$$

is the curvature of the free surface.

By (9.36), the kinematic condition (9.40c) becomes

$$\partial_t \eta = G(\eta)[\psi]. \tag{9.41}$$

To derive the second equation, differentiate (9.33) with respect to x:

$$\psi_x = \phi_x + \eta_x \phi_y \qquad \text{on } y = \eta(t,x). \tag{9.42}$$

Together with

$$G(\eta)[\psi] = \phi_y - \eta_x \phi_x \qquad \text{on } y = \eta(t,x), \tag{9.43}$$

this yields the linear system

$$\begin{cases} \psi_x = \phi_x + \eta_x \phi_y, \\ G(\eta)[\psi] = -\eta_x \phi_x + \phi_y. \end{cases}$$

Solving for ϕ_x and ϕ_y, we obtain

$$\begin{cases} \phi_x = \dfrac{\psi_x - \eta_x G(\eta)[\psi]}{1+\eta_x^2}, \\ \phi_y = \dfrac{G(\eta)[\psi] + \eta_x \psi_x}{1+\eta_x^2}, \end{cases} \qquad \text{on } y = \eta(t,x). \tag{9.44}$$

Next, differentiate (9.33) with respect to t:

$$\partial_t \psi = \partial_t \phi + \phi_y \, \partial_t \eta \qquad \text{on } y = \eta(t,x). \tag{9.45}$$

Using (9.40d) and (9.41), we get

$$\begin{aligned} \partial_t \psi &= -g\eta - \frac{1}{2}\left(\phi_x^2 + \phi_y^2\right) + \kappa H(\eta) + \phi_y G(\eta)[\psi] \\ &= -g\eta - \frac{1}{2}\phi_x^2 + \frac{1}{2}\phi_y^2 - \eta_x \phi_x \phi_y + \kappa H(\eta). \end{aligned} \tag{9.46}$$

On the other hand, by (9.42),

$$\psi_x^2 = (\phi_x + \eta_x \phi_y)^2 = \phi_x^2 + 2\eta_x \phi_x \phi_y + \eta_x^2 \phi_y^2,$$

while, by (9.43),

$$G(\eta)[\psi] + \eta_x \psi_x = (\phi_y - \eta_x \phi_x) + \eta_x(\phi_x + \eta_x \phi_y) = (1+\eta_x^2)\phi_y.$$

Hence

$$\frac{\left(G(\eta)[\psi] + \eta_x \psi_x\right)^2}{1+\eta_x^2} = (1+\eta_x^2)\phi_y^2.$$

Therefore

$$-\frac{1}{2}\psi_x^2 + \frac{1}{2}\frac{\left(G(\eta)[\psi] + \eta_x \psi_x\right)^2}{1+\eta_x^2} = -\frac{1}{2}\phi_x^2 - \eta_x \phi_x \phi_y + \frac{1}{2}\phi_y^2.$$

Substituting this identity into (9.46), we obtain

$$\partial_t \psi = -g\eta - \frac{1}{2}\psi_x^2 + \frac{1}{2}\frac{\left(\eta_x \psi_x + G(\eta)[\psi]\right)^2}{1+\eta_x^2} + \kappa H(\eta). \tag{9.47}$$

Thus the water-wave problem can be written as the Zakharov system

$$\begin{cases} \partial_t \eta = G(\eta)[\psi], \\ \partial_t \psi = -g\eta - \dfrac{1}{2}\psi_x^2 + \dfrac{1}{2}\dfrac{\left(\eta_x \psi_x + G(\eta)[\psi]\right)^2}{1+\eta_x^2} + \kappa H(\eta). \end{cases} \tag{9.48}$$

10 Water waves 2 - 23/03/2026

Exercise 10.1 (Hamiltonian formulation)**.** Define the Hamiltonian

$$\mathcal{H}(\eta,\psi) := \frac{1}{2}\int_{\Omega_\eta} |\nabla\phi|^2 \,\mathrm{d}x\mathrm{d}y + \int_{\Omega_\eta} gy \,\mathrm{d}x\mathrm{d}y + \kappa \int_{\mathbb{T}} \sqrt{1+\eta_x^2}\,\mathrm{d}x, \tag{10.1}$$

where the first term is the kinetic energy, the second the gravitational potential energy, and the third the capillarity energy. Show that, up to irrelevant constants,

$$\mathcal{H}(\eta,\psi) = \frac{1}{2}\int_{\mathbb{T}} \psi\, G(\eta)[\psi]\,\mathrm{d}x + \frac{g}{2}\int_{\mathbb{T}} \eta^2\,\mathrm{d}x + \kappa\int_{\mathbb{T}} \sqrt{1+\eta_x^2}\,\mathrm{d}x, \tag{10.2}$$

and deduce that the Zakharov system is Hamiltonian, namely

$$\partial_t \eta = \nabla_\psi \mathcal{H}(\eta,\psi), \qquad \psi_t = -\nabla_\eta \mathcal{H}(\eta,\psi). \tag{10.3}$$

We continue the discussion by presenting general properties of the Dirichlet–Neumann operator, leaving the proof as an exercise.

Proposition 10.2 (Properties of the Dirichlet–Neumann operator)**.** *For the Dirichlet–Neumann operator the following properties hold.*

1. *for every fixed η, the map $\psi \mapsto G(\eta)[\psi]$ is linear;*
2. *$G(\eta)$ is symmetric, i.e.*
$$\langle \psi_1, G(\eta)[\psi_2]\rangle_{L^2(\mathbb{T})} = \langle G(\eta)[\psi_1], \psi_2\rangle_{L^2(\mathbb{T})}; \tag{10.4}$$
3. *$G(\eta) \geq 0$ and its kernel consists of the constant functions.*

Exercise 10.3. Prove the above proposition.

Remark 10.4. For fixed η, the operator $G(\eta)$ is a first-order pseudodifferential operator. It is linear in ψ, but nonlinear in η. On suitable Sobolev spaces, namely sufficiently high regularity in η, the map $\eta \mapsto G(\eta)$ is smooth.

10.1 Travelling waves

We now look for 2π-periodic travelling waves for the Zakharov system (9.48). More precisely, we seek solutions of the form

$$\eta(t,x) = \check{\eta}(x+ct), \qquad \psi(t,x) = \check{\psi}(x+ct), \tag{10.5}$$

where $c \in \mathbb{R}$ is the wave speed to be determined.

Dropping the checks and denoting again by x the travelling variable, we have

$$\partial_t \eta = c\eta_x, \qquad \psi_t = c\psi_x.$$

Therefore (9.48) becomes

$$\begin{cases} c\,\eta_x = G(\eta)[\psi], \\ c\,\psi_x = -g\eta - \dfrac{1}{2}\psi_x^2 + \dfrac{1}{2}\dfrac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1+\eta_x^2} + \kappa H(\eta). \end{cases} \tag{10.6}$$

To apply Crandall–Rabinowotz, it is convenient to rewrite the problem as $F(c,\eta,\psi)=0$, where

$$F(c,\eta,\psi) := \begin{pmatrix} F_1(c,\eta,\psi) \\ F_2(c,\eta,\psi) \end{pmatrix} := \begin{pmatrix} c\,\eta_x - G(\eta)[\psi] \\ c\,\psi_x + g\eta + \dfrac{1}{2}\psi_x^2 - \dfrac{1}{2}\dfrac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1+\eta_x^2} - \kappa H(\eta) \end{pmatrix}. \tag{10.7}$$

Observe that

$$F(c,0,0) = 0 \qquad \text{for every } c \in \mathbb{R}.$$

Remark 10.5 (Even–odd reduction)**.** For the bifurcation analysis it is natural to restrict to even surface profiles and odd traces of the velocity potential. Set

$$H^s_{\text{even}}(\mathbb{T}) := \{f \in H^s(\mathbb{T}) : f(-x) = f(x)\},$$
$$H^s_{\text{odd}}(\mathbb{T}) := \{f \in H^s(\mathbb{T}) : f(-x) = -f(x)\}.$$

We now check the parity of the two component of F.

Lemma 10.6. *If η is even and ψ is odd, then $G(\eta)[\psi]$ is odd.*

Proof. Assume $\eta(-x) = \eta(x)$ and $\psi(-x) = -\psi(x)$. Let ϕ solve

$$\begin{cases} \Delta\phi = 0 & \text{in } \Omega_\eta, \\ \phi = \psi & \text{on } y = \eta(x), \\ \phi_y = 0 & \text{on } y = -h. \end{cases}$$

Then the function $\widetilde{\phi}(x, y) := -\phi(-x, y)$ satisfies the same boundary value problem. By uniqueness,

$$\widetilde{\phi} = \phi, \qquad \text{hence} \qquad \phi(-x, y) = -\phi(x, y).$$

Differentiating gives

$$\phi_x(-x, y) = \phi_x(x, y), \qquad \phi_y(-x, y) = -\phi_y(x, y).$$

Since η is even, η_x is odd:

$$\eta_x(-x) = -\eta_x(x).$$

Therefore

$$\begin{aligned} G(\eta)[\psi](-x) &= \phi_y(-x, \eta(-x)) - \eta_x(-x)\phi_x(-x, \eta(-x)) \\ &= -\phi_y(x, \eta(x)) + \eta_x(x)\phi_x(x, \eta(x)) \\ &= -G(\eta)[\psi](x). \end{aligned}$$

So $G(\eta)[\psi]$ is odd. □

Proposition 10.7. *If η is even and ψ is odd, then $F_1(c, \eta, \psi)$ is odd and $F_2(c, \eta, \psi)$ is even.*

Proof. From Lemma 10.6, η being odd, and ψ being even we have that

$$c\eta_x - G(\eta)[\psi]$$

is odd, and

$$c\psi_x + g\eta + \frac{1}{2}\psi_x^2 - \frac{1}{2}\frac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1 + \eta_x^2} - \kappa H(\eta)$$

is even. Indeed, we have

$$\frac{\big(\eta_x(-x)\psi_x(-x) + G(\eta)[\psi](-x)\big)^2}{1 + \eta_x^2(-x)} = \frac{\big(-\eta_x(x)\psi_x(x) - G(\eta)[\psi](x)\big)^2}{1 + \eta_x^2(x)} = \frac{1}{2}\frac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1 + \eta_x^2} \tag{10.8}$$

and

$$\frac{\eta_x}{\sqrt{1 + \eta_x^2}} \tag{10.9}$$

is an odd function, hence its derivative $H(\eta)$ is even. □

Based on the above discussion we fix s sufficiently large, and work with

$$F : \mathbb{R} \times H^s_{\text{even}}(\mathbb{T}) \times H^s_{\text{odd}}(\mathbb{T}) \longrightarrow H^{s-1}_{\text{odd}}(\mathbb{T}) \times H^{s-2}_{\text{even}}(\mathbb{T}).$$

We lower regularity in the codomain since the function F lowers regularity due to the presence of derivatives.

Proposition 10.8 (Linearization at the trivial branch)**.** *Let $s > 5/2$. Then the differential of F with respect to (η, ψ) at $(c, 0, 0)$ is*

$$d_{(\eta,\psi)}F(c,0,0)[h,\varphi] = \begin{pmatrix} c\,h_x - G(0)[\varphi] \\ c\,\varphi_x + gh - \kappa h_{xx} \end{pmatrix}. \tag{10.10}$$

Proof. From (10.7),

$$F_1(c,\eta,\psi) = c\eta_x - G(\eta)[\psi].$$

At $(\eta, \psi) = (0, 0)$, the differential is simply

$$dF_1(c,0,0)[h,\varphi] = c\,h_x - G(0)[\varphi].$$

For the second component,

$$F_2(c,\eta,\psi) = c\psi_x + g\eta + \frac{1}{2}\psi_x^2 - \frac{1}{2}\frac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1+\eta_x^2} - \kappa H(\eta).$$

At $(\eta, \psi) = (0, 0)$, both nonlinear quadratic terms have zero differential, because they vanish to second order. Thus the only nontrivial point is the curvature term. Since

$$H(\eta) = \partial_x\left(\frac{\eta_x}{\sqrt{1+\eta_x^2}}\right),$$

we have

$$dH(0)[h] = \partial_x\big(h_x\big) = h_{xx}.$$

Hence

$$dF_2(c,0,0)[h,\varphi] = c\,\varphi_x + gh - \kappa h_{xx}.$$

This proves (10.10). □

Proposition 10.9 (Mode-by-mode analysis of the linearized operator)**.** *Let $(h, \varphi) \in H^s_{\text{even}}(\mathbb{T}) \times H^s_{\text{odd}}(\mathbb{T})$, and $(f, g) \in H^{s-1}_{\text{odd}}(\mathbb{T}) \times H^{s-2}_{\text{even}}(\mathbb{T})$, and consider*

$$d_{(\eta,\psi)}F(c,0,0)[h,\varphi] = (f,g). \tag{10.11}$$

Write the Fourier expansions

$$h(x) = h_0 + \sum_{n\geq 1} h_n\cos(nx), \quad \varphi(x) = \sum_{n\geq 1}\varphi_n\sin(nx),$$
$$f(x) = \sum_{n\geq 1} f_n\sin(nx), \quad g(x) = g_0 + \sum_{n\geq 1} g_n\cos(nx).$$

Then (10.11) is equivalent to

$$g\,h_0 = g_0, \tag{10.12}$$

$$\begin{pmatrix} -cn & -n\tanh(nh) \\ g+\kappa n^2 & cn \end{pmatrix}\begin{pmatrix} h_n \\ \varphi_n \end{pmatrix} = \begin{pmatrix} f_n \\ g_n \end{pmatrix}, \qquad n \geq 1. \tag{10.13}$$

Proof. From (10.10),

$$d_{(\eta,\psi)}F(c,0,0)[h,\varphi] = \begin{pmatrix} c\,h_x - G(0)[\varphi] \\ c\,\varphi_x + gh - \kappa h_{xx} \end{pmatrix}.$$

Since

$$h_x = -\sum_{n\geq 1} nh_n\sin(nx), \qquad \varphi_x = \sum_{n\geq 1} n\varphi_n\cos(nx),$$

and

$$G(0)[\varphi] = \sum_{n\geq 1} n\tanh(nh)\,\varphi_n\sin(nx),$$

the first component becomes

$$c\,h_x - G(0)[\varphi] = \sum_{n\geq 1}\big(-cn\,h_n - n\tanh(nh)\varphi_n\big)\sin(nx).$$

Thus

$$-cn\,h_n - n\tanh(nh)\varphi_n = f_n \qquad (n \geq 1).$$

For the second component,

$$-\kappa h_{xx} = -\kappa\sum_{n\geq 1}(-n^2 h_n)\cos(nx) = \kappa\sum_{n\geq 1} n^2 h_n \cos(nx).$$

Therefore

$$c\,\varphi_x + gh - \kappa h_{xx} = gh_0 + \sum_{n\geq 1}\big((g+\kappa n^2)h_n + cn\,\varphi_n\big)\cos(nx).$$

Hence

$$gh_0 = g_0, \qquad (g+\kappa n^2)h_n + cn\,\varphi_n = g_n \quad (n \geq 1).$$

This is exactly (10.12)–(10.13). □

To understand possible bifurcation values for c we investigate when the matrices above are singular.

Proposition 10.10 (Dispersion relation and possible bifurcation speeds)**.** *For $n \geq 1$, the matrix in (10.13) is singular if and only if*

$$c^2 = \frac{(g+\kappa n^2)\tanh(nh)}{n}. \tag{10.14}$$

Hence the possible wave speeds are

$$c_n^{\pm} = \pm\sqrt{\frac{(g+\kappa n^2)\tanh(nh)}{n}}, \qquad n \geq 1. \tag{10.15}$$

Proof. The determinant of the matrix in (10.13) is

$$\begin{aligned}\det\begin{pmatrix} -cn & -n\tanh(nh) \\ g+\kappa n^2 & cn \end{pmatrix} &= (-cn)(cn) - \big(-n\tanh(nh)\big)(g+\kappa n^2) \\ &= -c^2n^2 + n(g+\kappa n^2)\tanh(nh).\end{aligned}$$

Thus the determinant vanishes if and only if

$$c^2n^2 = n(g+\kappa n^2)\tanh(nh),$$

that is,

$$c^2 = \frac{(g+\kappa n^2)\tanh(nh)}{n}.$$

This proves (10.14) and hence (10.15). □

Remark 10.11. If $c \neq c_n^{\pm}$ for every $n \geq 1$, then the linearized operator

$$d_{(\eta,\psi)}F(c,0,0)$$

is invertible on the even–odd subspace. Therefore the possible bifurcation values along the trivial branch $(c,0,0)$ are precisely the speeds $c_n^{\pm}$.

Remark 10.12. Moreover, if for a fixed n the value $c_n^{\pm}$ is simple, namely

$$c_n^{\pm} \neq c_m^{\pm} \qquad \text{for every } m \neq n,$$

then the kernel is one-dimensional, which is exactly the situation needed in the Crandall–Rabinowitz theorem. To ensure this for some n, it is possible to study the map $n \mapsto c_n^{\pm}$ and understand whether or not it is injective!

Remark 10.13. In the infinite-depth case $h = +\infty$, one has $\tanh(nh) = 1$, hence

$$G(0) = |D|, \qquad c_n^{\pm} = \pm\sqrt{\frac{g + \kappa n^2}{n}}.$$

If in addition $\kappa = 0$, then

$$c_n^{\pm} = \pm\sqrt{\frac{g}{n}}.$$

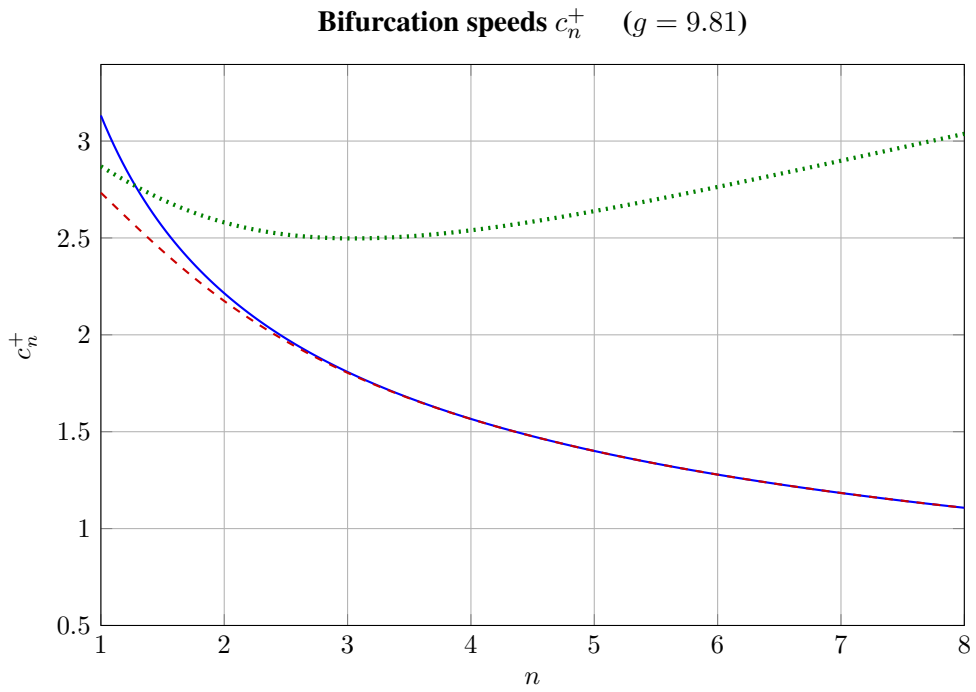

Figure 3: The three bifurcation speeds c_n^+ as functions of $n \geq 1$. **Case 1** (blue, solid): $h = \infty$, $\kappa = 0$ – strictly decreasing. **Case 2** (red, dashed): $h = 1$, $\kappa = 0$ – also decreasing. **Case 3** (green, dotted): $h = 1$, $\kappa = 1$ – decreases initially, then grows like $\sqrt{\kappa}\,\sqrt{n}$ (dash-dotted asymptote), eventually making $n \mapsto c_n^+$ injective.

Remark 10.14. At a simple bifurcation value $c = c_n^{\pm}$, the kernel of the linearized operator is generated by the n-th mode. Indeed, from

$$-cn\,h_n - n\tanh(nh)\varphi_n = 0$$

we get

$$\varphi_n = -\frac{c}{\tanh(nh)}\,h_n.$$

Hence one may choose the kernel generator in the form

$$\left(\cos(nx),\ -\frac{c_n^{\pm}}{\tanh(nh)}\sin(nx)\right).$$

10.2 Existence of small periodic travelling waves

Collecting now all the discussion together, let

$$X := H^s_{\text{even}}(\mathbb{T}), \qquad Y := H^s_{\text{odd}}(\mathbb{T}), \qquad Z := H^{s-1}_{\text{odd}}(\mathbb{T}) \times H^{s-2}_{\text{even}}(\mathbb{T}),$$

with $s > 3$, and let

$$F : \mathbb{R} \times X \times Y \to Z$$

be the travelling-wave operator introduced in the previous subsection:

$$F(c,\eta,\psi)=\begin{pmatrix} c\,\eta_x - G(\eta)[\psi] \\ c\,\psi_x + g\eta + \dfrac{1}{2}\psi_x^2 - \dfrac{1}{2}\dfrac{\big(\eta_x\psi_x + G(\eta)[\psi]\big)^2}{1+\eta_x^2} - \kappa H(\eta) \end{pmatrix}.$$

Recall that

$$F(c,0,0)=0 \qquad \text{for every } c\in\mathbb{R}.$$

Theorem 10.15 (Existence of small periodic travelling waves)**.** *Let $s>3$ and fix $n_0\geq 1$ and $\sigma\in\{+1,-1\}$. Set*

$$c_0 := c_{n_0}^{\sigma} := \sigma\sqrt{\frac{(g+\kappa n_0^2)\tanh(n_0 h)}{n_0}}. \tag{10.16}$$

Assume that c_0 is a simple bifurcation speed, namely

$$c_{n_0}^{\sigma} \neq c_n^{\tau} \qquad \textit{for every } (n,\tau)\neq(n_0,\sigma). \tag{10.17}$$

Then there exist $\varepsilon_0>0$ and C^1 maps

$$\varepsilon\mapsto c(\varepsilon)\in\mathbb{R}, \qquad \varepsilon\mapsto(\eta_\varepsilon,\psi_\varepsilon)\in X\times Y, \qquad |\varepsilon|<\varepsilon_0,$$

such that

$$c(0)=c_0, \qquad (\eta_0,\psi_0)=(0,0), \qquad F\big(c(\varepsilon),\eta_\varepsilon,\psi_\varepsilon\big)=0,$$

and $(\eta_\varepsilon,\psi_\varepsilon)\neq(0,0)$ for $\varepsilon\neq 0$.

Equivalently, for every $\varepsilon\neq 0$ sufficiently small, the pair

$$\eta(t,x)=\eta_\varepsilon(x+c(\varepsilon)t), \qquad \psi(t,x)=\psi_\varepsilon(x+c(\varepsilon)t)$$

is a nontrivial 2π-periodic travelling-wave solution of the water-wave system.

Moreover,

$$c(\varepsilon)=c_0+o(\varepsilon), \tag{10.18}$$

and

$$(\eta_\varepsilon,\psi_\varepsilon)=\varepsilon\,w^*+o(\varepsilon) \qquad \textit{in } X\times Y, \tag{10.19}$$

where

$$w^*(x)=\left(\cos(n_0x), -\sigma\sqrt{\frac{g+\kappa n_0^2}{n_0\tanh(n_0h)}}\,\sin(n_0x)\right). \tag{10.20}$$

Proof. We verify the hypotheses of the Crandall–Rabinowitz theorem for the map F at the point $(c_0,0,0)$.

Step 1: Regularity and trivial branch. By the previous subsection, the map

$$(\eta,\psi)\mapsto G(\eta)[\psi]$$

is smooth on suitable Sobolev spaces, and the curvature map

$$\eta\mapsto H(\eta)=\partial_x\left(\frac{\eta_x}{\sqrt{1+\eta_x^2}}\right)$$

is smooth from $H^s(\mathbb{T})$ to $H^{s-2}(\mathbb{T})$ for $s>3$. Hence

$$F\in C^2(\mathbb{R}\times X\times Y,Z),$$

indeed F is actually smoother than needed. Moreover, $F(c,0,0)=0$ for every $c\in\mathbb{R}$.

Step 2: Linearization at the trivial branch. From the previous subsection,

$$L_0 := d_{(\eta,\psi)}F(c_0,0,0) \tag{10.21}$$

is given by

$$L_0(h,\varphi) = \begin{pmatrix} c_0\, h_x - G(0)[\varphi] \\ c_0\, \varphi_x + g h - \kappa h_{xx} \end{pmatrix}. \tag{10.22}$$

We expand

$$h(x) = h_0 + \sum_{n\geq 1} h_n \cos(nx), \qquad \varphi(x) = \sum_{n\geq 1} \varphi_n \sin(nx),$$

and use the formula

$$G(0)[\sin(nx)] = n\tanh(nh)\sin(nx).$$

Then

$$L_0(h,\varphi) = \left(\sum_{n\geq 1} \big(-c_0 n\, h_n - n\tanh(nh)\varphi_n\big)\sin(nx),\ g h_0 + \sum_{n\geq 1} \big((g+\kappa n^2)h_n + c_0 n\,\varphi_n\big)\cos(nx) \right).$$

Thus, for each $n \geq 1$, the n-th mode is governed by the matrix

$$M_n(c_0) = \begin{pmatrix} -c_0 n & -n\tanh(nh) \\ g+\kappa n^2 & c_0 n \end{pmatrix}. \tag{10.23}$$

Its determinant is

$$\det M_n(c_0) = -c_0^2 n^2 + n(g+\kappa n^2)\tanh(nh).$$

Hence

$$\det M_n(c_0) = 0 \quad \Longleftrightarrow \quad c_0^2 = \frac{(g+\kappa n^2)\tanh(nh)}{n}.$$

By assumption (10.17), this happens only for $n = n_0$.

Step 3: Computation of the kernel. For $n \neq n_0$, the matrix $M_n(c_0)$ is invertible, so all corresponding Fourier coefficients vanish in the kernel. At $n = n_0$, the system

$$M_{n_0}(c_0)\begin{pmatrix} h_{n_0} \\ \varphi_{n_0} \end{pmatrix} = 0$$

reduces to

$$-c_0 n_0\, h_{n_0} - n_0\tanh(n_0 h)\varphi_{n_0} = 0.$$

Hence

$$\varphi_{n_0} = -\frac{c_0}{\tanh(n_0 h)}\, h_{n_0} = -\sigma\sqrt{\frac{g+\kappa n_0^2}{n_0\tanh(n_0 h)}}\, h_{n_0}.$$

Therefore

$$\ker L_0 = \operatorname{span}\{w^*\}, \tag{10.24}$$

where w^* is given by (10.20). In particular, $\dim\ker L_0 = 1$.

Step 4: Computation of the range. Let

$$(f,g) \in Z, \qquad f(x) = \sum_{n\geq 1} f_n \sin(nx), \qquad g(x) = g_0 + \sum_{n\geq 1} g_n \cos(nx).$$

The equation $L_0(h,\varphi) = (f,g)$ is equivalent to $g h_0 = g_0$ and, for each $n \geq 1$,

$$M_n(c_0)\begin{pmatrix} h_n \\ \varphi_n \end{pmatrix} = \begin{pmatrix} f_n \\ g_n \end{pmatrix}. \tag{10.25}$$

For $n \neq n_0$, this can be solved uniquely since $M_n(c_0)$ is invertible. For $n = n_0$, solvability is equivalent to the compatibility condition that the right-hand side be orthogonal to a left kernel vector of $M_{n_0}(c_0)$.

Set

$$\alpha_0 := \sqrt{\frac{g + \kappa n_0^2}{n_0 \tanh(n_0 h)}}. \tag{10.26}$$

A direct computation shows that $\ell_0 := (\sigma\alpha_0, 1)$ satisfies $\ell_0 M_{n_0}(c_0) = 0$. Therefore (10.25) at $n = n_0$ is solvable if and only if

$$\sigma\alpha_0 f_{n_0} + g_{n_0} = 0. \tag{10.27}$$

It follows that

$$\operatorname{Ran} L_0 = \{(f, g) \in Z : \ \sigma\alpha_0 f_{n_0} + g_{n_0} = 0\}. \tag{10.28}$$

Hence $\operatorname{Ran} L_0$ is a closed hyperplane of codimension 1 in Z.

Step 5: Transversality. We compute

$$\partial_c d_{(\eta,\psi)} F(c, 0, 0)[h, \varphi] = \begin{pmatrix} h_x \\ \varphi_x \end{pmatrix}.$$

Evaluating at $(c_0, 0, 0)$ on the kernel generator w^*, we obtain

$$\partial_c d_{(\eta,\psi)} F(c_0, 0, 0)[w^*] = (-n_0 \sin(n_0 x), -\sigma\alpha_0 n_0 \cos(n_0 x)).$$

For this pair, the compatibility quantity in (10.27) is

$$\sigma\alpha_0(-n_0) + (-\sigma\alpha_0 n_0) = -2\sigma\alpha_0 n_0 \neq 0.$$

Thus

$$\partial_c d_{(\eta,\psi)} F(c_0, 0, 0)[w^*] \notin \operatorname{Ran} L_0. \tag{10.29}$$

Step 6: Application of the Crandall–Rabinowitz theorem. All the hypotheses of the Crandall–Rabinowitz theorem are now verified:

1. $F \in C^1$;
2. $F(c, 0, 0) = 0$ for every c;
3. $\dim \ker L_0 = 1$ by (10.24);
4. $\operatorname{codim} \operatorname{Ran} L_0 = 1$ by (10.28);
5. the transversality condition holds by (10.29).

Therefore there exists a local bifurcating branch

$$\varepsilon \mapsto \big(c(\varepsilon), \eta_\varepsilon, \psi_\varepsilon\big)$$

of nontrivial zeros of F, with

$$\big(c(0), \eta_0, \psi_0\big) = (c_0, 0, 0),$$

and the expansions

$$c(\varepsilon) = c_0 + o(\varepsilon), \qquad (\eta_\varepsilon, \psi_\varepsilon) = \varepsilon w^* + o(\varepsilon).$$

This is exactly the desired conclusion. □

11 Hopf bifurcation 1 - 24/03/2026

We investigate now a Theorem concerning bifurcation from periodic orbits, which reflects the fact the now the kernel is no longer one-dimensional but two-dimensional, together with the codimension of the closed range.

Example 11.1 (Motivating example)**.** Consider the complex equation

$$\dot{z} = (\mu + i - |z|^2)z, \qquad z \in \mathbb{C}, \qquad \mu \in \mathbb{R}. \tag{11.1}$$

Write $z = \rho e^{i\theta}$, with $\rho \geq 0$. Then

$$\dot{\rho} = (\mu - \rho^2)\rho, \qquad \dot{\theta} = 1. \tag{11.2}$$

Hence:

1. if $\mu < 0$, then $\rho = 0$ is asymptotically stable;
2. if $\mu = 0$, then the equilibrium $\rho = 0$ loses hyperbolicity;
3. if $\mu > 0$, then $\rho = 0$ becomes unstable and a stable periodic orbit of radius $\sqrt{\mu}$ appears.

This is the basic picture of Hopf bifurcation.

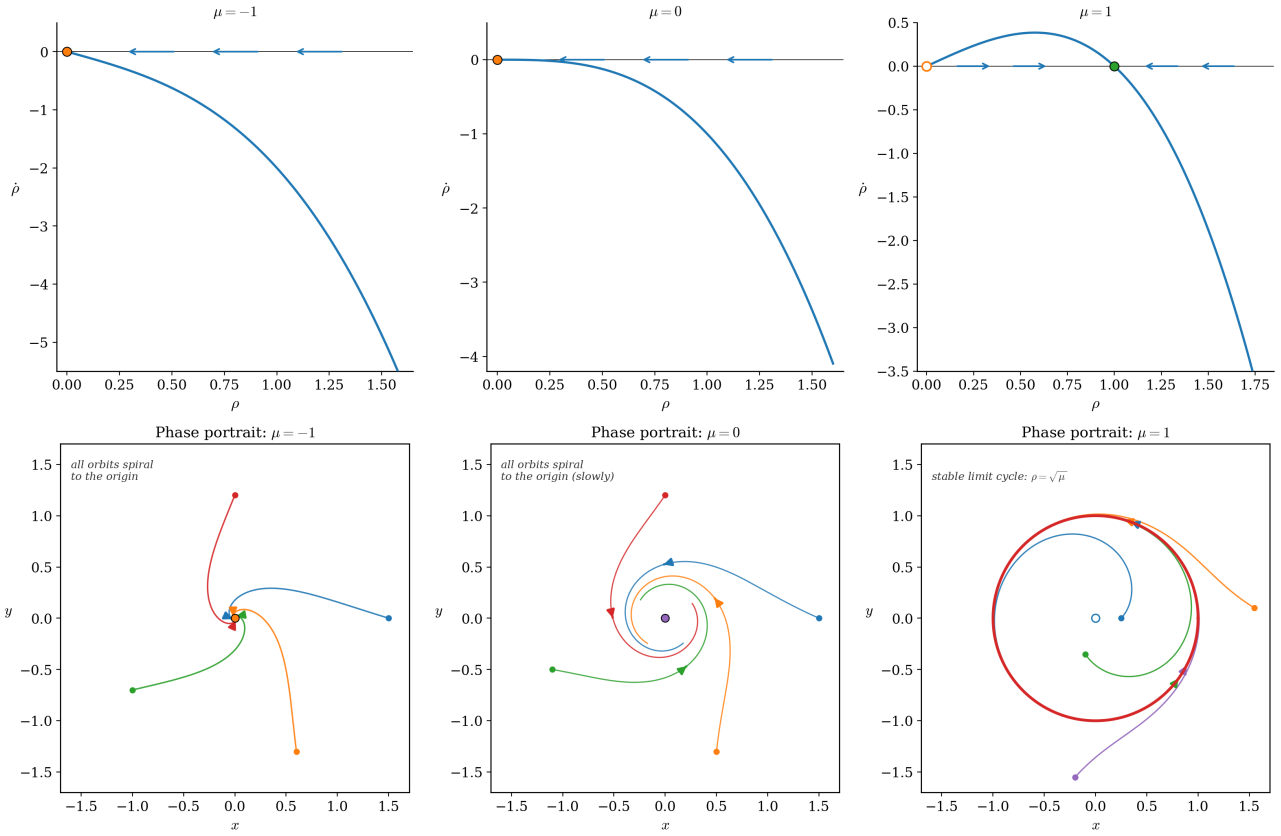

Figure 4: Hopf bifurcation: radial dynamics $\dot{\rho} = (\mu - \rho^2)\rho$ (top) and planar phase portraits (bottom) for $\mu = -1$, $\mu = 0$, and $\mu = 1$.

11.1 Abstract Hopf bifurcation

Let X, Y be Banach spaces and let

$$F \in C^2(\mathbb{R}^2 \times X, Y)$$

satisfy

$$F(\omega, \mu, 0) = 0 \qquad \forall (\omega, \mu) \in \mathbb{R}^2.$$

Fix a point $(\omega_0, \mu_0) \in \mathbb{R}^2$, viewed as a possible bifurcation point, and set

$$L := d_u F(\omega_0, \mu_0, 0) \in \mathcal{L}(X, Y). \tag{11.3}$$

A necessary condition for bifurcation is that L is not invertible.

Define $V := \ker L$, and $R := \operatorname{Ran} L$, and assume that $\dim V = 2$,, $\operatorname{codim} R = 2$,, and R is closed in Y. Choose closed complementary subspaces

$$X = W \oplus V, \qquad Y = R \oplus Z,$$

with $\dim Z = 2$, and let

$$Q : Y \to R, \qquad P : Y \to Z$$

be the corresponding projections.

Theorem 11.2 (Abstract Hopf bifurcation theorem)**.** *Assume that there exists $\bar{v} \in V \setminus \{0\}$ such that*

$$P\big(d^2_{\mu u} F(\omega_0, \mu_0, 0)[\bar{v}]\big), \qquad P\big(d^2_{\omega u} F(\omega_0, \mu_0, 0)[\bar{v}]\big) \tag{11.4}$$

are linearly independent in Z. Then (ω_0, μ_0) is a bifurcation point. More precisely, there exist $\omega, \mu \in C^1((-\varepsilon, \varepsilon), \mathbb{R})$ such that $\omega(0) = \omega_0$, and $\mu(0) = \mu_0$, and there exists a family of nontrivial solutions

$$u_s = s\bar{v} + w(\omega(s), \mu(s), s\bar{v}) \tag{11.5}$$

of

$$F(\omega(s), \mu(s), u_s) = 0, \tag{11.6}$$

where

$$w(\omega, \mu, v) \in W, \qquad w(\omega_0, \mu_0, 0) = 0, \qquad d_v w(\omega_0, \mu_0, 0) = 0.$$

In particular,

$$u_s = s\bar{v} + O(s^2).$$

Proof. Write $u = v + w$, with$v \in V$, and $w \in W$. Then $F(\omega, \mu, u) = 0$ is equivalent to the system

$$\begin{cases} QF(\omega, \mu, v + w) = 0, & \text{(range equation)}, \\ PF(\omega, \mu, v + w) = 0, & \text{(bifurcation equation)}. \end{cases} \tag{11.7}$$

Step 1: Solve the range equation. Define

$$G(\omega, \mu, v, w) := QF(\omega, \mu, v + w). \tag{11.8}$$

Then $G(\omega_0, \mu_0, 0, 0) = 0$, and moreover,

$$d_w G(\omega_0, \mu_0, 0, 0)[\hat{w}] = Q\, d_u F(\omega_0, \mu_0, 0)[\hat{w}] = QL\hat{w}, \qquad \hat{w} \in W. \tag{11.9}$$

Since

$$L|_W : W \to R$$

is bijective, and W, R are Banach spaces, the bounded inverse theorem implies that

$$L|_W \in \mathcal{L}(W, R)$$

is an isomorphism. Hence, by the implicit function theorem, there exists a unique map

$$w = w(\omega, \mu, v) \in W, \qquad w \in C^2,$$

defined near $(\omega_0, \mu_0, 0)$, such that

$$QF\big(\omega, \mu, v + w(\omega, \mu, v)\big) = 0. \tag{11.10}$$

Moreover, since $F(\omega,\mu,0)=0$ for all (ω,μ), we obtain $w(\omega,\mu,0)=0$. We now compute $d_v w(\omega_0,\mu_0,0)$. Differentiate (11.10) with respect to v at $(\omega_0,\mu_0,0)$ in the direction $\hat{v}\in V$:

$$QL\big(\hat{v}+d_v w(\omega_0,\mu_0,0)[\hat{v}]\big)=0. \tag{11.11}$$

Since $\hat{v}\in V=\ker L$, this reduces to

$$QL\,d_v w(\omega_0,\mu_0,0)[\hat{v}]=0.$$

As $QL|_W = L|_W$ is invertible, it follows that

$$d_v w(\omega_0,\mu_0,0)=0. \tag{11.12}$$

Step 2: Solve the bifurcation equation. Substituting the solution of the range equation into the second equation of (11.7), we get

$$PF\big(\omega,\mu,v+w(\omega,\mu,v)\big)=0. \tag{11.13}$$

Since $V\cong\mathbb{R}^2$, and $Z\cong\mathbb{R}^2$, we restrict ourselves to the one-dimensional slice $v=s\bar{v}$. Define

$$h(\omega,\mu,s):=PF\big(\omega,\mu,s\bar{v}+w(\omega,\mu,s\bar{v})\big). \tag{11.14}$$

Then $h(\omega,\mu,0)=0$ and we introduce

$$\chi(\omega,\mu,s):=\begin{cases}\dfrac{h(\omega,\mu,s)}{s}, & s\neq 0,\\ \partial_s h(\omega,\mu,0), & s=0.\end{cases} \tag{11.15}$$

Since $h\in C^2$, the map χ is of class C^1. Indeed, for $s\neq 0$,

$$\begin{aligned} h(\omega,\mu,s)-h(\omega,\mu,0) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\tau}h(\omega,\mu,\tau s)\,\mathrm{d}\tau\\ &= s\int_0^1 \partial_s h(\omega,\mu,\tau s)\,\mathrm{d}\tau, \end{aligned} \tag{11.16}$$

and therefore

$$\chi(\omega,\mu,s)=\int_0^1 \partial_s h(\omega,\mu,\tau s)\,\mathrm{d}\tau. \tag{11.17}$$

Moreover,

$$\begin{aligned} \chi(\omega_0,\mu_0,0) &= \partial_s h(\omega_0,\mu_0,0)\\ &= P\,d_u F(\omega_0,\mu_0,0)\big[\bar{v}+d_v w(\omega_0,\mu_0,0)[\bar{v}]\big]\\ &= P\,d_u F(\omega_0,\mu_0,0)[\bar{v}]\\ &= PL\bar{v}\\ &= 0, \end{aligned} \tag{11.18}$$

where we used (11.12) and $\bar{v}\in\ker L$.

Step 3: Implicit Function Theorem for χ. We now compute χ and its derivatives at $(\omega_0,\mu_0,0)$. One finds

$$\begin{aligned} \partial_\omega\chi(\omega_0,\mu_0,0) &= P\,d^2_{\omega u}F(\omega_0,\mu_0,0)[\bar{v}]+PL\big(d^2_{\omega v}w(\omega_0,\mu_0,0)[\bar{v}]\big)\\ &= P\,d^2_{\omega u}F(\omega_0,\mu_0,0)[\bar{v}], \end{aligned} \tag{11.19}$$

because $L\big(d^2_{\omega v}w(\omega_0,\mu_0,0)[\bar{v}]\big)\in R$ and P vanishes on R. Similarly,

$$\begin{aligned} \partial_\mu\chi(\omega_0,\mu_0,0) &= P\,d^2_{\mu u}F(\omega_0,\mu_0,0)[\bar{v}]+PL\big(d^2_{\mu v}w(\omega_0,\mu_0,0)[\bar{v}]\big)\\ &= P\,d^2_{\mu u}F(\omega_0,\mu_0,0)[\bar{v}]. \end{aligned} \tag{11.20}$$

By the assumption (11.4), the linear map

$$(\dot\omega, \dot\mu) \longmapsto \dot\omega\, \partial_\omega \chi(\omega_0, \mu_0, 0) + \dot\mu\, \partial_\mu \chi(\omega_0, \mu_0, 0)$$

is an isomorphism from $\mathbb{R}^2$ onto Z. Hence, by the implicit function theorem, there exist C^1 functions $\omega(s)$ and $\mu(s)$, defined for $|s|$ small, such that

$$\chi(\omega(s), \mu(s), s) = 0, \qquad \omega(0) = \omega_0, \qquad \mu(0) = \mu_0.$$

Since $h(\omega, \mu, s) = s\,\chi(\omega, \mu, s)$, we also have $h(\omega(s), \mu(s), s) = 0$. Together with (11.10), this yields $F(\omega(s), \mu(s), u_s) = 0$, where

$$u_s = s\bar v + w(\omega(s), \mu(s), s\bar v).$$

Finally, (11.12) implies $w(\omega(s), \mu(s), s\bar v) = O(s^2)$, hence $u_s = s\bar v + O(s^2)$. □

Remark 11.3. In the dynamical Hopf problem, periodic solutions are defined up to time translation. Thus the conclusion is best interpreted either after fixing a phase, or equivalently as the existence of a local branch of periodic orbits modulo time translation.

11.2 Dynamical Hopf bifurcation

We now present an application of the above abstract theorem relative to autonomous systems parameter dependent. To do so, we consider the autonomous system

$$\dot x = f(\mu, x), \qquad x \in \mathbb{R}^n, \qquad \mu \in \mathbb{R}, \tag{11.21}$$

where

$$f \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n), \qquad f(\mu, 0) = 0 \qquad \forall \mu \in \mathbb{R}.$$

The linearized equation at the equilibrium $x = 0$ is

$$\dot x = A(\mu)x, \qquad A(\mu) := d_x f(\mu, 0). \tag{11.22}$$

Assume:

(H1) for $\mu = \mu_0$, the matrix $A(\mu_0)$ is nonsingular and has a pair of simple purely imaginary eigenvalues

$$\pm i\omega_0, \qquad \omega_0 > 0; \tag{11.23}$$

Remark 11.4. Let $z = y_1 + iy_2, \qquad y_1, y_2 \in \mathbb{R}^n$, be an eigenvector corresponding to $i\omega_0$, so that $A(\mu_0)z = i\omega_0 z$. Then

$$A(\mu_0)y_1 = -\omega_0 y_2, \qquad A(\mu_0)y_2 = \omega_0 y_1. \tag{11.24}$$

Thus the real plane

$$E := \operatorname{span}_{\mathbb{R}}\{y_1, y_2\}$$

is invariant under $A(\mu_0)$ and is filled by periodic orbits of the linearized system.

(H2) $A(\mu_0)$ has no other eigenvalues of the form

$$\pm ik\omega_0, \qquad k \in \mathbb{Z}, \qquad |k| \geq 2; \tag{11.25}$$

Remark 11.5. Since $i\omega_0$ is a simple eigenvalue, it is stable under perturbations. Equivalently, if

$$p(\lambda, \mu) := \det(A(\mu) - \lambda I),$$

then

$$p(i\omega_0, \mu_0) = 0, \qquad \partial_\lambda p(i\omega_0, \mu_0) \neq 0.$$

Hence, by the implicit function theorem, there exists a unique C^1 branch

$$\mu \longmapsto \lambda(\mu)$$

of eigenvalues near μ_0 such that $\lambda(\mu_0) = i\omega_0$.

(H3) if $\lambda(\mu) = \alpha(\mu) + i\beta(\mu)$ is the branch of eigenvalues of $A(\mu)$ such that $\lambda(\mu_0) = i\omega_0$, then

$$\alpha'(\mu_0) \neq 0. \tag{11.26}$$

Namely, the branch is crossing the imaginary axis in a transversal way.

Remark 11.6. If $\alpha'(\mu_0) > 0$, then the equilibrium changes stability as μ crosses μ_0: for $\mu < \mu_0$ one has $\alpha(\mu) < 0$ and the equilibrium is linearly stable, while for $\mu > \mu_0$ one has $\alpha(\mu) > 0$ and the equilibrium is linearly unstable.

12 Hopf bifurcation 2 - 30/03/2026

Theorem 12.1 (Dynamical Hopf bifurcation)**.** *Under assumptions* $(H1), (H2), (H3)$*, from* $\mu = \mu_0$ *bifurcates a branch of nontrivial periodic solutions of the autonomous system*

$$\dot{x} = f(\mu, x), \qquad f \in C^2(\mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n), \tag{12.1}$$

with $f(\mu, 0) = 0$ *for all* $\mu \in \mathbb{R}$*, with period close to* $\frac{2\pi}{\omega_0}$*.*

More precisely, there exist $\omega, \mu \in C^1((-\varepsilon, \varepsilon), \mathbb{R})$ *such that*

$$\omega(s) \to \omega_0, \qquad \mu(s) \to \mu_0 \qquad \text{as } s \to 0,$$

and there exists a family of nontrivial solutions of (11.21) *having period* $\frac{2\pi}{\omega(s)}$*.*

Proof. We look for 2π-periodic solutions of the rescaled equation

$$\omega x'(t) - f(\mu, x(t)) = 0. \tag{12.2}$$

Define

$$F(\omega, \mu, x) := \omega x'(t) - f(\mu, x(t)), \tag{12.3}$$

so that

$$F : \mathbb{R}^2 \times C^1(\mathbb{T}; \mathbb{R}^n) \to C^0(\mathbb{T}; \mathbb{R}^n), \qquad F \in C^2,$$

and $F(\omega, \mu, 0) = 0$. Set $A_0 := A(\mu_0)$. Then

$$d_x F(\omega_0, \mu_0, 0)[h] = \omega_0 h'(t) - A_0 h(t). \tag{12.4}$$

Denote this linearized operator by

$$Lh := \omega_0 h'(t) - A_0 h(t). \tag{12.5}$$

Step 1: Computation of the kernel. Write the Fourier expansion

$$h(t) = \sum_{k \in \mathbb{Z}} h_k e^{ikt}, \qquad h_k \in \mathbb{C}^n, \qquad h_{-k} = \overline{h_k},$$

so that h is real-valued. Then

$$Lh = 0 \quad \Longleftrightarrow \quad (ik\omega_0 I - A_0)h_k = 0 \qquad \forall k \in \mathbb{Z}. \tag{12.6}$$

Let $z \in \mathbb{C}^n \setminus \{0\}$ satisfy

$$A_0 z = i\omega_0 z. \tag{12.7}$$

Then $A_0 \bar{z} = -i\omega_0 \bar{z}$.

- For $k = 1$ one has $h_1 \in \operatorname{span}_{\mathbb{C}}\{z\}$;
- For $k = -1$ one has $h_{-1} \in \operatorname{span}_{\mathbb{C}}\{\bar{z}\}$;
- For $|k| \geq 2$ assumption 11.2 implies $h_k = 0$;
- For $k = 0$, since A_0 is nonsingular, we have $h_0 = 0$.

Therefore

$$V := \ker L = \left\{ a z e^{it} + \bar{a}\, \bar{z} e^{-it} : a \in \mathbb{C} \right\}, \tag{12.8}$$

which is a two-dimensional real vector space.

Step 2: Computation of the range. Let

$$g(t) = \sum_{k \in \mathbb{Z}} g_k e^{ikt}.$$

To solve $Lh = g$ we must solve, for each $k \in \mathbb{Z}$,

$$(ik\omega_0 I - A_0)h_k = g_k. \tag{12.9}$$

For $|k| \neq 1$, the operator $ik\omega_0 I - A_0$ is invertible, so there is no obstruction. For $k = 1$, solvability requires

$$g_1 \in \operatorname{Ran}(i\omega_0 I - A_0),$$

while for $k = -1$, solvability requires

$$g_{-1} \in \operatorname{Ran}(-i\omega_0 I - A_0).$$

Since $i\omega_0$ is a simple eigenvalue of A_0, its algebraic and geometric multiplicities coincide, hence

$$\mathbb{C}^n = \ker(i\omega_0 I - A_0) \oplus \operatorname{Ran}(i\omega_0 I - A_0). \tag{12.10}$$

Pick $\eta \in \mathbb{C}^n$ the normalized dual eigenvector, i.e. $\langle \eta, z \rangle = 1$, and define the spectral projection $\Pi : \mathbb{C}^n \to \ker(i\omega_0 I - A_0)$ as

$$\Pi(\xi) := \langle \eta, \xi \rangle z. \tag{12.11}$$

Similarly, define $\bar{\Pi}(\xi) := \overline{\Pi(\bar{\xi})}$. Then

$$R := \operatorname{Ran} L = \left\{ g \in C^0(\mathbb{T}; \mathbb{R}^n) : \Pi(g_1) = 0, \quad \bar{\Pi}(g_{-1}) = 0 \right\} \tag{12.12}$$

and a complementary subspace is

$$Z := \left\{ \Pi(g_1)e^{it} + \bar{\Pi}(g_{-1})e^{-it} : g \in C^0(\mathbb{T}; \mathbb{R}^n) \right\}, \tag{12.13}$$

and therefore $Y = R \oplus Z$, with $\dim Z = 2$. On Y we define $P : Y \to Z$ be the corresponding projection, which has the form

$$P(g) = \Pi(g_1)e^{it} + \bar{\Pi}(g_{-1})e^{-it} = \Pi(g_1)e^{it} + \overline{\Pi(g_1)}e^{-it}. \tag{12.14}$$

Step 3: Verification of the transversality condition. Take

$$\bar{v} := ze^{it} + \bar{z}e^{-it} \in V. \tag{12.15}$$

Since $F(\omega, \mu, x) = \omega x' - f(\mu, x)$, we have

$$d^2_{\omega x} F(\omega_0, \mu_0, 0)[h] = h'.$$

Hence

$$d^2_{\omega x} F(\omega_0, \mu_0, 0)[\bar{v}] = ize^{it} - i\bar{z}e^{-it}, \tag{12.16}$$

from which it follows from the definition of P that

$$P\big(d^2_{\omega x} F(\omega_0, \mu_0, 0)[\bar{v}]\big) = ize^{it} - i\bar{z}e^{-it}. \tag{12.17}$$

Next,

$$d^2_{\mu x} F(\omega_0, \mu_0, 0)[h] = -\partial_\mu A(\mu_0)h.$$

Differentiate the eigenvalue equation

$$A(\mu)z(\mu) = \lambda(\mu)z(\mu)$$

at $\mu = \mu_0$:

$$\partial_\mu A(\mu_0)z + A_0 z'(\mu_0) = \lambda'(\mu_0)z + i\omega_0 z'(\mu_0). \tag{12.18}$$

Therefore

$$\big(\partial_\mu A(\mu_0) - \lambda'(\mu_0)I\big)z = (i\omega_0 I - A_0)z'(\mu_0) \in \operatorname{Ran}(i\omega_0 I - A_0). \tag{12.19}$$

Applying Π and using (12.11), we obtain

$$\Pi\big(\partial_\mu A(\mu_0)z\big) = \lambda'(\mu_0)z. \tag{12.20}$$

Thus

$$P\big(d^2_{\mu x}F(\omega_0,\mu_0,0)[\bar v]\big) = -\lambda'(\mu_0)ze^{it} - \overline{\lambda'(\mu_0)}\,\bar z e^{-it}. \tag{12.21}$$

Under the real-linear identification

$$aze^{it} + \bar a\,\bar z e^{-it} \longleftrightarrow a \in \mathbb{C},$$

the vectors (12.17) and (12.21) correspond to

$$i, \qquad -\lambda'(\mu_0).$$

They are linearly independent over $\mathbb{R}$ if and only if

$$\lambda'(\mu_0) \notin i\mathbb{R}.$$

Since

$$\lambda'(\mu_0) = \alpha'(\mu_0) + i\beta'(\mu_0),$$

this is equivalent to

$$\alpha'(\mu_0) \neq 0,$$

which is precisely assumption 11.2.

Step 4: conclusion. All assumptions of Theorem 11.2 are satisfied. Therefore there exist

$$\omega(s), \qquad \mu(s),$$

and nontrivial 2π-periodic solutions

$$u_s(t) = s\bar v + o(s) = s\big(ze^{it} + \bar z e^{-it}\big) + o(s)$$

of the rescaled equation (12.2). Returning to the original time variable, we obtain a family of nontrivial periodic solutions of (11.21), with period

$$\frac{2\pi}{\omega(s)}.$$

□

12.1 A Van der Pol-type example

Example 12.2 (A Van der Pol-type planar system)**.** Consider the system

$$\begin{cases} x' = y + \mu x - x^3, \\ y' = -x. \end{cases} \tag{12.22}$$

Equivalently, eliminating y, one obtains

$$x'' - (\mu - 3x^2)x' + x = 0.$$

Then the following hold:

1. $\mu = 0$ is a Hopf bifurcation value;
2. if $\mu \leq 0$, then (12.22) has no nontrivial periodic solutions;
3. if $\mu > 0$, then a branch of nontrivial periodic solutions bifurcates from the origin.

Proof. The linearization of (12.22) at the origin is

$$A_\mu = \begin{pmatrix} \mu & 1 \\ -1 & 0 \end{pmatrix}.$$

The eigenvalues are

$$\lambda_\pm(\mu) = \frac{\mu \pm \sqrt{\mu^2 - 4}}{2}.$$

Hence at $\mu = 0$ one has

$$\lambda_\pm(0) = \pm i,$$

and

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Re}\lambda_\pm(\mu)\Big|_{\mu=0} = \frac{1}{2} \neq 0.$$

Therefore the dynamical Hopf bifurcation theorem applies at $(\omega_0, \mu_0) = (1, 0)$. In particular, there is a local branch of nontrivial periodic solutions with associated frequency $\omega(s)$ and parameter $\mu(s)$ satisfying $\omega(0) = 1$ and $\mu(0) = 0$.

We now show that for $\mu \leq 0$ there are no nontrivial periodic solutions. Consider the energy

$$E(x, y) := \frac{1}{2}(x^2 + y^2).$$

Along solutions,

$$\begin{aligned} \dot{E} &= xx' + yy' \\ &= x(y + \mu x - x^3) + y(-x) \\ &= \mu x^2 - x^4 \\ &= x^2(\mu - x^2). \end{aligned}$$

If $\mu \leq 0$, then

$$\dot{E} \leq 0 \qquad \text{for all } (x, y).$$

If a solution were periodic, integrating $\dot{E}$ over one period would give

$$0 = \int_0^T \dot{E}(t)\,\mathrm{d}t = \int_0^T x(t)^2\big(\mu - x(t)^2\big)\,\mathrm{d}t.$$

Since the integrand is everywhere nonpositive, it must vanish identically. Hence

$$x \equiv 0.$$

Then $y' = -x$ implies $y' \equiv 0$, and from

$$x' = y + \mu x - x^3$$

we obtain $y \equiv 0$. Therefore the only periodic solution for $\mu \leq 0$ is the trivial one.

It follows that the Hopf branch produced above cannot stay on the slice $\mu = 0$. Indeed, if $\mu(s) \equiv 0$, then for $s \neq 0$ small we would get nontrivial periodic solutions at $\mu = 0$, contradicting the previous argument. Hence, for $s \neq 0$ small,

$$\mu(s) > 0.$$

So the branch is genuinely one-sided. In the usual Hopf picture one expects

$$\mu'(0) = 0, \qquad \mu(s) = c\,s^2 + o(s^2)$$

with $c > 0$, that is, a supercritical branch tangent to the critical line $\mu = 0$.

We conclude with a qualitative phase-plane picture for $\mu > 0$, which strongly suggests the existence of a stable periodic orbit. The nullclines are

$$y' = 0 \iff x = 0,$$

and

$$x' = 0 \iff y = -x(\mu - x^2) = x^3 - \mu x.$$

Thus the y-axis is one nullcline, while the second one is the odd cubic

$$y = x^3 - \mu x.$$

Near the origin the linearization has eigenvalues with positive real part, so the origin is an unstable focus. Hence, if we start on the positive y-axis at a point $(0, a)$ with $a > 0$ small, the trajectory spirals outward and, after one turn around the origin, it meets again the positive y-axis at a height larger than a.

On the other hand, for large excursions the quartic term dominates in

$$\dot{E} = \mu x^2 - x^4 = x^2(\mu - x^2).$$

Therefore, when $|x| > \sqrt{\mu}$ the energy decreases. If the trajectory makes a sufficiently large loop, then during much of one turn it travels in the region $|x| > \sqrt{\mu}$, so the next intersection with the positive y-axis occurs at a smaller height.

This suggests considering the first-return map on the positive y-axis: to an initial value $a > 0$ one associates the next intersection height $P(a)$. The previous discussion indicates that

$$P(a) > a \quad \text{for } a > 0 \text{ small}, \qquad P(a) < a \quad \text{for } a \text{ large}.$$

By continuity, there should exist $a_* > 0$ such that

$$P(a_*) = a_*.$$

The corresponding trajectory is then a closed orbit. Moreover, since nearby small intersections are pushed outward and large ones are pushed inward, this closed orbit is expected to be asymptotically stable.

Thus the Hopf branch lies on the $\mu > 0$ side and corresponds to a small stable limit cycle surrounding the unstable focus at the origin. □

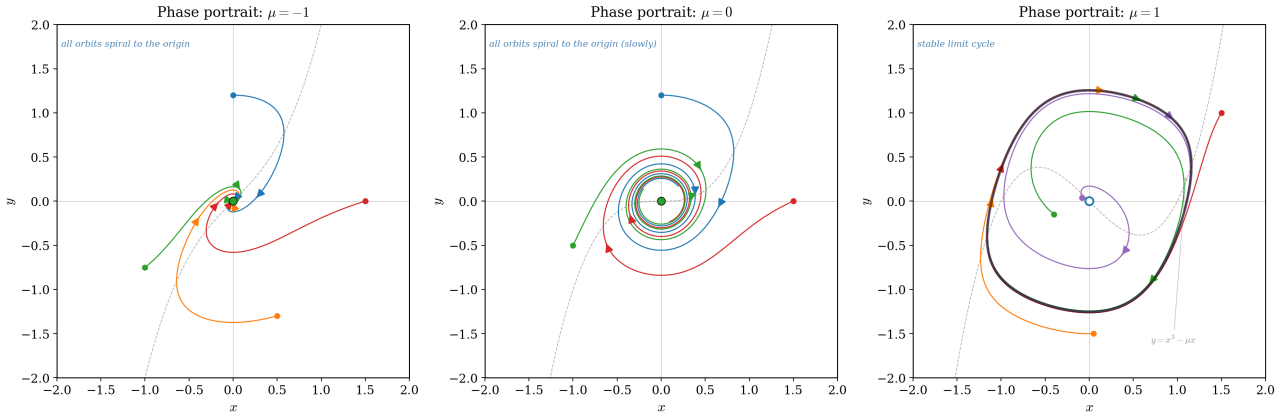

Figure 5: Phase portraits of the Van der Pol-type system $\dot{x} = y + \mu x - x^3$, $\dot{y} = -x$ for $\mu = -1, 0, 1$. For $\mu > 0$ a stable limit cycle of amplitude $\rho \approx \sqrt{\mu}$ emerges: trajectories starting inside spiral outward, while those starting outside spiral inward, both converging to the periodic orbit (black curve). The dashed gray curve is the nullcline $y = x^3 - \mu x$.

Exercise 12.3 (A planar system)**.** Let

$$\phi, \psi \in C^\infty(\mathbb{R}), \qquad \phi(0) = \psi(0) = 0.$$

Discuss the Hopf bifurcation for the system

$$\begin{cases} x' = \phi(y), \\ y' = \psi(x) + \mu y. \end{cases} \tag{12.23}$$

13 Brouwer degree 1 - 13/04/2026

We look for a topological method to count the number of solutions to the equation

$$f(x) = y.$$

Example 13.1 (Winding number for holomorphic functions on $\mathbb{C}$)**.** Let $f \in H(\mathbb{C})$ be a holomorphic function, and recall the definition of the winding number $n(\gamma, z_0)$.

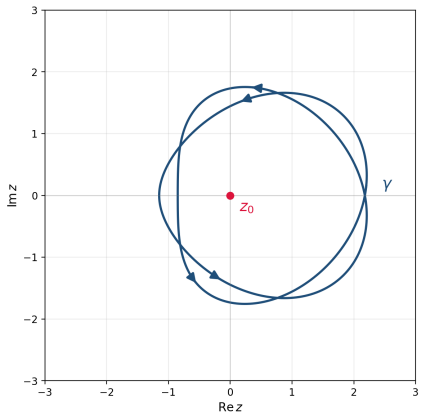

Figure 6: A curve γ around $z_0 = 0$.

Let $\gamma : [0,1] \to \mathbb{C} \setminus \{z_0\}$ be a closed path around z_0, and define

$$n(\gamma, z_0) := \frac{1}{2\pi i} \int_\gamma \frac{\mathrm{d}z}{z - z_0} \in \mathbb{Z}. \tag{13.1}$$

Then, for f, assuming $0 \notin f(\gamma)$, we compute

$$n(f(\gamma), 0) = \frac{1}{2\pi i} \int_\gamma \frac{f'(z)}{f(z)} \,\mathrm{d}z = \sum_k n(\gamma, z_k)\, \alpha_k, \tag{13.2}$$

where z_k are the zeros of f, each with multiplicity α_k.

So, if γ is a simple closed curve, then $n(f(\gamma), 0)$ counts exactly the number of zeros of f enclosed by γ, counted with multiplicity.

It might be hard to compute exactly the zeros of f, however, the Rouchè Theorem tells us that the number of zeros is stable under small perturbation.

For this reason we recall the concept of homotopy. Namely, if $f, g : \gamma \to \mathbb{C} \setminus \{0\}$, a homotopy between f and g is a continuous map

$$H : [0,1] \times \gamma \to \mathbb{C} \setminus \{0\}$$

such that

$$H(0, z) = f(z), \qquad H(1, z) = g(z), \qquad z \in \gamma.$$

We recall that, if f and g are homotopic, then

$$n(f(\gamma), 0) = n(g(\gamma), 0). \tag{13.3}$$

Example 13.2 (A one-dimensional motivating example)**.** Consider a continuous map $f : [0,1] \to \mathbb{R}$ and fix $y \in \mathbb{R}$. A first naive idea is to count the number of solutions of

$$f(x) = y.$$

This number is not stable under deformation: two intersections can be created or annihilated, and a zero can disappear by crossing the boundary.

If y is a regular value, a better idea is to count zeros with sign. Namely, if $f \in C^1([0,1])$ and every solution of $f(x) = y$ is simple, one considers

$$\sum_{x \in f^{-1}(y)} \operatorname{sign} f'(x). \tag{13.4}$$

Under deformations, simple zeros can only be created or destroyed in pairs with opposite signs, so this signed count is stable as long as no zero crosses the boundary. Brouwer degree is the higher-dimensional analogue of this signed counting principle.

Definition 13.3. Let $U \subset \mathbb{R}^n$ be open and bounded, and fix $y \in \mathbb{R}^n$. We define

$$D_y(U, \mathbb{R}^n) := \left\{ f \in C^0(\overline{U}, \mathbb{R}^n) \ : \ y \notin f(\partial U) \right\}. \tag{13.5}$$

Definition 13.4 (Axiomatic approach)**.** A degree map

$$\deg(\cdot, U, y) : D_y(U, \mathbb{R}^n) \to \mathbb{R}$$

is any map satisfying the following properties:

(D1) *Translation invariance:*

$$\deg(f, U, y) = \deg(f - y, U, 0).$$

(D2) *Normalization:*

$$\deg(\mathrm{Id}, U, y) = \begin{cases} 1, & y \in U, \\ 0, & y \in \mathbb{R}^n \setminus U. \end{cases}$$

(D3) *Additivity:* if $U_1, U_2 \subset U$ are open and disjoint, and

$$y \notin f\big(U \setminus (U_1 \cup U_2)\big),$$

then

$$\deg(f, U, y) = \deg(f, U_1, y) + \deg(f, U_2, y).$$

(D4) *Homotopy invariance:* if

$$H(t) := (1-t)f + tg \in D_y(U, \mathbb{R}^n) \qquad \forall\, t \in [0,1],$$

then

$$\deg(f, U, y) = \deg(g, U, y).$$

Proposition 13.5 (Properties of a degree map)**.** *Let* deg *be a degree map. Then:*

1. $$\deg(f, \varnothing, y) = 0.$$

2. *If U_i, $1 \le i \le N$, are pairwise disjoint open subsets of U such that*

$$y \notin f\Big(U \setminus \bigcup_{i=1}^{N} U_i\Big),$$

then

$$\deg(f, U, y) = \sum_{i=1}^{N} \deg(f, U_i, y).$$

3. *If*

$$\sup_{x \in \partial U} |f(x) - g(x)| < \operatorname{dist}\big(y, f(\partial U)\big),$$

then

$$\deg(f, U, y) = \deg(g, U, y).$$

In particular, if $f = g$ on ∂U, then

$$\deg(f, U, y) = \deg(g, U, y).$$

Proof. For (1), apply (D3) with $U_1 = U$ and $U_2 = \varnothing$. Then

$$\deg(f, U, y) = \deg(f, U, y) + \deg(f, \varnothing, y),$$

hence $\deg(f, \varnothing, y) = 0$.

For (3), consider the linear homotopy

$$H(t, x) := (1 - t)f(x) + tg(x), \qquad (t, x) \in [0, 1] \times \overline{U}.$$

For every $x \in \partial U$ and every $t \in [0, 1]$,

$$\begin{aligned} |H(t, x) - y| &= |(1 - t)(f(x) - y) + t(g(x) - y)| \\ &= |f(x) - y + t(g(x) - f(x))| \\ &\geq |f(x) - y| - t\,|g(x) - f(x)| \\ &\geq \operatorname{dist}\big(y, f(\partial U)\big) - |f(x) - g(x)| \\ &> 0. \end{aligned}$$

Thus $H(t) \in D_y(U, \mathbb{R}^n)$ for all $t \in [0, 1]$, and by (D4)

$$\deg(f, U, y) = \deg(g, U, y).$$

The final statement follows by taking $f = g$ on ∂U. □

Exercise 13.6. Prove point (2) of the previous proposition.

Exercise 13.7. Prove that if $y \notin f(U)$, then

$$\deg(f, U, y) = 0.$$

The converse is false. Equivalently, show that

$$\deg(f, U, y) \neq 0 \qquad \Longrightarrow \qquad y \in f(U).$$

Proposition 13.8. *The two maps*

$$D_y(U, \mathbb{R}^n) \ni f \longmapsto \deg(f, U, y)$$

and

$$\mathbb{R}^n \setminus f(\partial U) \ni y \longmapsto \deg(f, U, y)$$

are locally constant.

Proof. Fix $f \in D_y(U, \mathbb{R}^n)$ and let $g \in C^0(\overline{U}, \mathbb{R}^n)$ satisfy

$$\|f - g\|_{L^\infty(\partial U)} < \operatorname{dist}\big(y, f(\partial U)\big).$$

Then, by point (3) of the previous proposition,

$$\deg(f, U, y) = \deg(g, U, y).$$

Hence the map

$$f \longmapsto \deg(f, U, y)$$

is locally constant.

Now fix $f \in D_y(U, \mathbb{R}^n)$ and let $\bar{y} \in \mathbb{R}^n$ satisfy

$$|\bar{y} - y| < \operatorname{dist}\big(y, f(\partial U)\big).$$

Define

$$y(t) := (1 - t)\bar{y} + ty, \qquad H(t) := f - y(t).$$

Then $H(t) \in D_0(U, \mathbb{R}^n)$ for every $t \in [0, 1]$. By translation invariance and homotopy invariance,

$$\deg(f, U, \bar{y}) = \deg(f - \bar{y}, U, 0) = \deg(f - y, U, 0) = \deg(f, U, y).$$

Hence the map

$$y \longmapsto \deg(f, U, y)$$

is locally constant on $\mathbb{R}^n \setminus f(\partial U)$. □

Corollary 13.9. *Let $y_1, y_2 \in \mathbb{R}^n$, let $f_1 \in D_{y_1}(U, \mathbb{R}^n)$ and $f_2 \in D_{y_2}(U, \mathbb{R}^n)$, and define*

$$H(t) := (1-t)f_1 + tf_2.$$

Assume that there exists a continuous path

$$y : [0,1] \to \mathbb{R}^n, \qquad y(0) = y_1, \quad y(1) = y_2,$$

such that

$$H(t) \in D_{y(t)}(U, \mathbb{R}^n) \qquad \forall\, t \in [0,1].$$

Then

$$t \longmapsto \deg(H(t), U, y(t))$$

is constant. In particular,

$$\deg(f_1, U, y_1) = \deg(f_2, U, y_2).$$

Proof. Define

$$\widehat{H}(t) := H(t) - y(t).$$

Then $\widehat{H}(t) \in D_0(U, \mathbb{R}^n)$ for every $t \in [0,1]$, and by translation invariance

$$\deg(H(t), U, y(t)) = \deg(\widehat{H}(t), U, 0).$$

By the previous proposition, the map

$$t \longmapsto \deg(\widehat{H}(t), U, 0)$$

is locally constant, hence constant on the connected interval $[0,1]$. □

13.1 Construction of a degree map

Definition 13.10. Let $f \in C^1(U, \mathbb{R}^n)$.

1. A point $x \in U$ is called a *regular point* of f if $df(x)$ is invertible.
2. A point $y \in f(U)$ is called a *regular value* of f if every $x \in f^{-1}(y)$ is a regular point.

Remark 13.11. If $y \notin f(U)$, then y is automatically a regular value.

Proposition 13.12. *Let $U \subset \mathbb{R}^n$ be open and bounded, let*

$$f \in C^1(U, \mathbb{R}^n) \cap C^0(\overline{U}, \mathbb{R}^n),$$

let $y \notin f(\partial U)$, and assume that y is a regular value of f. Then

$$f^{-1}(y) = \{x \in U : \ f(x) = y\}$$

is a finite set. More precisely, if

$$f^{-1}(y) = \{x_1, \dots, x_m\},$$

then there exist pairwise disjoint neighborhoods U_{x_i} of x_i and a neighborhood U_y of y such that

$$f : U_{x_i} \to U_y$$

is invertible for every i, and

$$f^{-1}(U_y) = \bigcup_{i=1}^{m} U_{x_i}.$$

Proof. We first show that $f^{-1}(y)$ is finite. Assume by contradiction that there is an infinite sequence $(x_k) \subset f^{-1}(y)$. Since $\overline{U}$ is compact, up to a subsequence we may assume that

$$x_k \to x \in \overline{U}.$$

By continuity,

$$f(x) = \lim_{k\to\infty} f(x_k) = y.$$

Since $y \notin f(\partial U)$, necessarily $x \in U$. Because y is a regular value, $df(x)$ is invertible. Hence the inverse function theorem yields a neighborhood $V_x \subset U$ of x such that

$$f : V_x \to f(V_x)$$

is injective. This is impossible, because infinitely many distinct points x_k belong to V_x for k large and all satisfy $f(x_k) = y$. Therefore $f^{-1}(y)$ is finite.

Now let

$$f^{-1}(y) = \{x_1, \dots, x_m\}.$$

Since $df(x_i)$ is invertible for every i, the inverse function theorem gives neighborhoods $\widetilde{U}_{x_i}$ of x_i and neighborhoods V_i of y such that

$$f : \widetilde{U}_{x_i} \to V_i$$

is a C^1-diffeomorphism. Shrinking the neighborhoods if necessary, we may assume that the $\widetilde{U}_{x_i}$ are pairwise disjoint. Set

$$U_y := \bigcap_{i=1}^{m} V_i.$$

Replacing $\widetilde{U}_{x_i}$ by

$$U_{x_i} := \widetilde{U}_{x_i} \cap f^{-1}(U_y),$$

we obtain pairwise disjoint neighborhoods U_{x_i} of x_i such that

$$f : U_{x_i} \to U_y$$

is invertible for every i.

It remains to show that

$$f^{-1}(U_y) = \bigcup_{i=1}^{m} U_{x_i}.$$

If not, there would exist a sequence $(z_k) \subset \overline{U} \setminus \bigcup_{i=1}^{m} U_{x_i}$ such that

$$f(z_k) \to y.$$

By compactness, up to a subsequence,

$$z_k \to z \in \overline{U} \setminus \bigcup_{i=1}^{m} U_{x_i}.$$

Again by continuity,

$$f(z) = y,$$

so $z \in f^{-1}(y) = \{x_1, \dots, x_m\}$, a contradiction. □

Remark 13.13 (Reduction to the differential at a regular value)**.** Let

$$f \in C^1(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n),$$

and assume that y is a regular value. Write

$$f^{-1}(y) = \{x_1, \dots, x_N\}.$$

Choose pairwise disjoint balls $B_{\rho_i}(x_i) \Subset U$ around these points. By additivity,

$$\deg(f, U, y) = \sum_{i=1}^{N} \deg\big(f, B_{\rho_i}(x_i), y\big). \tag{13.6}$$

Near x_i we have the first-order expansion

$$f(x) = f(x_i) + df(x_i)(x - x_i) + o(|x - x_i|) = y + df(x_i)(x - x_i) + o(|x - x_i|). \tag{13.7}$$

If we set

$$L_i(x) := y + df(x_i)(x - x_i),$$

then on $\partial B_{\rho_i}(x_i)$ one has

$$\|f - L_i\|_{L^\infty(\partial B_{\rho_i}(x_i))} \le \varepsilon(\rho_i)\rho_i, \qquad \varepsilon(\rho_i) \to 0 \quad \text{as } \rho_i \to 0,$$

while invertibility of $df(x_i)$ gives a lower bound

$$|L_i(x) - y| = |df(x_i)(x - x_i)| \ge c_i\, \rho_i \qquad \text{for } x \in \partial B_{\rho_i}(x_i).$$

Hence, for ρ_i small enough, the stability property yields

$$\deg\big(f, B_{\rho_i}(x_i), y\big) = \deg\big(L_i, B_{\rho_i}(x_i), y\big).$$

After translating x_i to the origin, this becomes

$$\deg\big(L_i, B_{\rho_i}(x_i), y\big) = \deg\big(df(x_i), B_{\rho_i}(0), 0\big).$$

Therefore

$$\deg(f, U, y) = \sum_{i=1}^{N} \deg\big(df(x_i), B_{\rho_i}(0), 0\big). \tag{13.8}$$

So the computation of the degree reduces to the computation of the degree of invertible matrices.

Exercise 13.14. Show that

$$\deg\big(df(x_i), B_{\rho_i}(0), 0\big) = \deg\big(df(x_i), B_1(0), 0\big). \tag{13.9}$$

14 The Brouwer mapping degree 2 - 14/04/2026

Proposition 14.1. *Let $L_1, L_2 \in GL(n)$. Then L_1 and L_2 are homotopic inside $GL(n)$ if and only if*

$$\operatorname{sign}\det L_1 = \operatorname{sign}\det L_2.$$

Proof. We show that every $L \in GL(n)$ is homotopic in $GL(n)$ to

$$\operatorname{diag}(\operatorname{sign}\det L, 1, \dots, 1).$$

Adding a multiple of one row to another and multiplying a row by a positive constant can be realized by continuous paths inside $GL(n)$. Indeed, if E_{ij} denotes the elementary matrix with entry 1 in position (i,j) and zero elsewhere, then

$$\gamma(t) := (I + tE_{ij})L, \qquad t \in [0,1],$$

satisfies

$$\det\gamma(t) = \det(I + tE_{ij})\det L = \det L \neq 0,$$

so $\gamma(t) \in GL(n)$ for all t. Similarly, multiplying one row by a positive constant is realized by a path of diagonal matrices with positive diagonal entries.

Using these operations, L is homotopic in $GL(n)$ to a diagonal matrix

$$\operatorname{diag}(m_1, \dots, m_n), \qquad m_i \in \{\pm 1\}.$$

Now the 2×2 path

$$\begin{pmatrix} \pm\cos(\pi t) & \mp\sin(\pi t) \\ \sin(\pi t) & \cos(\pi t) \end{pmatrix}, \qquad t \in [0,1], \tag{14.1}$$

shows that

$$\operatorname{diag}(\pm 1, 1) \quad \text{and} \quad \operatorname{diag}(\mp 1, -1)$$

are homotopic inside $GL(2)$. Applying this to successive 2×2 diagonal blocks, we can reduce $\operatorname{diag}(m_1, \dots, m_n)$ to

$$\operatorname{diag}(\operatorname{sign}\det L, 1, \dots, 1).$$

This proves the "if" direction. Conversely, if L_1 and L_2 are homotopic in $GL(n)$, then the determinant cannot change sign along the homotopy, since otherwise it would vanish somewhere along the path, contradicting invertibility. □

Theorem 14.2. *Let $L \in GL(n)$. Then*

$$\deg\big(L, B_1(0), 0\big) = \operatorname{sign}\det L. \tag{14.2}$$

Proof. If $\det L > 0$, then L is homotopic in $GL(n)$ to the identity. Hence, by homotopy invariance and normalization,

$$\deg\big(L, B_1(0), 0\big) = \deg\big(I, B_1(0), 0\big) = 1.$$

Assume now that $\det L < 0$. By the previous proposition, it is enough to compute the degree of

$$M_- := \operatorname{diag}(-1, 1, \dots, 1).$$

Set

$$U_1 := (-1,1)^n, \quad U_2 := (1,3) \times (-1,1)^{n-1}, \quad U := (-1,3) \times (-1,1)^{n-1}. \tag{14.3}$$

Define

$$g(r) := 2 - |r-1|, \qquad h(r) := 1 - r^2,$$

and consider the two maps

$$f_1(x) := \big(1 - g(x_1)h(x_2)\cdots h(x_n), x_2, \dots, x_n\big), \qquad f_2(x) := (1, x_2, \dots, x_n).$$

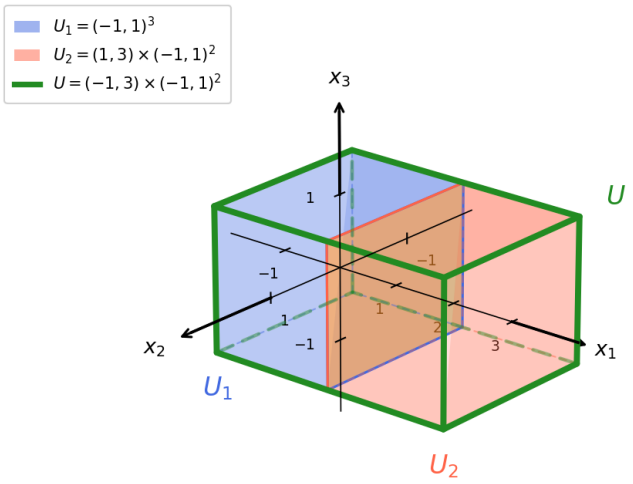

Figure 7: Visual representation of the sets U_1, U_2, and U for $n = 3$.

Then $f_1 = f_2$ on ∂U, hence

$$\deg(f_1, U, 0) = \deg(f_2, U, 0) = 0.$$

Moreover,

$$f_1^{-1}(0) = \{0, (2, 0, \dots, 0)\},$$

and these are regular zeros. A direct computation gives

$$df_1(0) = M_-, \qquad df_1(2, 0, \dots, 0) = I.$$

Using additivity and the reduction to the differential at regular values, after shrinking to small neighborhoods of the two zeros and translating the second neighborhood to the origin, we obtain

$$0 = \deg(f_1, U, 0) = \deg(M_-, B_1(0), 0) + \deg(I, B_1(0), 0).$$

Since $\deg(I, B_1(0), 0) = 1$, it follows that

$$\deg(M_-, B_1(0), 0) = -1.$$

Therefore

$$\deg\big(L, B_1(0), 0\big) = -1$$

whenever $\det L < 0$. □

14.1 Generalization to continuos functions

Remark 14.3. We have a formula for the degree of a C^1 function at a regular value. If

$$f \in C^1(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n)$$

and y is a regular value of f, then

$$\deg(f, U, y) = \sum_{x \in f^{-1}(y)} \operatorname{sign} \det(df(x)). \tag{14.4}$$

We now have to show that a degree map actually exists, and in doing so we must remove both the regular value assumption and the extra smoothness assumption. We proceed in two steps:

1. first we assume $f \in C^2$ and remove the regular value assumption by proving that the degree is locally constant with respect to y;
2. then we remove the C^2 assumption by proving invariance under small perturbations and approximating continuous maps by smooth ones.

14.2 Step 1: admitting critical values.

Our goal is to show that the degree is locally constant with respect to $y \in \mathbb{R}^n \setminus f(\partial U)$.

The main obstruction is the following: if we connect y to a nearby regular value $\tilde{y}$ by a path, then this path may cross critical values. Hence we look for an integral representation of the degree which still makes sense without assuming that y is regular.

Definition 14.4 (Family of mollifiers)**.** A family of mollifiers is a collection

$$\{\phi_\varepsilon\}_{\varepsilon>0} \subset C_c^\infty(\mathbb{R}^n)$$

such that

$$\operatorname{supp}\phi_\varepsilon \subset B_\varepsilon(0), \qquad \int_{\mathbb{R}^n} \phi_\varepsilon(x)\,\mathrm{d}x = 1 \tag{14.5}$$

for every $\varepsilon > 0$.

Lemma 14.5 (Integral representation for regular values)**.** *Let*

$$f \in D_y(U,\mathbb{R}^n) \cap C^1(U,\mathbb{R}^n),$$

and assume that y is a regular value of f. Then there exists $\varepsilon_0 > 0$ such that

$$\deg(f,U,y) = \int_U \phi_\varepsilon(f(x)-y)\,\det(df(x))\,\mathrm{d}x \tag{14.6}$$

for every $0 < \varepsilon < \varepsilon_0$. Moreover, if

$$0 < \varepsilon < \operatorname{dist}(y, f(\partial U)),$$

then

$$\operatorname{supp}\big(\phi_\varepsilon(f(\cdot)-y)\big) \Subset U. \tag{14.7}$$

Proof. Assume first that

$$f^{-1}(y) = \varnothing.$$

Then we can set

$$\varepsilon_0 := \operatorname{dist}(y, f(\overline{U})).$$

Hence

$$\phi_\varepsilon(f(x)-y) = 0 \qquad \forall x \in \overline{U},$$

for every $0 < \varepsilon < \varepsilon_0$, and therefore

$$\int_U \phi_\varepsilon(f(x)-y)\,\det(df(x))\,\mathrm{d}x = 0 = \deg(f,U,y).$$

Assume now that

$$f^{-1}(y) = \{x_1,\dots,x_N\}.$$

Since y is a regular value, the inverse function theorem implies that there exists $\varepsilon_0 > 0$ such that

$$f^{-1}(B_{\varepsilon_0}(y)) = \bigcup_{i=1}^N U(x_i),$$

where the sets $U(x_i)$ are pairwise disjoint neighborhoods of the points x_i. Shrinking ε_0 if necessary, we may also assume that

$$f|_{U(x_i)} : U(x_i) \to B_{\varepsilon_0}(y)$$

is a bijection and that

$$\det(df(x)) \neq 0 \qquad \forall x \in U(x_i), \quad i = 1,\dots,N.$$

For $0 < \varepsilon < \varepsilon_0$ we have

$$\phi_\varepsilon(f(x) - y) = 0 \qquad \forall x \in U \setminus \bigcup_{i=1}^{N} U(x_i),$$

hence

$$\int_U \phi_\varepsilon(f(x) - y)\ \det(df(x))\,\mathrm{d}x = \sum_{i=1}^{N} \int_{U(x_i)} \phi_\varepsilon(f(x) - y)\ \det(df(x))\,\mathrm{d}x. \tag{14.8}$$

Using on each $U(x_i)$ the change of variables

$$\tilde{x} = f(x) - y,$$

we obtain

$$\int_{U(x_i)} \phi_\varepsilon(f(x) - y)\ \det(df(x))\,\mathrm{d}x = \operatorname{sign}\det(df(x_i)) \int_{B_\varepsilon(0)} \phi_\varepsilon(\tilde{x})\,\mathrm{d}\tilde{x} = \operatorname{sign}\det(df(x_i)). \tag{14.9}$$

Summing over i gives

$$\int_U \phi_\varepsilon(f(x) - y)\ \det(df(x))\,\mathrm{d}x = \sum_{i=1}^{N} \operatorname{sign}\det(df(x_i)) = \deg(f, U, y). \tag{14.10}$$

Finally, if

$$0 < \varepsilon < \operatorname{dist}(y, f(\partial U)),$$

then the set where $\phi_\varepsilon(f(x) - y)$ may be nonzero stays a positive distance away from ∂U, which implies

$$\operatorname{supp}\big(\phi_\varepsilon(f(\cdot) - y)\big) \Subset U.$$

□

With this representation in hand, we now want to prove that the degree is locally constant with respect to y, still keeping the regular value assumption. The idea is to write

$$\deg(f, U, y_2) - \deg(f, U, y_1)$$

as the integral of a divergence and conclude by Stokes' theorem.

Lemma 14.6. *Let*

$$f \in C^2(U, \mathbb{R}^n), \qquad u \in C^1(\mathbb{R}^n, \mathbb{R}^n).$$

Then

$$(\operatorname{div} u)(f)\ \det(df) = \operatorname{div} D_f(u), \tag{14.11}$$

where $D_f(u) \in C^1(U, \mathbb{R}^n)$ is defined componentwise by

$$D_f(u)_j := \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, u(f), \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big). \tag{14.12}$$

Proof. We compute the divergence of $D_f(u)$ explicitly. Recall that

$$D_f(u)_j = \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, u(f), \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big). \tag{14.13}$$

Hence

$$\operatorname{div} D_f(u) = \sum_{j=1}^{n} \partial_{x_j} D_f(u)_j. \tag{14.14}$$

Now fix $j \in \{1, \dots, n\}$. Differentiating the determinant with respect to x_j and using multilinearity with respect to the columns, we obtain

$$\begin{aligned}\partial_{x_j} D_f(u)_j = &\sum_{k=1}^{j-1} \det\big(\partial_{x_1} f, \dots, \partial_{x_j}\partial_{x_k} f, \dots, u(f), \dots, \partial_{x_n} f\big) \\ &+ \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, \partial_{x_j}(u(f)), \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big) \\ &+ \sum_{k=j+1}^{n} \det\big(\partial_{x_1} f, \dots, u(f), \dots, \partial_{x_j}\partial_{x_k} f, \dots, \partial_{x_n} f\big). \end{aligned} \tag{14.15}$$

Summing in j, this gives

$$\begin{aligned}\operatorname{div} D_f(u) = \sum_{j=1}^{n} \Bigg[&\sum_{k=1}^{j-1} \det\big(\partial_{x_1} f, \dots, \partial_{x_j}\partial_{x_k} f, \dots, u(f), \dots, \partial_{x_n} f\big) \\ &+ \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, \partial_{x_j}(u(f)), \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big) \\ &+ \sum_{k=j+1}^{n} \det\big(\partial_{x_1} f, \dots, u(f), \dots, \partial_{x_j}\partial_{x_k} f, \dots, \partial_{x_n} f\big) \Bigg]. \end{aligned} \tag{14.16}$$

We now examine the terms containing second derivatives of f. Let $j \neq k$. Since $f \in C^2(U, \mathbb{R}^n)$, mixed partial derivatives commute:

$$\partial_{x_j}\partial_{x_k} f = \partial_{x_k}\partial_{x_j} f. \tag{14.17}$$

Moreover, the determinant is alternating with respect to the columns. Hence the term coming from $\partial_{x_j} D_f(u)_j$ in which $\partial_{x_j}\partial_{x_k} f$ appears in the k-th column cancels with the term coming from $\partial_{x_k} D_f(u)_k$ in which $\partial_{x_k}\partial_{x_j} f$ appears in the j-th column. More precisely, for $j \neq k$ we have

$$\begin{aligned}&\det\big(\dots, \partial_{x_j}\partial_{x_k} f, \dots, u(f), \dots\big) \\ &\qquad = -\det\big(\dots, u(f), \dots, \partial_{x_k}\partial_{x_j} f, \dots\big). \end{aligned} \tag{14.18}$$

Therefore all off-diagonal terms cancel pairwise.

After these cancellations, only the terms where the derivative hits the column $u(f)$ remain. Thus

$$\operatorname{div} D_f(u) = \sum_{j=1}^{n} \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, \partial_{x_j}(u(f)), \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big). \tag{14.19}$$

Now we use the chain rule. Since $u \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, we have

$$\partial_{x_j}(u(f)) = du(f)\, \partial_{x_j} f. \tag{14.20}$$

Hence

$$\operatorname{div} D_f(u) = \sum_{j=1}^{n} \det\big(\partial_{x_1} f, \dots, \partial_{x_{j-1}} f, du(f)\, \partial_{x_j} f, \partial_{x_{j+1}} f, \dots, \partial_{x_n} f\big). \tag{14.21}$$

We now apply the identity from the next exercise. Let

$$M = df \qquad \text{and} \qquad A = du(f).$$

If $M_1, \dots, M_n$ denote the columns of df, that identity gives

$$\sum_{j=1}^{n} \det\big(M_1, \dots, M_{j-1}, AM_j, M_{j+1}, \dots, M_n\big) = \operatorname{tr}(A) \det(M). \tag{14.22}$$

Therefore,

$$\operatorname{div} D_f(u) = \operatorname{tr}(du(f))\, \det(df). \tag{14.23}$$

Finally,

$$\operatorname{tr}(du(f)) = (\operatorname{div} u)(f), \tag{14.24}$$

and thus

$$\operatorname{div} D_f(u) = (\operatorname{div} u)(f)\ \det(df). \tag{14.25}$$

This proves the claim. □

Exercise 14.7. Let $A, M \in \mathbb{R}^{n\times n}$, and write $M_1, \dots, M_n$ for the columns of M. Prove that

$$\sum_{j=1}^{n} \det\big(M_1, \dots, M_{j-1}, AM_j, M_{j+1}, \dots, M_n\big) = \operatorname{tr}(A)\det(M). \tag{14.26}$$

Hint. Define

$$\psi(t) := \det(M + tAM),$$

and compute $\psi'(0)$ in two different ways. □

To remove the hypothesis that y is a regular value we will use the following theorem.

Theorem 14.8 (Sard)**.** *Let*

$$f \in C^1(U, \mathbb{R}^n).$$

Then the set of critical values of f has Lebesgue measure zero.

Lemma 14.9. *Let*

$$f \in C^2(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n).$$

Then $\deg(f, U, \cdot)$ *is constant in each ball contained in* $\mathbb{R}^n \setminus f(\partial U)$*, whenever it is defined by the determinant formula.*

Proof. Fix

$$\tilde{y} \in \mathbb{R}^n \setminus f(\partial U), \qquad \rho := \operatorname{dist}(\tilde{y}, f(\partial U)).$$

Consider the largest ball $B_\rho(\tilde{y})$ contained in

$$\mathbb{R}^n \setminus f(\partial U).$$

Take two regular values

$$y^1, y^2 \in B_\rho(\tilde{y}).$$

For $\varepsilon > 0$ small enough, Lemma above gives

$$\deg(f, U, y^2) - \deg(f, U, y^1) = \int_U \big(\phi_\varepsilon(f(x) - y^2) - \phi_\varepsilon(f(x) - y^1)\big)\det(df(x))\,\mathrm{d}x. \tag{14.27}$$

Set

$$\phi(y) := \phi_\varepsilon(y - y^1), \qquad z := y^1 - y^2, \tag{14.28}$$

and define

$$u(y) := z\int_0^1 \phi(y + tz)\,\mathrm{d}t. \tag{14.29}$$

Then

$$\operatorname{div} u(y) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\phi(y + tz)\,\mathrm{d}t \tag{14.30}$$

$$= \phi(y + z) - \phi(y) \tag{14.31}$$

$$= \phi_\varepsilon(y - y^2) - \phi_\varepsilon(y - y^1). \tag{14.32}$$

Therefore

$$\deg(f, U, y^2) - \deg(f, U, y^1) = \int_U (\operatorname{div} u)(f(x))\ \det(df(x))\,\mathrm{d}x. \tag{14.33}$$

By the previous lemma,

$$(\operatorname{div} u)(f)\ \det(df) = \operatorname{div} D_f(u), \tag{14.34}$$

hence

$$\deg(f, U, y^2) - \deg(f, U, y^1) = \int_U \operatorname{div} D_f(u)\,\mathrm{d}x. \tag{14.35}$$

If $\varepsilon > 0$ is chosen so small that

$$\varepsilon < \rho - \max\{|y^1 - \tilde{y}|,\ |y^2 - \tilde{y}|\}, \tag{14.36}$$

then u is supported inside $B_\rho(\tilde{y})$, and consequently $D_f(u)$ vanishes in a neighborhood of ∂U. Extending it by zero outside U and integrating over a cube $Q \supset U$, we get

$$\int_U \operatorname{div} D_f(u)\,\mathrm{d}x = \int_Q \operatorname{div} D_f(u)\,\mathrm{d}x = 0 \tag{14.37}$$

by Stokes' theorem. Thus

$$\deg(f, U, y^2) = \deg(f, U, y^1). \tag{14.38}$$

□

Remark 14.10. As a consequence, if

$$f \in C^2(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n) \qquad \text{and} \qquad y \notin f(\partial U),$$

we may define $\deg(f, U, y)$ by approximation with nearby regular values. Indeed, choose a regular value $\tilde{y}$ such that

$$|\tilde{y} - y| < \operatorname{dist}(y, f(\partial U)),$$

and set

$$\deg(f, U, y) := \deg(f, U, \tilde{y}). \tag{14.39}$$

This is well defined by the previous lemma, and Sard's theorem guarantees that such regular values exist arbitrarily close to y.

15 The Brouwer mapping degree 3 - 21/04/2026

15.1 Step 2: admitting continuous functions.

We now remove the C^2 assumption. The idea is the following. Fix

$$f \in D_y(U, \mathbb{R}^n), \qquad \rho := \operatorname{dist}(y, f(\partial U)) > 0,$$

and choose

$$f_1, f_2 \in C^2(U, \mathbb{R}^n)$$

such that

$$\|f_i - f\|_\infty < \rho, \qquad i = 1, 2.$$

Then

$$H(t, x) := (1-t)f_1(x) + tf_2(x) \in D_y(U, \mathbb{R}^n) \cap C^2(U, \mathbb{R}^n), \qquad t \in [0, 1].$$

If we prove that $t \mapsto \deg(H(t), U, y)$ is locally constant, then it is constant on $[0, 1]$. This will show that the degree obtained from smooth approximations does not depend on the chosen approximation.

Lemma 15.1. *Let*

$$f \in C^2(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n).$$

Then for every

$$g \in C^2(U, \mathbb{R}^n)$$

there exists $\varepsilon > 0$ such that

$$\deg(f + tg, U, y) = \deg(f, U, y) \tag{15.1}$$

for every $t \in (-\varepsilon, \varepsilon)$.

Proof. We first consider the case where y is a regular value of f. If

$$f^{-1}(y) = \varnothing,$$

then the same remains true for $f + tg$ whenever

$$|t| \, \|g\|_\infty < \operatorname{dist}(y, f(\overline{U})).$$

Hence the degree is zero for all such t.

Assume now that

$$f^{-1}(y) = \{x_1, \dots, x_N\}.$$

By the implicit function theorem there exist pairwise disjoint neighborhoods $U(x_i)$ and $\varepsilon_1 > 0$ such that, for every $|t| < \varepsilon_1$, the equation

$$(f + tg)(x) = y$$

has a unique solution in each $U(x_i)$. Shrinking the sets $U(x_i)$ if necessary, we may also assume that the sign of

$$\det d(f + tg)$$

is constant on each $U(x_i)$ for all $|t| < \varepsilon_1$.

Next set

$$\varepsilon_2 := \frac{\operatorname{dist}\big(y, f(\overline{U} \setminus \bigcup_{i=1}^N U(x_i))\big)}{\|g\|_\infty}, \tag{15.2}$$

with the usual convention that if $g = 0$ the statement is trivial. Then, for $|t| < \varepsilon_2$, one has

$$|(f + tg)(x) - y| > 0 \qquad \forall x \in \overline{U} \setminus \bigcup_{i=1}^N U(x_i).$$

Hence no new zeros appear outside the chosen neighborhoods. Therefore the number of zeros, counted with the sign of the Jacobian, does not change, and so

$$\deg(f + tg, U, y) = \deg(f, U, y)$$

for $|t| < \varepsilon$, where

$$\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}.$$

It remains to consider the case where y is a critical value of f. Set

$$\rho := \operatorname{dist}(y, f(\partial U)).$$

By Sard's theorem there exists a regular value $\tilde{y} \in B_{\rho/3}(y)$. Then

$$\deg(f, U, y) = \deg(f, U, \tilde{y}).$$

By the regular-value case there exists $\tilde{\varepsilon} > 0$ such that

$$\deg(f, U, \tilde{y}) = \deg(f + tg, U, \tilde{y})$$

for all $|t| < \tilde{\varepsilon}$.

Now set

$$\varepsilon := \min\left\{\tilde{\varepsilon}, \frac{\rho}{3\|g\|_\infty}\right\}. \tag{15.3}$$

For $x \in \partial U$ and $|t| < \varepsilon$ we have

$$|\tilde{y} - (f + tg)(x)| \geq |y - f(x)| - |\tilde{y} - y| - |t|\,|g(x)| \tag{15.4}$$

$$\geq \rho - \frac{\rho}{3} - \frac{\rho}{3} \tag{15.5}$$

$$= \frac{\rho}{3} > 0. \tag{15.6}$$

Thus

$$|\tilde{y} - y| < \operatorname{dist}(\tilde{y}, (f + tg)(\partial U)),$$

and by Step 1,

$$\deg(f + tg, U, \tilde{y}) = \deg(f + tg, U, y).$$

Combining the above equalities yields

$$\deg(f + tg, U, y) = \deg(f, U, y)$$

for all $|t| < \varepsilon$. □

Definition 15.2 (Bröwer degree)**.** We can now define the degree for every continuous map. Let

$$f \in D_y(U, \mathbb{R}^n), \qquad \rho := \operatorname{dist}(y, f(\partial U)).$$

Choose

$$\tilde{f} \in C^2(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n)$$

such that

$$\|f - \tilde{f}\|_\infty < \rho.$$

We define

$$\deg(f, U, y) := \deg(\tilde{f}, U, y). \tag{15.7}$$

To check that this is well posed, let $\tilde{f}_1, \tilde{f}_2$ be two such approximations and consider

$$H(t) := (1-t)\tilde{f}_1 + t\tilde{f}_2.$$

Then

$$\|H(t) - f\|_\infty < \rho \qquad \forall t \in [0,1],$$

hence $H(t) \in D_y(U, \mathbb{R}^n)$ for every t. By the previous lemma, $t \mapsto \deg(H(t), U, y)$ is locally constant, hence constant on the connected interval $[0,1]$. Therefore

$$\deg(\tilde{f}_1, U, y) = \deg(\tilde{f}_2, U, y),$$

and the definition is independent of the chosen approximation.

Theorem 15.3. *There exists a unique degree map satisfying the axioms* (D1)–(D4)*. Moreover,*

$$\deg(\cdot, U, y) : D_y(U, \mathbb{R}^n) \to \mathbb{Z}$$

is constant on each connected component, and for every

$$f \in D_y(U, \mathbb{R}^n)$$

one has the explicit formula

$$\deg(f, U, y) = \sum_{x \in \tilde{f}^{-1}(y)} \operatorname{sign} J\tilde{f}(x), \tag{15.8}$$

where

$$\tilde{f} \in C^2(U, \mathbb{R}^n) \cap D_y(U, \mathbb{R}^n), \qquad \|f - \tilde{f}\|_\infty < \operatorname{dist}(y, f(\partial U)),$$

and y is a regular value of $\tilde{f}$.

Proof. The previous discussion shows that the degree is well defined. By construction, it is locally constant with respect to the first argument. Since $\mathbb{Z}$ is discrete, it follows that

$$\deg(\cdot, U, y) : D_y(U, \mathbb{R}^n) \to \mathbb{Z}$$

is constant on connected components.

The normalization axiom (D2) is immediate. Translation invariance (D1) and additivity (D3) hold because they hold for smooth approximations and the definition is independent of the approximation. Finally, homotopy invariance (D4) follows from local constancy.

Uniqueness follows as well: any degree map satisfying (D1)–(D4) must agree with the determinant formula at regular values, and the above construction extends that formula uniquely to all of $D_y(U, \mathbb{R}^n)$. □

Exercise 15.4. Investigate the zeros of

$$f(x_1, x_2) = (x_1 - 2x_2 + \cos(x_1 + x_2), x_2 + 2x_1 + \sin(x_1 + x_2)). \tag{15.9}$$

15.2 Application: Bröwer fixed point

Theorem 15.5 (Brouwer fixed-point theorem)**.** *Let $U \subset \mathbb{R}^n$ be an open set such that $\overline{U}$ is homeomorphic to $\overline{B_1(0)}$, and let*

$$f : \overline{U} \to \overline{U}$$

be continuous. Then there exists $x \in \overline{U}$ such that $f(x) = x$.

Proof. Let

$$\phi : \overline{U} \to \overline{B_1(0)}$$

be a homeomorphism, and define

$$g := \phi \circ f \circ \phi^{-1} : \overline{B_1(0)} \to \overline{B_1(0)}.$$

Then g is continuous. Moreover, $x \in \overline{U}$ is a fixed point of f if and only if $\phi(x) \in \overline{B_1(0)}$ is a fixed point of g. Hence it is enough to prove the result for $\overline{U} = \overline{B_1(0)}$, that is, for a continuous map

$$g : \overline{B_1(0)} \to \overline{B_1(0)}.$$

If there exists $x \in \partial B_1(0)$ such that $g(x) = x$, then we are done.

Assume therefore that

$$g(x) \neq x \qquad \text{for every } x \in \partial B_1(0).$$

We claim that g must then have a fixed point in $B_1(0)$.

Consider the homotopy

$$H : [0,1] \times \overline{B_1(0)} \to \mathbb{R}^n, \qquad H(t,x) := x - tg(x). \tag{15.10}$$

We show that

$$H(t,x) \neq 0 \qquad \text{for every } (t,x) \in [0,1] \times \partial B_1(0).$$

Indeed, if $H(t,x) = 0$ for some $x \in \partial B_1(0)$, then $x = tg(x)$. Taking norms and using $g(x) \in \overline{B_1(0)}$, we get

$$1 = |x| = t|g(x)| \leq t \leq 1.$$

Hence $t = 1$ and $|g(x)| = 1$, so $x = g(x)$, contrary to our assumption on $\partial B_1(0)$.

Therefore

$$H(t,\cdot) \in D_0(B_1(0), \mathbb{R}^n) \qquad \text{for every } t \in [0,1].$$

By homotopy invariance of the degree,

$$\deg(I - g, B_1(0), 0) = \deg(I, B_1(0), 0) = 1.$$

Thus g has a fixed point in $B_1(0)$, and therefore f has a fixed point in $\overline{U}$. □

Corollary 15.6 (Brouwer fixed point on compact convex sets)**.** *Let $C \subset \mathbb{R}^n$ be nonempty, compact, and convex, and let*

$$f : C \to C$$

be continuous. Then there exists $x \in C$ such that

$$f(x) = x.$$

Remark 15.7. More generally, if K is a topological space homeomorphic to a compact convex subset of $\mathbb{R}^n$, then every continuous map $f : K \to K$ has a fixed point. Indeed, homeomorphisms preserve fixed points, so one reduces immediately to Corollary 15.6.

Theorem 15.8 (Tietze extension)**.** *Let K be a compact subset of a metric space X, and let Y be a Banach space. Then, any function $F \in C(K,Y)$ admits and extension $\overline{F} \in C(X,Y)$.*

Theorem 15.9. *Let $U \subset \mathbb{R}^n$ be a bounded open set with $0 \in U$, and assume that n is odd. Let*

$$f : \partial U \to \mathbb{R}^n$$

be continuous and suppose that $0 \notin f(\partial U)$. Then there exist $x \in \partial U$ and $\lambda \in \mathbb{R} \setminus \{0\}$ such that

$$f(x) = \lambda x.$$

Proof. By Tietze's Theorem, let

$$\overline{f} \in C(\overline{U}, \mathbb{R}^n)$$

be a continuous extension of f to $\overline{U}$. Assume by contradiction that

$$f(x) \neq \lambda x \qquad \text{for every } x \in \partial U,\ \lambda \in \mathbb{R}.$$

Since $0 \notin f(\partial U)$, this assumption already excludes the case $\lambda = 0$.

Consider the two homotopies

$$H_+(t,x) := (1-t)\overline{f}(x) + tx, \qquad (t,x) \in [0,1] \times \overline{U}, \tag{15.11}$$

and

$$H_-(t,x) := (1-t)\overline{f}(x) - tx, \qquad (t,x) \in [0,1] \times \overline{U}. \tag{15.12}$$

We claim that both homotopies avoid 0 on $[0,1] \times \partial U$.

For H_+: if $H_+(t,x) = 0$ for some $x \in \partial U$, then

$$(1-t)f(x) = -tx.$$

If $t = 1$, then $x = 0$, impossible since $0 \in U$ and hence $0 \notin \partial U$. If $t \in [0,1)$, then

$$f(x) = -\frac{t}{1-t}\, x,$$

contradicting our assumption.

Similarly, for H_-: if $H_-(t,x) = 0$, then either $t = 1$ and again $x = 0$, impossible, or $t \in [0,1)$ and

$$f(x) = \frac{t}{1-t}\, x,$$

again a contradiction.

Therefore

$$H_\pm(t,\cdot) \in D_0(U, \mathbb{R}^n) \qquad \text{for every } t \in [0,1].$$

Thus $\overline{f}$ is homotopic both to I and to $-I$ in $D_0(U, \mathbb{R}^n)$. By homotopy invariance,

$$\deg(\overline{f}, U, 0) = \deg(I, U, 0) = 1 \tag{15.13}$$

and also

$$\deg(\overline{f}, U, 0) = \deg(-I, U, 0). \tag{15.14}$$

Since 0 is the unique zero of $-I$ and

$$\det(-I) = (-1)^n,$$

the determinant formula gives

$$\deg(-I, U, 0) = (-1)^n = -1,$$

because n is odd. Hence

$$1 = \deg(I, U, 0) = \deg(\overline{f}, U, 0) = \deg(-I, U, 0) = -1,$$

a contradiction. □

Corollary 15.10 (Hairy ball theorem)**.** *Assume that n is odd. Let*

$$V : \mathbb{S}^{n-1} \to \mathbb{R}^n$$

be a continuous tangent vector field on $\mathbb{S}^{n-1}$, namely

$$V(x) \in T_x\mathbb{S}^{n-1} = \{v \in \mathbb{R}^n : v \cdot x = 0\} \qquad \textit{for every } x \in \mathbb{S}^{n-1}.$$

Then there exists $x_0 \in \mathbb{S}^{n-1}$ such that

$$V(x_0) = 0.$$

Proof. Assume by contradiction that

$$V(x) \neq 0 \qquad \text{for every } x \in \mathbb{S}^{n-1}.$$

Apply Theorem 15.9 with

$$U = B_1(0), \qquad f = V\big|_{\partial B_1(0)}.$$

Then there exist $x \in \mathbb{S}^{n-1}$ and $\lambda \neq 0$ such that

$$V(x) = \lambda x.$$

But $V(x) \in T_x\mathbb{S}^{n-1}$, so

$$0 = V(x) \cdot x = \lambda |x|^2 = \lambda,$$

a contradiction. □

Exercise 15.11 (Outward pointing vector field)**.** Let $R > 0$, pick $f \in C\big(\overline{B_R(0)}, \mathbb{R}^n\big)$ and assume that

$$f(x) \cdot x > 0 \qquad \text{for every } x \in \partial B_R(0).$$

Then there exists $x \in B_R(0)$ such that $f(x) = 0$.

Exercise 15.12 (Fundamental theorem of algebra)**.** Let

$$P(z) = z^p + c_1 z^{p-1} + \cdots + c_p, \qquad p \geq 1,$$

with $c_1, \ldots, c_p \in \mathbb{C}$. Then there exists $z_0 \in \mathbb{C}$ such that

$$P(z_0) = 0.$$

16 The Brouwer mapping degree 4 - 27/04/2026

Theorem 16.1 (Borsuk odd mapping)**.** *Let $U \subset \mathbb{R}^n$ be open, bounded, symmetric, with $0 \in U$. Let $f \in C(\overline{U})$ be odd and assume that $0 \notin f(\partial U)$. Then $\deg(f, U, 0)$ is odd.*

Proof. We may assume that $f \in C^1(\overline{U})$ and $\det df(0) \neq 0$. To see this, approximate $f \in C(\overline{U})$ by $g_1 \in C^1(\overline{U})$, consider the odd part

$$g_2(x) := \frac{1}{2}\big(g_1(x) - g_1(-x)\big),$$

and choose a δ which is not an eigenvalue of $g_2'(0)$. Then

$$\widetilde{f} = g_2 - \delta \cdot \mathrm{id}$$

is in $C^1(\overline{U})$, odd, with $\det d\widetilde{f}(0) \neq 0$, and close to f if δ and $\|g_1 - f\|_0$ are chosen sufficiently small. Hence

$$\deg(f, U, 0) = \deg(\widetilde{f}, U, 0).$$

Let $f \in C^1(\overline{U})$ and $\det df(0) \neq 0$. To prove the theorem, it suffices to show that there is an odd $g \in C^1(\overline{U})$ sufficiently close to f such that 0 is not a critical value of g in U, since then

$$\deg(f, U, 0) = \deg(g, U, 0) = \operatorname{sgn}\det dg(0) + \sum_{0 \neq x \in g^{-1}(0)} \operatorname{sgn}\det dg(x),$$

where the sum is even since $g(x) = 0$ iff $g(-x) = 0$ and $\det dg(\cdot)$ is even.

Such a map g will be defined by induction as follows. Consider

$$U_k := \{x \in U : \ x_i \neq 0 \text{ for some } i \leq k\}$$

and an odd $\varphi \in C^1(\mathbb{R})$ such that

$$\varphi'(0) = 0 \qquad \text{and} \qquad \varphi(t) = 0 \iff t = 0.$$

Consider

$$\widetilde{f}(x) = \frac{f(x)}{\varphi(x_1)}$$

on the open bounded $U_1 = \{x \in U : \ x_1 \neq 0\}$. By Sard Theorem, we find a regular value y^1 in $\widetilde{f}(U_1)$ such that $|y^1|$ is as small as necessary in the sequel. Hence, 0 is a regular value for

$$g_1(x) := f(x) - \varphi(x_1)y^1$$

on U_1, since

$$g_1(x) = \varphi(x_1)\big(\widetilde{f}(x) - y^1\big)$$

for $x \in U_1$ such that $g_1(x) = 0$.

Now, suppose that we have already an odd $g_k \in C^1(\overline{U})$ close to f on $\overline{U}$ such that 0 is a regular value in $g_k(U_n)$ for some $k < n$. Then we define

$$g_{k+1}(x) = g_k(x) - \varphi(x_{k+1})y^{k+1}$$

with $|y^{k+1}|$ small and such that 0 is a regular value for g_{k+1} on $\{x \in U : \ x_{k+1} \neq 0\}$.

Evidently, $g_{k+1} \in C^1(\overline{U})$ is odd and close to f on $\overline{U}$. If $x \in U_{k+1}$ and $x_{k+1} = 0$, then $x \in U_k$,

$$g_{k+1}(x) = g_k(x) \qquad \text{and} \qquad dg_{k+1}(x) = dg_k(x),$$

hence $\det dg_{k+1}(x) \neq 0$, and therefore 0 is a regular value in $g_{k+1}(U_{k+1})$. Thus, $g = g_n$ is odd, close to f on $\overline{U}$, and such that 0 is a regular value for g in $U \setminus \{0\}$, $U_n = U \setminus \{0\}$. By the induction step you see that we also have

$$dg(0) = dg_1(0) = df(0);$$

hence 0 is a regular value for g in the whole U. □

Exercise 16.2 (Even mapping theorem)**.** Let $U \subset \mathbb{R}^n$ be open, bounded, and symmetric with respect to the origin, with $0 \in U$. Let

$$f : \partial U \to \mathbb{R}^n \setminus \{0\} \tag{16.1}$$

be a continuous even map, that is,

$$f(-x) = f(x) \qquad \text{for every } x \in \partial U. \tag{16.2}$$

If $F \in C(\overline{U}, \mathbb{R}^n)$ is any continuous extension of f, then $\deg(F, U, 0)$ is even whenever n is even, and $\deg(F, U, 0) = 0$ whenever n is odd.

Corollary 16.3 (Borsuk–Ulam Theorem)**.** *Let $f : \mathbb{S}^n \to \mathbb{R}^n$ be continuous. Then there exists a point $x \in \mathbb{S}^n$ such that*

$$f(x) = f(-x).$$

Proof. Suppose, by contradiction, that

$$g(x) := f(x) - f(-x) \neq 0 \qquad \text{for every } x \in \mathbb{S}^n.$$

Define $G : \overline{B_1(0)} \to \mathbb{R}^n$ by

$$G(0) := 0, \qquad G(x) := |x|\, g\left(\frac{x}{|x|}\right) \quad \text{for } x \neq 0.$$

Then G is continuous on $\overline{B_1(0)}$ and $G|_{\mathbb{S}^n} = g$. Now define

$$\widetilde{G} : \overline{B_1(0)} \to \mathbb{R}^{n+1}, \qquad \widetilde{G}(x) := (G(x), 0).$$

Then $\widetilde{G}$ is continuous, $\widetilde{G}(x) \neq 0$ for every $x \in \mathbb{S}^n$, and

$$\widetilde{G}(-x) = (-G(x), 0) = -\widetilde{G}(x) \qquad \text{for every } x \in \mathbb{S}^n.$$

Thus $\widetilde{G}|_{\mathbb{S}^n}$ is odd and hence the odd mapping theorem proved above applies and yields

$$\deg(\widetilde{G}, B_1(0), 0) \neq 0.$$

Moreover, $\widetilde{G}(\mathbb{S}^n)$ is compact and does not contain 0. Therefore there exists $\rho > 0$ such that

$$B_\rho(0) \cap \widetilde{G}(\mathbb{S}^n) = \varnothing.$$

By local constancy of the degree,

$$\deg(\widetilde{G}, B_1(0), y) = \deg(\widetilde{G}, B_1(0), 0) \neq 0 \qquad \text{for every } y \in B_\rho(0).$$

Hence, by the existence property of the degree,

$$B_\rho(0) \subset \widetilde{G}(B_1(0)) \subset \mathbb{R}^n \times \{0\}.$$

This is impossible, since $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+1}$ contains no nonempty open ball of $\mathbb{R}^{n+1}$.

The contradiction shows that there exists $x \in \mathbb{S}^n$ such that $g(x) = 0$, that is, $f(x) = f(-x)$. □

Theorem 16.4 (Invariance of dimension)**.** *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be continuous and injective, with U open and bounded. Then $f(U)$ is open.*

Proof. Fix $x_0 \in U$. It is enough to show that there exists $r > 0$ such that $f(B_r(x_0))$ contains a ball centered at $f(x_0)$. By translating the domain and the range, we may assume

$$x_0 = 0, \qquad f(0) = 0.$$

Choose $r > 0$ such that $\overline{B_r(0)} \subset U$. Since f is injective on U, it is in particular injective on $B_r(0)$.

Consider the homotopy

$$h : [0,1] \times \overline{B_r(0)} \to \mathbb{R}^n, \qquad h(t,x) := f\left(\frac{x}{1+t}\right) - f\left(-\frac{t}{1+t}x\right).$$

It is continuous in (t,x), and

$$h(0,x) = f(x), \qquad h(1,x) = f\left(\frac{x}{2}\right) - f\left(-\frac{x}{2}\right).$$

In particular, $h(1,\cdot)$ is odd on $\partial B_r(0)$.

We claim that

$$h(t,x) \neq 0 \qquad \text{for every } (t,x) \in [0,1] \times \partial B_r(0).$$

Indeed, if $h(t,x) = 0$, then by injectivity of f,

$$\frac{x}{1+t} = -\frac{t}{1+t}x,$$

hence

$$x = -tx.$$

Therefore

$$(1+t)x = 0.$$

Since $t \in [0,1]$, we have $1 + t > 0$, and thus $x = 0$, contradicting $x \in \partial B_r(0)$.

Hence the homotopy h never vanishes on the boundary. By homotopy invariance of the degree,

$$\deg(f, B_r(0), 0) = \deg(h(0,\cdot), B_r(0), 0) = \deg(h(1,\cdot), B_r(0), 0).$$

Since $h(1,\cdot)$ is odd on $\partial B_r(0)$, the odd mapping theorem implies that

$$\deg(h(1,\cdot), B_r(0), 0)$$

is odd, in particular nonzero. Therefore

$$\deg(f, B_r(0), 0) \neq 0.$$

Now $0 \notin f(\partial B_r(0))$, because $f(0) = 0$ and f is injective. Since $f(\partial B_r(0))$ is compact, there exists $\delta > 0$ such that

$$B_\delta(0) \cap f(\partial B_r(0)) = \varnothing.$$

By local constancy of the degree,

$$\deg(f, B_r(0), y) = \deg(f, B_r(0), 0) \neq 0 \qquad \text{for every } y \in B_\delta(0).$$

Hence, by the existence property,

$$B_\delta(0) \subset f(B_r(0)) \subset f(U).$$

Thus $0 = f(x_0)$ is an interior point of $f(U)$. Since $x_0 \in U$ was arbitrary, $f(U)$ is open. □

Exercise 16.5. Let $m < n$ and let $U \subseteq \mathbb{R}^n$ be open. Then there is no continuous injective map $f : U \to \mathbb{R}^m$.

Exercise 16.6. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 map. Assume $\det df(x) \neq 0$ for all $x \in \mathbb{R}^n$ and that $|f(x)| \to \infty$ as $|x| \to \infty$. Prove that f is surjective.

17 The Leray-Schauder degree 1 - 04/05/2026

We would like to extend the notion of degree for maps between Banach spaces.

Example 17.1. Let $F : \overline{B_1(0)} \subseteq \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ given by

$$F(x) = (\sqrt{1 - \|x\|_2^2}, x_1, x_2, \dots). \tag{17.1}$$

We have that $F(\overline{B_1(0)}) \subseteq \overline{B_1(0)}$ but there is no fixed point, unlike in the finite dimensional case via the Bröwer fixed point Theorem. Indeed, if $F(x) = x$ and $x \in \partial B_1(0)$ then $x = 0$, while if $x \in B_1(0)$ then $F(x) \in \partial B_1(0)$. Hence, no fixed point is possible.

We start to extend the mapping degree from $\mathbb{R}^n$ to general Banach spaces. Naturally, we will first consider the finite dimensional case.

Proposition 17.2 (Degree on finite-dimensional Banach spaces)**.** *Let X be a real Banach space with $\dim X = n$, and let $U \subset X$ be open and bounded. Let*

$$f \in C(\overline{U}, X), \qquad y \in X, \qquad y \notin f(\partial U).$$

If $\phi : X \to \mathbb{R}^n$ is a linear isomorphism, define

$$\deg(f, U, y) := \deg\big(\phi \circ f \circ \phi^{-1}, \phi(U), \phi(y)\big). \tag{17.2}$$

Then this definition is independent of the choice of the isomorphism ϕ. Therefore the Brouwer degree is well-defined on finite-dimensional real Banach spaces.

Proof. Let $\phi, \psi : X \to \mathbb{R}^n$ be two linear isomorphisms. We prove that

$$\deg\big(\phi \circ f \circ \phi^{-1}, \phi(U), \phi(y)\big) = \deg\big(\psi \circ f \circ \psi^{-1}, \psi(U), \psi(y)\big).$$

Set $A := \psi \circ \phi^{-1} \in GL(n)$, and abbreviate

$$f_\phi := \phi \circ f \circ \phi^{-1}, \qquad y_\phi := \phi(y).$$

Then

$$\psi \circ f \circ \psi^{-1} = A \circ f_\phi \circ A^{-1}, \qquad \psi(U) = A(\phi(U)), \qquad \psi(y) = Ay_\phi.$$

We first prove the claim in the case where $f_\phi \in C^1(\phi(\overline{U}), \mathbb{R}^n)$ and y_ϕ is a regular value of f_ϕ. Then Ay_ϕ is a regular value of $A \circ f_\phi \circ A^{-1}$. Indeed, if

$$A \circ f_\phi \circ A^{-1}(z) = Ay_\phi,$$

then $z = Ax$ for some $x \in f_\phi^{-1}(y_\phi)$, and

$$d(A \circ f_\phi \circ A^{-1})(Ax) = A \circ df_\phi(x) \circ A^{-1}.$$

Therefore

$$\begin{aligned} \det d(A \circ f_\phi \circ A^{-1})(Ax) &= \det\big(A \circ df_\phi(x) \circ A^{-1}\big) && (17.3)\\ &= \det(A)\det(df_\phi(x))\det(A^{-1}) && (17.4)\\ &= \det(df_\phi(x)). && (17.5) \end{aligned}$$

Hence the sign of the determinant is preserved at corresponding preimages. Using the determinant formula for the degree, we obtain

$$\begin{aligned} \deg\big(A \circ f_\phi \circ A^{-1}, A(\phi(U)), Ay_\phi\big) &= \sum_{z \in (A \circ f_\phi \circ A^{-1})^{-1}(Ay_\phi)} \operatorname{sgn}\det d(A \circ f_\phi \circ A^{-1})(z) && (17.6)\\ &= \sum_{x \in f_\phi^{-1}(y_\phi)} \operatorname{sgn}\det(df_\phi(x)) && (17.7)\\ &= \deg(f_\phi, \phi(U), y_\phi). && (17.8) \end{aligned}$$

We now pass to the general continuous case. By the definition of the Brouwer degree in $\mathbb{R}^n$, choose $\widetilde{f}_\phi \in C^1(\phi(\overline{U}), \mathbb{R}^n)$ in the same connected component of

$$D_{y_\phi}\big(\phi(U), \mathbb{R}^n\big) := \{h \in C(\phi(\overline{U}), \mathbb{R}^n) : y_\phi \notin h(\partial\phi(U))\}$$

as f_ϕ, such that y_ϕ is a regular value of $\widetilde{f}_\phi$. Then $A \circ \widetilde{f}_\phi \circ A^{-1}$ belongs to the same connected component of

$$D_{Ay_\phi}\big(A(\phi(U)), \mathbb{R}^n\big)$$

as $A \circ f_\phi \circ A^{-1}$. Indeed, if H is a homotopy in $D_{y_\phi}(\phi(U), \mathbb{R}^n)$ joining f_ϕ to $\widetilde{f}_\phi$, then

$$(t, z) \mapsto A\, H(t, A^{-1}z)$$

is a homotopy in $D_{Ay_\phi}(A(\phi(U)), \mathbb{R}^n)$ joining $A \circ f_\phi \circ A^{-1}$ to $A \circ \widetilde{f}_\phi \circ A^{-1}$. By homotopy invariance and by the already proved regular case,

$$\begin{aligned}
\deg\big(A \circ f_\phi \circ A^{-1}, A(\phi(U)), Ay_\phi\big) &= \deg\big(A \circ \widetilde{f}_\phi \circ A^{-1}, A(\phi(U)), Ay_\phi\big) && (17.9)\\
&= \deg\big(\widetilde{f}_\phi, \phi(U), y_\phi\big) && (17.10)\\
&= \deg\big(f_\phi, \phi(U), y_\phi\big). && (17.11)
\end{aligned}$$

Since $A \circ f_\phi \circ A^{-1} = \psi \circ f \circ \psi^{-1}$, this proves

$$\deg\big(\psi \circ f \circ \psi^{-1}, \psi(U), \psi(y)\big) = \deg\big(\phi \circ f \circ \phi^{-1}, \phi(U), \phi(y)\big).$$

Thus the definition (17.2) is independent of the chosen isomorphism. □

Now, for the infinite dimensional case, we try to mimic the approximation of continuos functions with smooth functions. Namely, we approximate F with maps with finite dimensional operators.

17.1 Compact mappings

Definition 17.3 (Compact mappings)**.** Let X, Y be Banach spaces, and let $U \subset X$ be open. A map $F : U \to Y$ is called compact if it is continuous and, for every bounded set $B \subset U$, the set $F(B)$ is precompact in Y, that is, $\overline{F(B)}$ is compact in Y. We denote by $\mathcal{K}(U, Y)$ the space of compact mappings from U into Y.

Example 17.4 (Compact linear operators)**.** Let $T \in \mathcal{L}(X, Y)$. Then T is compact as a linear operator if and only if the map $x \mapsto Tx$ is compact in the sense of the previous definition.

Exercise 17.5 (An integral operator is compact)**.** Let $X = C^0([0, 1])$, and let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Define $F : X \to X$ by

$$F(u)(t) = \int_0^t f(u(s))\,\mathrm{d}s. \tag{17.12}$$

Then F is compact.

Example 17.6 (Finite-dimensional image)**.** Let $F : X \to Y$ be continuous and bounded, and assume that

$$\dim \operatorname{span} F(X) < \infty.$$

Then F is compact. Indeed, if $B \subset X$ is bounded, then $F(B)$ is bounded and contained in the finite-dimensional subspace

$$Y_0 := \operatorname{span} F(X).$$

Since bounded subsets of finite-dimensional normed spaces are precompact, $F(B)$ is precompact in Y_0, and therefore also in Y.

Proposition 17.7 (Uniform limit of compact mappings)**.** *Let X, Y be Banach spaces, let $U \subset X$ be open, and let $F_n : U \to Y$ be compact mappings. Assume that $\|F_n - F\|_{L^\infty(U,Y)} \to 0$. Then F is compact.*

Proof. First, F is continuous because it is the uniform limit of continuous maps. Let $B \subset U$ be bounded. We prove that $F(B)$ is totally bounded. Since Y is complete, total boundedness of $F(B)$ implies precompactness. Fix $\varepsilon > 0$. Choose $n_\varepsilon \in \mathbb{N}$ such that

$$\|F_{n_\varepsilon} - F\|_{L^\infty(U,Y)} < \frac{\varepsilon}{2}. \tag{17.13}$$

Since F_{n_ε} is compact, $F_{n_\varepsilon}(B)$ is precompact, hence totally bounded. Therefore there exist $y_1, \dots, y_N \in Y$ such that

$$F_{n_\varepsilon}(B) \subset \bigcup_{i=1}^{N} B_{\varepsilon/2}(y_i). \tag{17.14}$$

Let $x \in B$. Then there exists $i \in \{1, \dots, N\}$ such that

$$\|F_{n_\varepsilon}(x) - y_i\|_Y < \frac{\varepsilon}{2}. \tag{17.15}$$

Hence

$$\|F(x) - y_i\|_Y \le \|F(x) - F_{n_\varepsilon}(x)\|_Y + \|F_{n_\varepsilon}(x) - y_i\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \tag{17.16}$$

Thus

$$F(B) \subset \bigcup_{i=1}^{N} B_\varepsilon(y_i). \tag{17.17}$$

So $F(B)$ is totally bounded. Since Y is complete, $\overline{F(B)}$ is compact. Therefore F is compact. □

Corollary 17.8 (Differential of a compact mapping)**.** *Let X, Y be Banach spaces, let $U \subset X$ be open, and let*

$$F \in C^1(U, Y)$$

be compact. Then, for every $x_0 \in U$,

$$dF(x_0) \in \mathcal{L}(X, Y)$$

is a compact linear operator.

Exercise 17.9. Prove the above corollary.

Proposition 17.10 (Approximation by finite-dimensional compact mappings)**.** *Let X, Y be Banach spaces, let $U \subset X$ be open and bounded, and let $F : \overline{U} \to Y$ be compact. Then, for every $\varepsilon > 0$, there exists a compact map $F_\varepsilon : \overline{U} \to Y$ such that*

$$\|F_\varepsilon - F\|_{L^\infty(\overline{U},Y)} < \varepsilon \tag{17.18}$$

and

$$\dim \operatorname{span} F_\varepsilon(\overline{U}) < \infty. \tag{17.19}$$

Proof. Since U is bounded and F is compact, $F(\overline{U})$ is precompact. Hence there exist $y_1, \dots, y_N \in Y$ such that

$$F(\overline{U}) \subset \bigcup_{i=1}^{N} B_\varepsilon(y_i). \tag{17.20}$$

For $y \in F(\overline{U})$ define

$$\varphi_i(y) = \max\{\varepsilon - \|y - y_i\|_Y, 0\}. \tag{17.21}$$

Since the balls $B_\varepsilon(y_i)$ cover $F(\overline{U})$, for every $y \in F(\overline{U})$ there exists i such that $\varphi_i(y) > 0$. Hence

$$\sum_{j=1}^{N} \varphi_j(y) > 0 \qquad \text{for every } y \in F(\overline{U}). \tag{17.22}$$

Set

$$\psi_i(y) = \frac{\varphi_i(y)}{\sum_{j=1}^{N} \varphi_j(y)}. \tag{17.23}$$

Then

$$\psi_i(y) \geq 0, \qquad \sum_{i=1}^{N} \psi_i(y) = 1. \tag{17.24}$$

Define

$$F_\varepsilon(x) = \sum_{i=1}^{N} \psi_i(F(x))\, y_i. \tag{17.25}$$

Then $F_\varepsilon(\overline{U}) \subset \operatorname{span}\{y_1, \dots, y_N\}$. Thus F_ε has finite-dimensional image. Moreover, if $\psi_i(F(x)) \neq 0$, then $\|F(x) - y_i\|_Y < \varepsilon$. Therefore

$$\|F(x) - F_\varepsilon(x)\|_Y = \left\| \sum_{i=1}^{N} \psi_i(F(x))(F(x) - y_i) \right\|_Y \tag{17.26}$$

$$\leq \sum_{i=1}^{N} \psi_i(F(x)) \, \|F(x) - y_i\|_Y \tag{17.27}$$

$$< \varepsilon \sum_{i=1}^{N} \psi_i(F(x)) = \varepsilon. \tag{17.28}$$

Taking the supremum over $x \in \overline{U}$, we obtain

$$\|F_\varepsilon - F\|_{L^\infty(\overline{U},Y)} < \varepsilon.$$

Finally, F_ε is continuous and has finite-dimensional bounded image on bounded sets. Hence F_ε is compact. □

Example 17.11 (Finite-rank approximation of a compact operator on ℓ^2)**.** Let $X = Y = \ell^2(\mathbb{N})$, and define

$$T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N}), \qquad T(x_1, x_2, x_3, \dots) = \left(x_1, \frac{x_2}{2}, \frac{x_3}{3}, \dots, \frac{x_n}{n}, \dots\right).$$

Equivalently,

$$(Tx)_n = \frac{x_n}{n}. \tag{17.29}$$

Then $T \in \mathcal{L}(\ell^2, \ell^2)$ and $\|T\|_{\mathcal{L}(\ell^2,\ell^2)} = 1$. Indeed, for every $x \in \ell^2$,

$$\|Tx\|_{\ell^2}^2 = \sum_{n=1}^{\infty} \frac{|x_n|^2}{n^2} \leq \sum_{n=1}^{\infty} |x_n|^2 = \|x\|_{\ell^2}^2 . \tag{17.30}$$

Thus $\|T\|_{\mathcal{L}(\ell^2,\ell^2)} \leq 1$. Taking $x = e_1$, where $e_1 = (1, 0, 0, \dots)$, gives $\|Te_1\|_{\ell^2} = 1$, hence $\|T\|_{\mathcal{L}(\ell^2,\ell^2)} = 1$. For $N \in \mathbb{N}$, define the finite-rank operator $T_N : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ by

$$T_N(x_1, x_2, x_3, \dots) = \left(x_1, \frac{x_2}{2}, \dots, \frac{x_N}{N}, 0, 0, \dots\right).$$

Then $\operatorname{rank}(T_N) \leq N$, so T_N is finite-rank. In particular, T_N is compact. Moreover,

$$(T - T_N)x = \left(0, \dots, 0, \frac{x_{N+1}}{N+1}, \frac{x_{N+2}}{N+2}, \dots\right).$$

Hence, for every $x \in \ell^2$,

$$\|(T - T_N)x\|_{\ell^2}^2 = \sum_{n=N+1}^{\infty} \frac{|x_n|^2}{n^2} \leq \frac{1}{(N+1)^2} \sum_{n=N+1}^{\infty} |x_n|^2 \leq \frac{1}{(N+1)^2} \|x\|_{\ell^2}^2 . \tag{17.31}$$

Therefore

$$\|T - T_N\|_{\mathcal{L}(\ell^2,\ell^2)} \leq \frac{1}{N+1}.$$

In fact equality holds, because taking $x = e_{N+1}$ gives

$$\|(T - T_N)e_{N+1}\|_{\ell^2} = \frac{1}{N+1}.$$

Thus

$$\|T - T_N\|_{\mathcal{L}(\ell^2,\ell^2)} = \frac{1}{N+1} \to 0. \tag{17.32}$$

Consequently, T is the operator-norm limit of the compact finite-rank operators T_N. Since the operator-norm limit of compact linear operators is compact, it follows that T is compact.

Definition 17.12 (Compact and finite-dimensional classes)**.** Let X be a real Banach space and let $U \subset X$ be open and bounded. We denote by $\mathcal{K}(\overline{U}, X)$ the space of compact maps $F : \overline{U} \to X$. We also denote by $\mathcal{K}_{\text{fin}}(\overline{U}, X)$ the class of compact maps $F : \overline{U} \to X$ with finite-dimensional image, namely

$$\mathcal{K}_{\text{fin}}(\overline{U}, X) = \{F \in \mathcal{K}(\overline{U}, X) : \dim \operatorname{span} F(\overline{U}) < \infty\}.$$

Lemma 17.13 (Closedness of compact perturbations of the identity)**.** *Let X be a real Banach space, let $U \subset X$ be bounded and closed, and let $F \in \mathcal{K}(U, X)$. Then, $I + F$ maps closed subsets of U into closed subsets of X and it is proper, that is, inverse images of compact sets are compact.*

Proof. Let $A \subset U$ be closed, and let

$$y_n = (I + F)(x_n) \in (I + F)(A)$$

be such that $y_n \to y$ in X. Since $y_n - x_n = F(x_n)$ and $F(A)$ is precompact, after passing to a subsequence we may assume that $F(x_n) \to z$ for some $z \in X$. Hence

$$x_n = y_n - F(x_n) \to y - z.$$

Since A is closed, $x := y - z \in A$. By continuity of F,

$$y = \lim_{n\to\infty} (x_n + F(x_n)) = x + F(x).$$

Thus $y \in (I + F)(A)$. Hence $(I + F)(A)$ is closed.

Now let $K \subset X$ be compact and let $x_n \in (I + F)^{-1}(K)$. Then

$$y_n := x_n + F(x_n) \in K.$$

Passing to a subsequence, we may assume that $y_n \to y \in K$. Since $F(U)$ is precompact, passing to another subsequence we may assume that $F(x_n) \to z$. Therefore

$$x_n = y_n - F(x_n) \to y - z.$$

Hence every sequence in $(I + F)^{-1}(K)$ has a convergent subsequence, and the limit belongs to $(I + F)^{-1}(K)$ by continuity. Therefore $(I + F)^{-1}(K)$ is compact. □

Remark 17.14 (Why closedness of the boundary image is needed)**.** The closedness part of the previous lemma is used to convert the admissibility condition $y \notin (I + F)(\partial U)$ into a positive separation condition. Indeed, since ∂U is closed in $\overline{U}$ and F is compact, the map $I + F$ maps ∂U into a closed subset of X. Hence

$$(I + F)(\partial U) \quad \text{is closed in } X. \tag{17.33}$$

Therefore

$$\rho := \operatorname{dist}\big(y, (I + F)(\partial U)\big) > 0. \tag{17.34}$$

This positive distance is essential in the definition of the Leray–Schauder degree. It allows us to choose a finite-dimensional approximation F_1 of F such that

$$\|F_1 - F\|_{L^\infty(\overline{U},X)} < \rho, \tag{17.35}$$

and then guarantees that no boundary solution is created by the approximation: $y \notin (I + F_1)(\partial U)$. Without closedness of $(I + F)(\partial U)$, one could have

$$y \notin (I + F)(\partial U) \qquad \text{but} \qquad \operatorname{dist}\big(y, (I + F)(\partial U)\big) = 0,$$

and then arbitrarily small uniform perturbations could produce boundary solutions.

Remark 17.15 (Why compactness of the solution set is useful)**.** The properness part of the previous lemma implies that, for every $y \in X$, the solution set

$$\mathcal{S}_y := \{x \in \overline{U} : x + F(x) = y\} = (I + F)^{-1}(\{y\}) \tag{17.36}$$

is compact. Indeed, $\{y\}$ is compact in X, and $I + F$ is proper on $\overline{U}$.

This compactness is not strictly necessary for the formal definition of the degree, where the main point is the closedness of $(I + F)(\partial U)$. It becomes relevant when one wants to use the degree as a local invariant. For example, if $V \subset U$ is an open set containing all solutions, then compactness allows one to localize the degree on V. Similarly, if all solutions are isolated, compactness implies that only finitely many solutions can occur in $\overline{U}$, so the degree can be decomposed as a finite sum of local degrees.

18 The Leray-Schauder degree 2 - 05/05/2026

18.1 Definition of the degree map

Definition 18.1 (Leray–Schauder degree)**.** Let X be a real Banach space, let $U \subset X$ be open and bounded, and let $y \in X$. Define

$$\mathcal{D}_y(U,X) = \{F \in \mathcal{K}(\overline{U},X) : y \notin (I+F)(\partial U)\}.$$

Similarly, define

$$\mathcal{K}_{\mathrm{fin},y}(U,X) = \mathcal{K}_{\mathrm{fin}}(\overline{U},X) \cap \mathcal{D}_y(U,X).$$

Let $F \in \mathcal{D}_y(U,X)$. By the previous lemma,$(I+F)(\partial U)$ is closed. Hence

$$\rho := \operatorname{dist}\big(y,(I+F)(\partial U)\big) > 0. \tag{18.1}$$

Choose $F_1 \in \mathcal{K}_{\mathrm{fin}}(\overline{U},X)$ such that

$$\|F_1 - F\|_{L^\infty(\overline{U},X)} < \rho. \tag{18.2}$$

Then $F_1 \in \mathcal{K}_{\mathrm{fin},y}(U,X)$. Indeed, if $x \in \partial U$, then

$$\|x + F_1(x) - y\|_X \geq \|x + F(x) - y\|_X - \|F_1(x) - F(x)\|_X > \rho - \rho = 0.$$

Let $X_1 \subset X$ be a finite-dimensional subspace such that $F_1(\overline{U}) \subset X_1$, $y \in X_1$ and set $U_1 = U \cap X_1$. The Leray–Schauder degree is defined by

$$\deg(I+F,U,y) := \deg_{X_1}(I+F_1,U_1,y), \tag{18.3}$$

where the degree on the right-hand side is the Brouwer degree in the finite-dimensional vector space X_1.

Theorem 18.2 (Reduction property)**.** *Let $X_0 \subset X_1$ be finite-dimensional real vector spaces, let $U \subset X_1$ be open and bounded, and let $F : \overline{U} \to X_0$ be continuous. Let $y \in X_0$, and assume that $y \notin (I+F)(\partial U)$. Set $U_0 = U \cap X_0$. Then*

$$\deg_{X_1}(I+F,U,y) = \deg_{X_0}(I+F|_{\overline{U_0}},U_0,y). \tag{18.4}$$

Proof. We first prove the statement in the smooth regular case. After choosing linear coordinates, we may assume that

$$X_1 = \mathbb{R}^n = \mathbb{R}^m \oplus \mathbb{R}^{n-m}, \qquad X_0 = \mathbb{R}^m \oplus \{0\}.$$

Write

$$x = (x_1,x_2), \qquad x_1 \in \mathbb{R}^m, \quad x_2 \in \mathbb{R}^{n-m}.$$

Since F takes values in X_0, we can write

$$F(x_1,x_2) = (F_0(x_1,x_2),0).$$

Let $y = (y_1,0) \in X_0$. If $x + F(x) = y$, then

$$(x_1 + F_0(x_1,x_2),x_2) = (y_1,0).$$

Therefore necessarily $x_2 = 0$. Hence all solutions of $x + F(x) = y$ lie in $U_0 = U \cap X_0$. At such a solution $(x_1,0)$, the derivative of $I+F$ has block form

$$d(I+F)(x_1,0) = \begin{pmatrix} I_{\mathbb{R}^m} + d_{x_1}F_0(x_1,0) & d_{x_2}F_0(x_1,0) \\ 0 & I_{\mathbb{R}^{n-m}} \end{pmatrix}.$$

Consequently,

$$\det d(I+F)(x_1,0) = \det\big(I_{\mathbb{R}^m} + d_{x_1}F_0(x_1,0)\big). \tag{18.5}$$

By the determinant formula for the Brouwer degree, the contribution of each solution to the degree in X_1 is equal to the contribution of the same solution to the degree in X_0. Hence

$$\deg_{X_1}(I+F,U,y) = \deg_{X_0}(I+F|_{\overline{U_0}},U_0,y).$$

The general continuous case follows by approximation by smooth maps and homotopy invariance of the Brouwer degree. □

Proposition 18.3 (Well-posedness of the Leray–Schauder degree)**.** *The definition* (18.3) *is independent of the chosen finite-dimensional approximation F_1 and of the chosen finite-dimensional subspace X_1.*

Proof. Let $F \in \mathcal{D}_y(U, X)$ and set

$$\rho = \operatorname{dist}\big(y, (I+F)(\partial U)\big) > 0.$$

Let $F_1, F_2 \in \mathcal{K}_{\text{fin}}(\overline{U}, X)$ be such that

$$\|F_i - F\|_{L^\infty(\overline{U},X)} < \rho, \qquad i = 1, 2.$$

Let $X_i \subset X$ be finite-dimensional subspaces such that

$$F_i(\overline{U}) \subset X_i, \qquad y \in X_i.$$

Set

$$U_i = U \cap X_i, \qquad i = 1, 2, \qquad X_0 = X_1 + X_2, \qquad U_0 = U \cap X_0.$$

Then X_0 is finite-dimensional, and by the reduction property,

$$\deg_{X_i}(I + F_i, U_i, y) = \deg_{X_0}(I + F_i, U_0, y), \qquad i = 1, 2. \tag{18.6}$$

It remains to compare the two degrees in X_0. Define the finite-dimensional homotopy $H : [0,1] \times \overline{U_0} \to X_0$ by

$$H(t, x) = (1-t)F_1(x) + tF_2(x). \tag{18.7}$$

For $x \in \partial U$ and $t \in [0,1]$,

$$\|H(t,x) - F(x)\|_X \le (1-t)\,\|F_1(x) - F(x)\|_X + t\,\|F_2(x) - F(x)\|_X < \rho. \tag{18.8}$$

Therefore $y \notin (I + H(t,\cdot))(\partial U)$ for every $t \in [0,1]$. In particular,

$$y \notin (I + H(t,\cdot))(\partial U_0) \qquad \text{for every } t \in [0,1],$$

because $\partial_{X_0} U_0 \subset X_0 \cap \partial U$. By homotopy invariance of the Brouwer degree in the finite-dimensional space X_0,

$$\deg_{X_0}(I + F_1, U_0, y) = \deg_{X_0}(I + F_2, U_0, y). \tag{18.9}$$

Combining the previous identities gives

$$\deg_{X_1}(I + F_1, U_1, y) = \deg_{X_2}(I + F_2, U_2, y).$$

Thus the Leray–Schauder degree is well defined. □

Theorem 18.4 (Leray–Schauder degree)**.** *Let X be a real Banach space, let $U \subset X$ be open and bounded, and let $F \in \mathcal{D}_y(U, X)$. The Leray–Schauder degree has the following properties.*

1. *Translation.*
$$\deg(I + F, U, y) = \deg(I + F - y, U, 0). \tag{18.10}$$

2. *Normalization. If $y \notin \partial U$, then*
$$\deg(I, U, y) = \begin{cases} 1, & y \in U, \\ 0, & y \notin \overline{U}. \end{cases} \tag{18.11}$$

3. *Additivity. Let $U_1, U_2 \subset U$ be open and disjoint. Assume that $y \notin (I+F)\big(\overline{U} \setminus (U_1 \cup U_2)\big)$. Then*
$$\deg(I + F, U, y) = \deg(I + F, U_1, y) + \deg(I + F, U_2, y). \tag{18.12}$$

4. *Homotopy invariance. Let* $H : [0,1] \times \overline{U} \to X$ *be compact, and let* $y : [0,1] \to X$ *be continuous. Assume that*

$$y(t) \notin (I + H(t,\cdot))(\partial U) \qquad \text{for every } t \in [0,1]. \tag{18.13}$$

Then $\deg(I + H(t,\cdot), U, y(t))$ *is independent of* $t \in [0,1]$. *In particular,*

$$\deg(I + H(0,\cdot), U, y(0)) = \deg(I + H(1,\cdot), U, y(1)). \tag{18.14}$$

Proof. All four properties are inherited from the corresponding properties of the Brouwer degree through the finite-dimensional approximation defining the Leray–Schauder degree.

More precisely, once F is replaced by a sufficiently close finite-dimensional approximation F_1, and once a finite-dimensional subspace X_1 containing $F_1(\overline{U})$ and y is chosen, the Leray–Schauder degree is by definition

$$\deg(I + F, U, y) = \deg_{X_1}(I + F_1, U \cap X_1, y).$$

Translation, normalization, and additivity then follow from the corresponding Brouwer degree properties in X_1. The well-posedness proposition guarantees that the result does not depend on the particular approximation or finite-dimensional subspace chosen.

It remains to justify the homotopy invariance property. By the translation property, it is enough to treat the case where the target point is fixed and equal to 0. Indeed, define

$$K : [0,1] \times \overline{U} \to X, \qquad K(t,x) := H(t,x) - y(t).$$

Then K is compact, because $H([0,1] \times \overline{U})$ is precompact and $y([0,1])$ is compact. Moreover,

$$\deg(I + H(t,\cdot), U, y(t)) = \deg(I + K(t,\cdot), U, 0).$$

Thus we prove that $t \mapsto \deg(I + K(t,\cdot), U, 0)$ is constant.

Fix $t_0 \in [0,1]$. We first claim that the boundary condition is locally uniform in t near t_0. Namely, there exist $\delta > 0$ and $\rho > 0$ such that

$$\|x + K(t,x)\|_X \geq \rho \tag{18.15}$$

for every $x \in \partial U$ and every $t \in [0,1]$ with $|t - t_0| < \delta$. Suppose not. Then there exist sequences $t_n \to t_0, \qquad x_n \in \partial U$, such that

$$x_n + K(t_n, x_n) \to 0.$$

Since K is compact, after passing to a subsequence we may assume that

$$K(t_n, x_n) \to z$$

for some $z \in X$. Hence

$$x_n = \big(x_n + K(t_n, x_n)\big) - K(t_n, x_n) \to -z.$$

Since ∂U is closed, $x_0 := -z$ belongs to ∂U. By continuity of K, $K(t_n, x_n) \to K(t_0, x_0)$. Therefore $z = K(t_0, x_0)$, and so $x_0 + K(t_0, x_0) = 0$. This contradicts the assumption

$$0 \notin (I + K(t_0,\cdot))(\partial U).$$

Thus (18.15) holds. Choose an interval $J \subset [0,1]$ containing t_0 such that (18.15) holds for every $t \in J$. Since K is compact on $J \times \overline{U}$, we may choose a compact map with finite-dimensional image $K_1 : J \times \overline{U} \to X$ such that

$$\|K_1 - K\|_{L^\infty(J \times \overline{U}, X)} < \rho. \tag{18.16}$$

Let $X_1 \subset X$ be a finite-dimensional subspace such that

$$K_1(J \times \overline{U}) \subset X_1,$$

and set $U_1 := U \cap X_1$. For $x \in \partial U$ and $t \in J$, we have

$$\|x + K_1(t,x)\|_X \geq \|x + K(t,x)\|_X - \|K_1(t,x) - K(t,x)\|_X > \rho - \rho = 0.$$

Hence $0 \notin (I + K_1(t,\cdot))(\partial U)$ for every $t \in J$. In particular,

$$0 \notin (I + K_1(t,\cdot))(\partial_{X_1} U_1) \qquad \text{for every } t \in J,$$

because $\partial_{X_1} U_1 \subset X_1 \cap \partial U$. By the definition of the Leray–Schauder degree and by the well-posedness of the finite-dimensional approximation,

$$\deg(I + K(t,\cdot), U, 0) = \deg_{X_1}(I + K_1(t,\cdot), U_1, 0) \qquad \text{for every } t \in J.$$

The right-hand side is a Brouwer degree in the fixed finite-dimensional space X_1. Since $(t,x) \mapsto x + K_1(t,x)$ has no zero on $J \times \partial_{X_1} U_1$, the homotopy invariance of the Brouwer degree implies that $\deg_{X_1}(I + K_1(t,\cdot), U_1, 0)$ is independent of $t \in J$. Therefore $\deg(I + K(t,\cdot), U, 0)$ is locally constant in t. Since $[0,1]$ is connected, a locally constant integer-valued function on $[0,1]$ is constant. Hence

$$\deg(I + K(t,\cdot), U, 0)$$

is independent of $t \in [0,1]$. Returning to H and $y(t)$, we obtain $\deg(I + H(t,\cdot), U, y(t))$ independent of $t \in [0,1]$. In particular,

$$\deg(I + H(0,\cdot), U, y(0)) = \deg(I + H(1,\cdot), U, y(1)).$$

□

Proposition 18.5 (Other properties)**.** *Let X be a real Banach space, let $U \subset X$ be open and bounded, and let $F \in \mathcal{D}_y(U, X)$. Then the following properties hold.*

1. *Empty set.*
$$\deg(I + F, \varnothing, y) = 0. \tag{18.17}$$

2. *Iterated additivity. Let $U_1, \dots, U_N \subset U$ be pairwise disjoint open sets. If*
$$y \notin (I + F)\left(\overline{U} \setminus \bigcup_{i=1}^{N} U_i\right), \tag{18.18}$$
then
$$\deg(I + F, U, y) = \sum_{i=1}^{N} \deg(I + F, U_i, y). \tag{18.19}$$

3. *No solutions implies zero degree.*
$$y \notin (I + F)(U) \quad \Rightarrow \quad \deg(I + F, U, y) = 0. \tag{18.20}$$
Equivalently,
$$\deg(I + F, U, y) \neq 0, \quad \Rightarrow \quad \exists\, x \in U \ :\ x + F(x) = y. \tag{18.21}$$

4. *Local constancy in the map and in the point. If $G : \overline{U} \to X$ is compact and $z \in X$ satisfy*
$$\|F - G\|_{L^\infty(\partial U, X)} + \|y - z\|_X < \operatorname{dist}\big(y, (I + F)(\partial U)\big), \tag{18.22}$$
then
$$\deg(I + G, U, z) = \deg(I + F, U, y). \tag{18.23}$$
In particular, the degree is locally constant both with respect to the compact map and with respect to the point.

5. *Boundary invariance. If $G : \overline{U} \to X$ is compact and $G(x) = F(x)$ for every $x \in \partial U$, then*
$$\deg(I + G, U, y) = \deg(I + F, U, y). \tag{18.24}$$

Proof of point 4. Let $\rho = \operatorname{dist}\big(y, (I+F)(\partial U)\big) > 0$ and define the homotopy $H : [0,1] \times \overline{U} \to X$ and the function $y(t)$ by

$$H(t,x) = (1-t)F(x) + tG(x), \quad y_t = (1-t)y + tz. \tag{18.25}$$

For $x \in \partial U$ we have

$$\begin{aligned} \|x + H(t,x) - y_t - (x + F(x) - y)\|_X &= \|H(t,x) - F(x) - (y_t - y)\|_X && (18.26)\\ &\le t\,\|G(x) - F(x)\|_X + t\,\|z - y\|_X && (18.27)\\ &\le \|G - F\|_{L^\infty(\partial U, X)} + \|z - y\|_X < \rho. && (18.28) \end{aligned}$$

Since

$$\|x + F(x) - y\|_X \ge \rho \qquad \text{for every } x \in \partial U,$$

it follows that

$$x + H(t,x) \ne y_t \qquad \text{for every } x \in \partial U,\ t \in [0,1].$$

Therefore, by homotopy invariance,

$$\deg(I + F, U, y) = \deg(I + G, U, z).$$

□

Exercise 18.6. Prove the remaining points in the proposition above.

Exercise 18.7 (Linear Leray–Schauder degree formula)**.** Let $B = B_1(0)$ denote the open unit ball.

1. Let $X = \mathbb{R}^n$ and let $L \in \mathcal{L}(\mathbb{R}^n)$ be a linear map such that 1 is not an eigenvalue of L. Prove that
$$\deg(I - L, B, 0) = (-1)^m, \tag{18.29}$$
where m is the total algebraic multiplicity of the real eigenvalues $\lambda > 1$ of L.
2. Let $X = \ell^2(\mathbb{N};\mathbb{R})$, and let $L : X \to X$ be the diagonal compact operator
$$Lx = (\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n, \dots), \tag{18.30}$$
where $\lambda_n \in \mathbb{R}$, $\lambda_n \to 0$, and $\lambda_n \ne 1$ for every $n \in \mathbb{N}$. Prove that
$$\deg(I - L, B, 0) = (-1)^m, \tag{18.31}$$
where $m = \#\{n \in \mathbb{N} : \lambda_n > 1\}$.

Exercise 18.8. Let X be a real Banach space, let $B = B_1(0) \subset X$, and let $F_0 : \overline{B} \to X$ be compact. Set $F = I - F_0$ and assume that there exists a compact linear operator $L \in \mathcal{L}(X)$ such that

$$\|F_0(x) - Lx\| < \|x - Lx\| \qquad \text{for every } x \in \partial B. \tag{18.32}$$

Then $\deg(F, B, 0)$ is odd.

Exercise 18.9. There is no map

$$F = I - F_0 : \overline{B_1(0)} \to \partial B_1(0) \tag{18.33}$$

such that $F_0 : \overline{B_1(0)} \to X$ is compact and $F(x) = x$ for every $x \in \partial B_1(0)$.

Proof. Let $B = B_1(0)$ and suppose, by contradiction, that such a map exists. Since $F(\overline{B}) \subseteq \partial B$, we have $0 \notin F(\overline{B})$. Hence, by the existence property of the Leray–Schauder degree,

$$\deg(F, B, 0) = 0. \tag{18.34}$$

On the other hand, by assumption, $F(x) = x$ when $x \in \partial B$. Thus F and the identity map I agree on ∂B. By boundary invariance and normalization of the Leray–Schauder degree,

$$\deg(F, B, 0) = \deg(I, B, 0) = 1. \tag{18.35}$$

Equations (18.34) and (18.35) contradict each other. Hence no such map exists. □

19 The Leray-Schauder degree 3 - 11/05/2026

19.1 Application: Schauder fixed point

Theorem 19.1 (Leray–Schauder principle)**.** *Let X be a real Banach space and let $F \in \mathcal{K}(X, X)$. Assume that there exists $M > 0$ such that every solution of*

$$x = tF(x), \qquad t \in [0, 1], \tag{19.1}$$

satisfies the a priori bound $\|x\|_X \leq M$. Then F has a fixed point.

Proof. Choose $\rho > M$ and consider the compact homotopy

$$H : [0, 1] \times \overline{B_\rho(0)} \to X, \qquad H(t, x) = -tF(x). \tag{19.2}$$

We claim that

$$0 \notin (I + H(t, \cdot))(\partial B_\rho(0)) \qquad \text{for every } t \in [0, 1]. \tag{19.3}$$

Indeed, if $x \in \partial B_\rho(0)$ and $x + H(t, x) = 0$, then $x = tF(x)$. By the a priori bound, this implies

$$\rho = \|x\|_X \leq M, \tag{19.4}$$

This is impossible. Hence, by homotopy invariance and normalization of the Leray–Schauder degree,

$$\deg(I - F, B_\rho(0), 0) = \deg(I, B_\rho(0), 0) = 1. \tag{19.5}$$

By the existence property of the degree, there exists $x \in B_\rho(0)$ such that $x - F(x) = 0$. Thus $x = F(x)$, and F has a fixed point. □

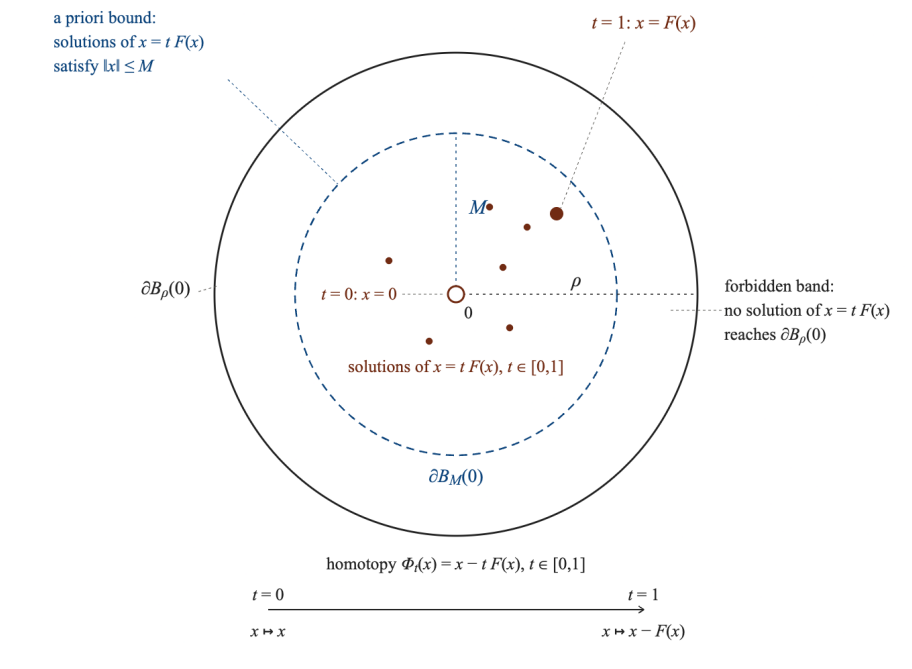

Figure 8: Geometric interpretation of the Leray-Schauder fixed point principle

Remark 19.2 (Geometric interpretation)**.** The a priori estimate is the geometric mechanism which makes the degree argument possible. Indeed, the zeros of the homotopy

$$x - tF(x) = 0, \qquad t \in [0, 1], \tag{19.6}$$

are exactly the solutions of $x = tF(x)$. The assumption says that all such zeros, for all values of the homotopy parameter t, lie in the fixed ball $B_M(0)$. Hence, after choosing $\rho > M$, no zero of the homotopy can touch the boundary $\partial B_\rho(0)$.

Thus the a priori estimate provides a boundary barrier: during the deformation from the identity map I to the map $I - F$, the zero set remains strictly inside $B_\rho(0)$. Consequently no zero can escape through the boundary, and the Leray–Schauder degree is preserved:

$$\deg(I - F, B_\rho(0), 0) = \deg(I, B_\rho(0), 0) = 1. \tag{19.7}$$

Since the final degree is nonzero, the final map $I - F$ must have a zero in $B_\rho(0)$, equivalently F has a fixed point.

Theorem 19.3 (Schauder fixed point)**.** *Let X be a real Banach space and let $C \subset X$ be closed, convex, and bounded. If $F \in \mathcal{K}(C, C)$, then there exists $x \in C$ such that $F(x) = x$.*

Proof. By the finite-dimensional approximation result for compact maps, there exist maps $F_n : C \to X$ such that

$$\|F_n - F\|_{L^\infty(C,X)} \to 0 \tag{19.8}$$

and $X_n := \operatorname{span} F_n(C)$ is finite-dimensional. Recalling the construction by partitions of unity of F_n, the approximating maps $F_n(x)$ are a convex combination of points of $F(C)$, so that

$$F_n(C) \subset \operatorname{conv}(F(C)), \tag{19.9}$$

where $\operatorname{conv}(F(C))$ is the convex hull of $F(C)$. Since $F(C) \subset C$ and C is convex, we have $\operatorname{conv}(F(C)) \subset C$. Hence $F_n(C) \subset X_n \cap C$, and in particular,

$$F_n(X_n \cap C) \subset X_n \cap C. \tag{19.10}$$

The set $X_n \cap C$ is closed, convex, bounded, and contained in the finite-dimensional space X_n. Therefore it is compact. By the finite-dimensional Brouwer fixed-point theorem applied in X_n, there exists $x_n \in X_n \cap C$ such that

$$x_n = F_n(x_n). \tag{19.11}$$

Since C is bounded and F is compact, the sequence $F(x_n)$ has a convergent subsequence. Passing to such a subsequence, still denoted by $F(x_n)$, there exists $z \in X$ such that

$$F(x_n) \to z. \tag{19.12}$$

Moreover, since $F(C) \subset C$ and C is closed, we have $z \in C$. Now

$$\begin{aligned} \|x_n - z\|_X &= \|F_n(x_n) - z\|_X && (19.13)\\ &\le \|F_n(x_n) - F(x_n)\|_X + \|F(x_n) - z\|_X && (19.14)\\ &\le \|F_n - F\|_{L^\infty(C,X)} + \|F(x_n) - z\|_X \to 0. && (19.15) \end{aligned}$$

Hence $x_n \to z$, and by continuity of F, $F(x_n) \to F(z)$. Since also $F(x_n) \to z$, we conclude that

$$F(z) = z. \tag{19.16}$$

Thus F has a fixed point in C. □

Theorem 19.4 (Another fixed-point theorem)**.** *Let X be a real Banach space, let $U \subset X$ be open and bounded, and let $F \in \mathcal{K}(\overline{U}, X)$. Assume that there exists $x_0 \in U$ such that, for every $x \in \partial U$ and every $\alpha \in (1, \infty)$.*

$$F(x) - x_0 \neq \alpha(x - x_0) \tag{19.17}$$

Then F has a fixed point in $\overline{U}$.

Proof. If F has a fixed point on ∂U, there is nothing to prove. Therefore assume that

$$F(x) \neq x \qquad \text{for every } x \in \partial U. \tag{19.18}$$

We prove that $\deg(I - F, U, 0) = 1$. Consider the compact homotopy

$$H : [0,1] \times \overline{U} \to X, \qquad H(t,x) = -x_0 - t(F(x) - x_0). \tag{19.19}$$

Then

$$I + H(t,\cdot) = x - x_0 - t(F(x) - x_0). \tag{19.20}$$

We claim that

$$0 \notin (I + H(t,\cdot))(\partial U) \qquad \text{for every } t \in [0,1]. \tag{19.21}$$

Indeed, suppose that, for some $x \in \partial U$ and some $t \in [0,1]$,

$$x - x_0 - t(F(x) - x_0) = 0. \tag{19.22}$$

If $t = 0$, then $x = x_0$, contradicting $x_0 \in U$ and $x \in \partial U$. If $0 < t < 1$, then

$$F(x) - x_0 = \frac{1}{t}(x - x_0), \tag{19.23}$$

with $1/t \in (1, \infty)$, contradicting (19.17). Finally, if $t = 1$, then $F(x) = x$, contradicting (19.18). Therefore the homotopy has no zero on the boundary. By homotopy invariance,

$$\deg(I - F, U, 0) = \deg(I - x_0, U, 0). \tag{19.24}$$

Since $x_0 \in U$, normalization gives

$$\deg(I - x_0, U, 0) = 1. \tag{19.25}$$

Hence $\deg(I - F, U, 0) = 1 \neq 0$, and by the existence property of the degree, there exists $x \in U$ such that $F(x) = x$. □

Corollary 19.5 (Conditions for fixed points)**.** *Let X be a real Banach space and let $F \in \mathcal{K}(\overline{B}_\rho(0), X)$, for some $\rho > 0$. Then F has a fixed point if one of the following conditions holds.*

(i) (Rothe condition)

$$F(\partial B_\rho(0)) \subseteq \overline{B}_\rho(0).$$

(ii) (Altman condition)

$$\|F(x) - x\|_X^2 \geq \|F(x)\|_X^2 - \|x\|_X^2 \qquad \textit{for every } x \in \partial B_\rho(0). \tag{19.26}$$

(iii) (Krasnosel'skii condition) *X is a Hilbert space and*

$$\langle F(x), x\rangle \leq \|x\|_X^2 \qquad \textit{for every } x \in \partial B_\rho(0). \tag{19.27}$$

Exercise 19.6. Prove the above corollary.

Exercise 19.7 (A semilinear Dirichlet problem by a priori bounds)**.** Let $f \in C^0([0,1] \times \mathbb{R}, \mathbb{R})$. Assume that there exists $R > 0$ such that

$$f(t,s) < 0 \qquad \text{for every } t \in [0,1] \text{ and every } s \geq R, \tag{19.28}$$

and

$$f(t,s) > 0 \qquad \text{for every } t \in [0,1] \text{ and every } s \leq -R. \tag{19.29}$$

Prove, using the Leray–Schauder principle, that the boundary value problem

$$\begin{cases} -u''(t) = f(t, u(t)), & t \in (0,1), \\ u(0) = u(1) = 0, \end{cases} \tag{19.30}$$

has at least one solution $u \in C^2([0,1])$.

19.2 Application: integral and differential equations

Lemma 19.8 (Compactness of a nonlinear integral operator)**.** *Let $I = [a, b] \subset \mathbb{R}$, let $\tau \in C^0(I, I)$, and let $f \in C^0(I \times I \times \mathbb{R}^n, \mathbb{R}^n)$. Define*

$$F : C^0(I, \mathbb{R}^n) \to C^0(I, \mathbb{R}^n), \qquad F(x)(t) = \int_a^{\tau(t)} f(t, s, x(s))\,\mathrm{d}s. \tag{19.31}$$

Then F is compact.

Proof. We first prove that F is continuous. Let $x_0 \in C^0(I, \mathbb{R}^n)$ and fix $\varepsilon > 0$. Choose $\rho > \|x_0\|_\infty + 1$ and set

$$Q = I \times I \times \overline{B_\rho(0)}.$$

Since Q is compact and f is continuous, f is uniformly continuous on Q. Hence there exists $\delta \in (0, 1)$ such that

$$|f(t, s, \xi) - f(t, s, \eta)| \le \frac{\varepsilon}{b - a} \tag{19.32}$$

whenever $(t, s, \xi), (t, s, \eta) \in Q$ and $|\xi - \eta| < \delta$. If $\|x - x_0\|_\infty < \delta$, then $x(s), x_0(s) \in \overline{B_\rho(0)}$ for every $s \in I$. Therefore, for every $t \in I$,

$$|F(x)(t) - F(x_0)(t)| \le \int_a^{\tau(t)} |f(t, s, x(s)) - f(t, s, x_0(s))|\,\mathrm{d}s \tag{19.33}$$

$$\le (b - a)\frac{\varepsilon}{b - a} = \varepsilon. \tag{19.34}$$

Thus

$$\|F(x) - F(x_0)\|_\infty \le \varepsilon, \tag{19.35}$$

and F is continuous.

We now prove compactness using Ascoli-Arzelà Theorem. Let $\mathcal{B} \subset C^0(I, \mathbb{R}^n)$ be bounded. Choose $\rho > 0$ such that $\|x\|_\infty \le \rho$ for every $x \in \mathcal{B}$. Set

$$M = \max_{(t,s,\xi) \in I \times I \times \overline{B_\rho(0)}} |f(t, s, \xi)|. \tag{19.36}$$

Then, for every $x \in \mathcal{B}$ and every $t \in I$,

$$|F(x)(t)| \le \int_a^{\tau(t)} |f(t, s, x(s))|\,\mathrm{d}s \le (b - a)M. \tag{19.37}$$

Thus $F(\mathcal{B})$ is uniformly bounded. It remains to prove equicontinuity. Fix $\varepsilon > 0$. Choose $\varepsilon_1, \varepsilon_2 > 0$ such that

$$(b - a)\varepsilon_1 < \frac{\varepsilon}{2}, \qquad M\varepsilon_2 < \frac{\varepsilon}{2}, \tag{19.38}$$

with the second condition omitted if $M = 0$. Since f is uniformly continuous on $I \times I \times \overline{B_\rho(0)}$ and τ is uniformly continuous on I, there exists $\delta > 0$ such that, whenever $|t - t_0| < \delta$,

$$|f(t, s, \xi) - f(t_0, s, \xi)| < \varepsilon_1 \qquad \text{for every } s \in I,\ \xi \in \overline{B_\rho(0)}, \tag{19.39}$$

and

$$|\tau(t) - \tau(t_0)| < \varepsilon_2. \tag{19.40}$$

Then, for $x \in \mathcal{B}$ and $|t - t_0| < \delta$,

$$|F(x)(t) - F(x)(t_0)| = \left| \int_a^{\tau(t)} f(t, s, x(s))\,\mathrm{d}s - \int_a^{\tau(t_0)} f(t_0, s, x(s))\,\mathrm{d}s \right| \tag{19.41}$$

$$\le \int_a^{\tau(t_0)} |f(t, s, x(s)) - f(t_0, s, x(s))|\,\mathrm{d}s \tag{19.42}$$

$$+ \left| \int_{\tau(t_0)}^{\tau(t)} |f(t, s, x(s))|\,\mathrm{d}s \right| \tag{19.43}$$

$$\le (b - a)\varepsilon_1 + M|\tau(t) - \tau(t_0)| \tag{19.44}$$

$$< \varepsilon. \tag{19.45}$$

Therefore $F(\mathcal{B})$ is equicontinuous. By the Ascoli-Arzelà theorem, $F(\mathcal{B})$ is precompact in $C^0(I,\mathbb{R}^n)$. Hence F is compact. □

Theorem 19.9 (Solvability of a nonlinear integral equation)**.** *Let F be the operator defined in (19.31). Let $\lambda \in \mathbb{R}$, $y \in C^0(I,\mathbb{R}^n)$, and $\rho > 0$. Define*

$$M(\rho) := (b-a) \max_{\substack{t,s\in I \\ \eta\in\overline{B_\rho(0)}}} |f(t,s,y(s)+\eta)|. \tag{19.46}$$

If

$$|\lambda| M(\rho) \le \rho, \tag{19.47}$$

then the integral equation

$$x - \lambda F(x) = y \tag{19.48}$$

has at least one solution $x \in C^0(I,\mathbb{R}^n)$.

Proof. Equation (19.48) is equivalent to the fixed point problem

$$x = T(x), \qquad T(x) := y + \lambda F(x). \tag{19.49}$$

Consider the closed convex bounded set

$$C = \overline{B_\rho(y)} = \{x \in C^0(I,\mathbb{R}^n) : \|x-y\|_\infty \le \rho\}.$$

We show that $T(C) \subset C$. Let $x \in C$. Then

$$x(s) = y(s) + \eta(s), \qquad |\eta(s)| \le \rho \quad \text{for every } s \in I.$$

Therefore, for every $t \in I$,

$$|T(x)(t) - y(t)| = |\lambda|\,|F(x)(t)| \le |\lambda| \int_a^{\tau(t)} |f(t,s,x(s))|\,\mathrm{d}s \le |\lambda| M(\rho) \le \rho. \tag{19.50}$$

Thus $\|T(x)-y\|_\infty \le \rho$, and hence $T(C) \subset C$.

Since F is compact, $T = y + \lambda F$ is compact. By the Schauder fixed-point theorem, T has a fixed point in C. Hence there exists $x \in C^0(I,\mathbb{R}^n)$ such that

$$x = y + \lambda F(x),$$

which is precisely (19.48). □

Theorem 19.10 (Peano local existence theorem)**.** *Let $I \subset \mathbb{R}$ be an interval, let $t_0 \in \mathring{I}$, let $x_0 \in \mathbb{R}^n$, and let $f \in C^0(I \times \mathbb{R}^n, \mathbb{R}^n)$. Then the initial value problem*

$$\begin{cases} \dot{x}(t) = f(t,x(t)), \\ x(t_0) = x_0, \end{cases} \tag{19.51}$$

has at least one local solution. More precisely, there exists $\varepsilon > 0$ such that $[t_0-\varepsilon, t_0+\varepsilon] \subset I$ and there exists $x \in C^1([t_0-\varepsilon, t_0+\varepsilon], \mathbb{R}^n)$ satisfying (19.51) on this interval.

Proof. Choose $\rho > 0$. Since $t_0 \in \mathring{I}$, there exists $\varepsilon_0 > 0$ such that $J_0 := [t_0-\varepsilon_0, t_0+\varepsilon_0] \subset I$. Set

$$M_0 := \max_{(t,\xi)\in J_0 \times \overline{B_\rho(x_0)}} |f(t,\xi)|.$$

This maximum is finite because f is continuous and $J_0 \times \overline{B_\rho(x_0)}$ is compact. Choose $\varepsilon \in (0, \varepsilon_0]$ such that

$$\varepsilon M_0 \le \rho. \tag{19.52}$$

Set $J := [t_0 - \varepsilon, t_0 + \varepsilon]$ and define $\Phi : C^0(J, \mathbb{R}^n) \to C^0(J, \mathbb{R}^n)$ by

$$\Phi(x)(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, \mathrm{d}s. \tag{19.53}$$

Note that Φ is a compact operator. Now define

$$C_\rho := \{x \in C^0(J, \mathbb{R}^n) : \|x - x_0\|_\infty \le \rho\}.$$

Then C_ρ is closed, convex, bounded, and nonempty. We claim that $\Phi(C_\rho) \subset C_\rho$. Indeed, if $x \in C_\rho$, then

$$x(s) \in \overline{B_\rho(x_0)} \qquad \text{for every } s \in J.$$

Therefore, for every $t \in J$,

$$|\Phi(x)(t) - x_0| = \left| \int_{t_0}^{t} f(s, x(s))\, \mathrm{d}s \right| \le \varepsilon M_0 \le \rho, \tag{19.54}$$

where we used (19.52). Thus $\|\Phi(x) - x_0\|_\infty \le \rho$, and hence $\Phi(x) \in C_\rho$. By the Schauder fixed-point theorem, there exists $x \in C_\rho$ such that $x = \Phi(x)$. Therefore

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, \mathrm{d}s \qquad \text{for every } t \in J. \tag{19.55}$$

Since $s \mapsto f(s, x(s))$ is continuous, the fundamental theorem of calculus implies that $x \in C^1(J, \mathbb{R}^n)$ and

$$\dot{x}(t) = f(t, x(t)) \qquad \text{for every } t \in J.$$

Evaluating (19.55) at $t = t_0$ gives $x(t_0) = x_0$. Thus x is a local solution of (19.51). □

Exercise 19.11 (Global existence under linear growth)**.** Assume, in addition to the hypotheses of the Peano local existence theorem, that $I = \mathbb{R}$ and that there exists $C > 0$ such that

$$|f(t, x)| \le C(1 + |x|) \qquad \text{for every } (t, x) \in \mathbb{R} \times \mathbb{R}^n. \tag{19.56}$$

Prove that the initial value problem (19.51) admits a global solution $x \in C^1(\mathbb{R}, \mathbb{R}^n)$.

20 Stationary Navier-Stokes 1 - 18/05/2026

20.1 Introduction to Navier–Stokes

Let $U \subset \mathbb{R}^3$ be a bounded domain with C^1 boundary. We consider the motion of an incompressible fluid inside U. The unknowns of the model are the velocity field $v = v(t,x) \in \mathbb{R}^3$, the pressure $p = p(t,x) \in \mathbb{R}$, and a prescribed external force density $K = K(t,x) \in \mathbb{R}^3$.

The first structural condition is incompressibility. This means that the fluid has constant density and, therefore, the velocity field has no local sources or sinks. Mathematically this is expressed by

$$\operatorname{div} v = 0. \tag{20.1}$$

Equivalently, the velocity field is divergence-free at every time.

The incompressible Navier–Stokes equations describe the balance between inertia, viscosity, pressure, and external forces. In vector notation they are

$$\rho \partial_t v = \nu \Delta v - \rho (v \cdot \nabla) v - \nabla p + K, \tag{20.2}$$

together with the incompressibility constraint $\operatorname{div} v = 0$. Here $\rho > 0$ is the density of the fluid and $\nu > 0$ is the viscosity. Equivalently, one may write the equation in the form

$$\rho\big(\partial_t v + (v \cdot \nabla) v\big) = \nu \Delta v - \nabla p + K. \tag{20.3}$$

This form emphasizes that the left-hand side is the material acceleration of the fluid, multiplied by the density.

After normalizing the density by taking $\rho = 1$, the system becomes

$$\partial_t v + (v \cdot \nabla) v = \nu \Delta v - \nabla p + K \tag{20.4}$$

with $\operatorname{div} v = 0$.

Let us briefly interpret the different terms. The term $\partial_t v$ measures the variation of the velocity at a fixed point in space. The nonlinear term $(v \cdot \nabla) v$ describes transport of momentum by the fluid itself: the velocity field moves and deforms the velocity field. The viscous term $\nu \Delta v$ acts as a diffusion term and tends to smooth out variations of the velocity. The pressure gradient ∇p is not an independent forcing term; rather, the pressure adjusts itself so as to enforce the incompressibility constraint $\operatorname{div} v = 0$. Finally, K represents the prescribed external force acting on the fluid.

20.2 Stationary solutions and weak formulation

We restrict to stationary solutions, namely solutions independent of time. Thus $\partial_t v = 0$, and the equation becomes, componentwise,

$$0 = \nu \partial_i \partial_i v_j - (v_i \partial_i) v_j - \partial_j p + K_j, \qquad j = 1,2,3, \tag{20.5}$$

where summation occurs over repeated indexes (Einstein summation convention). Together with incompressibility and homogeneous Dirichlet boundary conditions, the stationary problem is

$$\begin{cases} -\nu \Delta v + (v \cdot \nabla) v + \nabla p = K, & x \in U, \\ \operatorname{div} v = 0, & x \in U, \\ v = 0, & x \in \partial U. \end{cases} \tag{20.6}$$

Let w be a smooth divergence-free test vector field with compact support or zero trace on the boundary. Testing (20.5) against w and integrating by parts, the pressure term disappears:

$$\int_U \partial_j p \, w_j \, \mathrm{d}x = -\int_U p \, \partial_j w_j \, \mathrm{d}x = 0. \tag{20.7}$$

Here the boundary term vanishes because w has zero trace, and the last equality follows from $\operatorname{div} w = 0$. The viscous term gives

$$-\int_U \Delta v \cdot w \, \mathrm{d}x = \int_U \partial_j v_i \, \partial_j w_i \, \mathrm{d}x. \tag{20.8}$$

The nonlinear term is rewritten using $\operatorname{div} v = 0$:

$$\int_U (v_i \partial_i) v_j \, w_j \, \mathrm{d}x = \int_U v_i \partial_i v_j \, w_j \, \mathrm{d}x \tag{20.9}$$

$$= \int_U v_i \partial_i (v_j w_j) \, \mathrm{d}x - \int_U v_i v_j \, \partial_i w_j \, \mathrm{d}x \tag{20.10}$$

$$= - \int_U v_i v_j \, \partial_i w_j \, \mathrm{d}x. \tag{20.11}$$

Indeed,

$$\int_U v_i \partial_i (v_j w_j) \, \mathrm{d}x = - \int_U (\partial_i v_i) v_j w_j \, \mathrm{d}x = 0. \tag{20.12}$$

Therefore the weak formulation is expected to be

$$\nu \int_U \partial_j v_i \, \partial_j w_i \, \mathrm{d}x - \int_U v_i v_j \, \partial_i w_j \, \mathrm{d}x = \int_U K \cdot w \, \mathrm{d}x \tag{20.13}$$

for every admissible divergence-free test field w.

The pressure is not present in (20.13). This is the main advantage of testing only against divergence-free fields. The pressure can be recovered later under additional hypotheses, but it is not needed for the Leray-Schauder existence argument.

To make the structure of the argument visible already at this stage, we introduce the following notation formally. The rigorous functional setting will be fixed below. For sufficiently regular divergence-free vector fields, define the trilinear form

$$a(u, v, w) := \int_U u_k v_j \partial_k w_j \, \mathrm{d}x. \tag{20.14}$$

Thus the nonlinear term in (20.13) is written as

$$\int_U v_i v_j \, \partial_i w_j \, \mathrm{d}x = a(v, v, w). \tag{20.15}$$

We shall choose a Hilbert space X of divergence-free vector fields with zero boundary values, endowed with the scalar product

$$\langle u, w \rangle_X := \int_U \partial_j u_i \, \partial_j w_i \, \mathrm{d}x. \tag{20.16}$$

Once this space is defined, the trilinear form will determine a bilinear map $B : X \times X \to X$ by the Riesz representation identity

$$\langle B(u, v), w \rangle_X = a(u, v, w) \qquad \text{for every } w \in X. \tag{20.17}$$

Similarly, the force K will determine an element $\widetilde{K} \in X$ by

$$\langle \widetilde{K}, w \rangle_X = \int_U K \cdot w \, \mathrm{d}x \qquad \text{for every } w \in X. \tag{20.18}$$

With this notation, the weak stationary Navier-Stokes equation will become

$$\nu v - B(v, v) = \widetilde{K} \qquad \text{in } X. \tag{20.19}$$

20.3 Plan of the argument

The goal is to solve (20.13) by the Leray-Schauder principle. The strategy has three steps.

1. Choose a Hilbert space X of divergence-free vector fields with zero boundary values.
2. Prove that the nonlinear map $v \mapsto B(v, v)$ is compact from X to X.

3. Apply Leray-Schauder to the fixed-point equation

$$v = t\left(\nu^{-1}B(v,v) + \nu^{-1}\widetilde{K}\right), \qquad t \in [0,1], \tag{20.20}$$

and prove an a priori bound for all its solutions.

The key energy cancellation, and the reason why the a priori estimate will close, for the trilinear form (20.14) is

$$a(v,v,v) = 0. \tag{20.21}$$

20.4 Sobolev spaces toolbox

We define

$$H^1(U) := \{u \in L^2(U) : \partial_j u \in L^2(U),\ j = 1,2,3\},$$

where the derivatives are understood in the weak sense. Moreover,

$$H^1_0(U) := \overline{C_c^\infty(U)}^{H^1(U)}.$$

For vector fields we set

$$H^1_0(U;\mathbb{R}^3) := (H^1_0(U))^3.$$

Lemma 20.1 (Weak integration by parts)**.** *Let $u \in H^1_0(U)$ and $v \in H^1(U)$. Then*

$$\int_U u\,\partial_j v\,\mathrm{d}x = -\int_U \partial_j u\, v\,\mathrm{d}x. \tag{20.22}$$

Proof. First assume that $u \in C_c^\infty(U)$ and $v \in C^\infty(\overline{U})$. Since u has compact support in U, the classical integration by parts formula gives

$$\int_U u\,\partial_j v\,\mathrm{d}x = -\int_U \partial_j u\, v\,\mathrm{d}x. \tag{20.23}$$

Now let $u \in H^1_0(U)$ and $v \in H^1(U)$. By definition of $H^1_0(U)$, there exists $u_n \in C_c^\infty(U)$ such that

$$u_n \to u \qquad \text{in } H^1(U). \tag{20.24}$$

Since U has C^1 boundary, smooth functions are dense in $H^1(U)$. Hence there exists $v_n \in C^\infty(\overline{U})$ such that

$$v_n \to v \qquad \text{in } H^1(U). \tag{20.25}$$

For each n, the identity

$$\int_U u_n\,\partial_j v_n\,\mathrm{d}x = -\int_U \partial_j u_n\, v_n\,\mathrm{d}x \tag{20.26}$$

holds. Passing to the limit follows from the Cauchy-Schwarz inequality and the convergences

$$u_n \to u, \quad \partial_j u_n \to \partial_j u, \quad v_n \to v, \quad \partial_j v_n \to \partial_j v \qquad \text{in } L^2(U).$$

Thus (20.22) holds. □

Theorem 20.2 (Poincaré-Friedrichs inequality)**.** *There exists $C_P > 0$ such that*

$$\|u\|_{L^2(U)} \le C_P\,\|\nabla u\|_{L^2(U)} \qquad \textit{for every } u \in H^1_0(U). \tag{20.27}$$

Proof. Choose a rectangular box $Q \subset \mathbb{R}^3$ such that $U \subset Q$. Since $u \in H^1_0(U)$ is the H^1-limit of functions in $C_c^\infty(U)$, it is enough to prove the estimate for $u \in C_c^\infty(U)$.

Extend u by zero outside U. The extension, still denoted by u, belongs to $H^1_0(Q)$. Let $Q = (a_1,b_1)\times(a_2,b_2)\times(a_3,b_3)$ and set $d_1 = b_1 - a_1$. For fixed x_2, x_3, the one-dimensional fundamental theorem of calculus gives

$$u(x_1,x_2,x_3) = \int_{a_1}^{x_1} \partial_1 u(s,x_2,x_3)\,\mathrm{d}s. \tag{20.28}$$

Therefore, by Cauchy's inequality,

$$|u(x_1,x_2,x_3)|^2 \le d_1 \int_{a_1}^{b_1} |\partial_1 u(s,x_2,x_3)|^2 \,\mathrm{d}s. \tag{20.29}$$

Integrating over Q gives

$$\|u\|_{L^2(Q)}^2 \le d_1^2 \,\|\partial_1 u\|_{L^2(Q)}^2 \le d_1^2 \,\|\nabla u\|_{L^2(Q)}^2 . \tag{20.30}$$

Since the extension is zero outside U, this gives

$$\|u\|_{L^2(U)} \le d_1 \,\|\nabla u\|_{L^2(U)} . \tag{20.31}$$

Thus (20.27) holds with $C_P = d_1$. The general case follows by density. □

As a consequence, on $H_0^1(U)$ we may use the equivalent norm

$$\|u\|_{H_0^1(U)} := \left(\int_U |\nabla u|^2 \,\mathrm{d}x \right)^{1/2}. \tag{20.32}$$

Indeed, Poincaré-Friedrichs gives

$$\|u\|_{H^1(U)} \le (1 + C_P^2)^{1/2} \,\|u\|_{H_0^1(U)} ,$$

while the opposite inequality is immediate from the definition of the H^1-norm.

Theorem 20.3 (Rellich compactness)**.** *The embedding $H_0^1(U) \hookrightarrow L^2(U)$ is compact.*

Proof. We use the standard Rellich compactness theorem. Let (u_n) be bounded in $H_0^1(U)$. Since U is bounded, the theorem asserts that there exists a subsequence (u_{n_k}) and a function $u \in L^2(U)$ such that

$$u_{n_k} \to u \qquad \text{in } L^2(U). \tag{20.33}$$

Equivalently, every bounded sequence in $H_0^1(U)$ has a strongly convergent subsequence in $L^2(U)$. This is precisely compactness of the embedding. □

Theorem 20.4 (Ladyzhenskaya inequality)**.** *There exists a constant $C_L > 0$ such that*

$$\|u\|_{L^4(U)} \le C_L \,\|u\|_{L^2(U)}^{1/4} \,\|u\|_{H_0^1(U)}^{3/4} \qquad \textit{for every } u \in H_0^1(U). \tag{20.34}$$

Proof. By the Sobolev inequality in dimension three,

$$\|u\|_{L^6(U)} \le C_S \,\|u\|_{H_0^1(U)} \qquad \text{for every } u \in H_0^1(U). \tag{20.35}$$

Interpolating between L^2 and L^6, we get

$$\|u\|_{L^4(U)} \le \|u\|_{L^2(U)}^{1/4} \,\|u\|_{L^6(U)}^{3/4} . \tag{20.36}$$

Indeed,

$$\frac{1}{4} = \frac{1}{4}\cdot\frac{1}{2} + \frac{3}{4}\cdot\frac{1}{6}.$$

Therefore

$$\|u\|_{L^4(U)} \le C_S^{3/4} \,\|u\|_{L^2(U)}^{1/4} \,\|u\|_{H_0^1(U)}^{3/4} . \tag{20.37}$$

Thus (20.34) holds with $C_L = C_S^{3/4}$. □

Combining (20.34) with Poincaré-Friedrichs, we obtain

$$\|u\|_{L^4(U)} \le \beta \,\|u\|_{H_0^1(U)} \qquad \text{for every } u \in H_0^1(U), \tag{20.38}$$

for a constant $\beta > 0$ depending only on U. Indeed,

$$\|u\|_{L^4(U)} \le C_L \,\|u\|_{L^2(U)}^{1/4} \,\|u\|_{H_0^1(U)}^{3/4} \le C_L C_P^{1/4} \,\|u\|_{H_0^1(U)} . \tag{20.39}$$

Corollary 20.5 (Compact embedding into L^4)**.** *The embedding $H_0^1(U) \hookrightarrow L^4(U)$ is compact.*

Proof. Let (u_n) be bounded in $H_0^1(U)$. By Rellich compactness, after passing to a subsequence, there exists $u \in L^2(U)$ such that

$$u_n \to u \qquad \text{in } L^2(U). \tag{20.40}$$

The sequence (u_n) is bounded in $H_0^1(U)$, hence bounded in $L^6(U)$ by the Sobolev inequality. Since $u_n \to u$ in $L^2(U)$ and (u_n) is bounded in $L^6(U)$, also $u \in L^6(U)$, and $(u_n - u)$ is bounded in $L^6(U)$.

By interpolation,

$$\|u_n - u\|_{L^4(U)} \le \|u_n - u\|_{L^2(U)}^{1/4} \, \|u_n - u\|_{L^6(U)}^{3/4}. \tag{20.41}$$

The first factor tends to zero and the second factor is bounded. Therefore

$$u_n \to u \qquad \text{in } L^4(U). \tag{20.42}$$

Thus every bounded sequence in $H_0^1(U)$ has a strongly convergent subsequence in $L^4(U)$, which proves compactness of the embedding. □

21 Stationary Navier-Stokes 2 - 19/05/2026

21.1 Functional setting and the nonlinear operator

Let

$$\mathcal{V} := \{\varphi \in C_c^\infty(U;\mathbb{R}^3) : \operatorname{div}\varphi = 0\}.$$

We define

$$X := \overline{\mathcal{V}}^{H_0^1(U;\mathbb{R}^3)}.$$

By definition, X is a closed subspace of $H_0^1(U;\mathbb{R}^3)$. Moreover, since U has C^1 boundary, one has the standard identification

$$X = \{v \in H_0^1(U;\mathbb{R}^3) : \operatorname{div} v = 0 \text{ in } \mathcal{D}'(U)\}.$$

Hence X is a Hilbert space with scalar product

$$\langle u, v\rangle_X = \int_U \partial_j u_i \, \partial_j v_i \, \mathrm{d}x. \tag{21.1}$$

For $v \in X$ we use the norm

$$\|v\|_X := \|\nabla v\|_{L^2(U)}.$$

By the Poincaré inequality this norm is equivalent to the full H_0^1 norm. We define the trilinear form

$$a(u,v,w) = \int_U u_k v_j \partial_k w_j \, \mathrm{d}x, \qquad u,v,w \in X. \tag{21.2}$$

Lemma 21.1 (Boundedness of the trilinear form)**.** *There exists a constant $\alpha > 0$ such that*

$$|a(u,v,w)| \le \alpha \, \|u\|_{L^4(U)} \, \|v\|_{L^4(U)} \, \|w\|_X \qquad \textit{for every } u,v,w \in X. \tag{21.3}$$

In particular, by (20.38)*,*

$$|a(u,v,w)| \le \alpha\beta^2 \|u\|_X \, \|v\|_X \, \|w\|_X. \tag{21.4}$$

Proof. By Hölder's inequality,

$$|a(u,v,w)| = \left| \int_U u_k v_j \partial_k w_j \, \mathrm{d}x \right| \tag{21.5}$$

$$\le \sum_{j,k=1}^{3} \|u_k\|_{L^4(U)} \, \|v_j\|_{L^4(U)} \, \|\partial_k w_j\|_{L^2(U)} \, . \tag{21.6}$$

The last expression is bounded by a dimensional constant times

$$\|u\|_{L^4(U)} \, \|v\|_{L^4(U)} \, \|w\|_X.$$

Absorbing the dimensional constant into α gives (21.3). Combining this with (20.38) gives (21.4). □

Lemma 21.2 (Cancellation)**.** *For every $v \in X$,*

$$a(v,v,v) = 0. \tag{21.7}$$

Proof. First assume that v is smooth, divergence-free, and has zero trace on ∂U. Then

$$a(v,v,v) = \int_U v_k v_j \partial_k v_j \, \mathrm{d}x = \frac{1}{2} \int_U v_k \partial_k (v_j v_j) \, \mathrm{d}x = -\frac{1}{2} \int_U (\partial_k v_k) v_j v_j \, \mathrm{d}x = 0. \tag{21.8}$$

The boundary term vanishes because $v = 0$ on ∂U, and the last identity follows from $\partial_k v_k = 0$. For a general $v \in X$, the identity follows by approximation by smooth divergence-free vector fields and by the continuity estimate (21.4). □

Let $K \in L^2(U, \mathbb{R}^3)$. The map $w \mapsto \int_U K \cdot w \, \mathrm{d}x$ is a bounded linear functional on X. Indeed,

$$\left| \int_U K \cdot w \, \mathrm{d}x \right| \leq \|K\|_{L^2(U)} \, \|w\|_{L^2(U)} \leq C_P \, \|K\|_{L^2(U)} \, \|w\|_X. \tag{21.9}$$

Therefore, by the Riesz representation theorem, there exists a unique $\widetilde{K} \in X$ such that

$$\int_U K \cdot w \, \mathrm{d}x = \langle \widetilde{K}, w \rangle_X \qquad \text{for every } w \in X. \tag{21.10}$$

Moreover,

$$\|\widetilde{K}\|_X \leq C_P \, \|K\|_{L^2(U)} \, . \tag{21.11}$$

Similarly, for every fixed $u, v \in X$, the map $w \mapsto a(u, v, w)$ is a bounded linear functional on X. Hence there exists a unique $B(u, v) \in X$ such that

$$a(u, v, w) = \langle B(u, v), w \rangle_X \qquad \text{for every } w \in X. \tag{21.12}$$

The estimate (21.4) gives

$$\|B(u, v)\|_X \leq \alpha \beta^2 \|u\|_X \, \|v\|_X. \tag{21.13}$$

Thus the weak stationary Navier-Stokes equation is equivalent to

$$\nu v - B(v, v) = \widetilde{K} \qquad \text{in } X. \tag{21.14}$$

Indeed, testing (21.14) against $w \in X$ gives

$$\nu \langle v, w \rangle_X - a(v, v, w) = \langle \widetilde{K}, w \rangle_X = \int_U K \cdot w \, \mathrm{d}x, \tag{21.15}$$

which is precisely the weak formulation.

Proposition 21.3 (Compactness of the nonlinear map)**.** *The map*

$$F : X \to X, \qquad F(v) = B(v, v), \tag{21.16}$$

is compact.

Proof. First, F is continuous. Indeed, if $u, v \in X$, then by bilinearity,

$$F(u) - F(v) = B(u, u) - B(v, v) = B(u - v, u) + B(v, u - v). \tag{21.17}$$

Hence, using (21.13),

$$\|F(u) - F(v)\|_X \leq \alpha \beta^2 \|u - v\|_X \, \|u\|_X + \alpha \beta^2 \|v\|_X \, \|u - v\|_X. \tag{21.18}$$

Thus on bounded sets,

$$\|F(u) - F(v)\|_X \leq C_R \|u - v\|_X, \qquad \|u\|_X, \|v\|_X \leq R, \tag{21.19}$$

so F is locally Lipschitz. We now prove compactness. Let (v_n) be bounded in X. By the compact embedding

$$X \hookrightarrow L^4(U, \mathbb{R}^3),$$

after passing to a subsequence, there exists $v \in L^4(U, \mathbb{R}^3)$ such that

$$v_n \to v \qquad \text{in } L^4(U, \mathbb{R}^3). \tag{21.20}$$

In particular, (v_n) is Cauchy in $L^4(U, \mathbb{R}^3)$ and bounded in X. For $n, m \in \mathbb{N}$, using the definition of B and the Riesz representation,

$$\|B(v_n, v_n) - B(v_m, v_m)\|_X = \sup_{\|w\|_X \leq 1} |\langle B(v_n, v_n) - B(v_m, v_m), w \rangle_X| \tag{21.21}$$

$$= \sup_{\|w\|_X \leq 1} |a(v_n, v_n, w) - a(v_m, v_m, w)|. \tag{21.22}$$

By bilinearity in the first two variables,

$$a(v_n, v_n, w) - a(v_m, v_m, w) = a(v_n - v_m, v_n, w) + a(v_m, v_n - v_m, w). \tag{21.23}$$

Thus, by (21.3),

$$\|B(v_n, v_n) - B(v_m, v_m)\|_X \le \alpha \|v_n - v_m\|_{L^4} \|v_n\|_{L^4} + \alpha \|v_m\|_{L^4} \|v_n - v_m\|_{L^4} \tag{21.24}$$

$$\le \alpha\big(\|v_n\|_{L^4} + \|v_m\|_{L^4}\big) \|v_n - v_m\|_{L^4}. \tag{21.25}$$

Since (v_n) is bounded in L^4 and Cauchy in L^4, it follows that $B(v_n, v_n)$ is Cauchy in X. Since X is complete, $B(v_n, v_n)$ converges in X. Hence F maps bounded sequences into sequences with convergent subsequences. Therefore $F : X \to X$ is compact. □

21.2 Existence by the Leray-Schauder principle

Theorem 21.4 (Existence of weak stationary Navier-Stokes solutions)**.** *Let $U \subset \mathbb{R}^3$ be a bounded domain with C^1 boundary, let $\nu > 0$, and let $K \in L^2(U, \mathbb{R}^3)$. Then there exists $v \in X$ such that*

$$\nu\langle v, w\rangle_X - a(v, v, w) = \int_U K \cdot w \, \mathrm{d}x \qquad \text{for every } w \in X. \tag{21.26}$$

Equivalently, there exists a weak velocity solution of the divergence-free formulation of the stationary Navier-Stokes problem (20.6). Under the standard pressure-recovery theorem, one can then recover a pressure p and obtain a weak solution of (20.6).

Proof. Equation (21.14) is equivalent to the fixed-point problem

$$v = \nu^{-1} B(v, v) + \nu^{-1} \widetilde{K}. \tag{21.27}$$

Define

$$\Phi : X \to X, \qquad \Phi(v) = \nu^{-1} B(v, v) + \nu^{-1} \widetilde{K}. \tag{21.28}$$

The map $v \mapsto B(v, v)$ is compact by the previous proposition, and the constant map $v \mapsto \widetilde{K}$ is compact. Hence Φ is compact. We apply the Leray-Schauder principle. Let $v \in X$ solve

$$v = t\Phi(v), \qquad t \in [0, 1]. \tag{21.29}$$

Equivalently,

$$\nu v = tB(v, v) + t\widetilde{K}. \tag{21.30}$$

Taking the inner product in X with v, we get

$$\nu\|v\|_X^2 = t\langle B(v, v), v\rangle_X + t\langle \widetilde{K}, v\rangle_X = ta(v, v, v) + t\langle \widetilde{K}, v\rangle_X = t\langle \widetilde{K}, v\rangle_X. \tag{21.31}$$

Using Cauchy's inequality,

$$\nu\|v\|_X^2 \le \|\widetilde{K}\|_X \|v\|_X. \tag{21.32}$$

Thus every solution of (21.29) satisfies

$$\|v\|_X \le \nu^{-1} \|\widetilde{K}\|_X. \tag{21.33}$$

The bound is independent of $t \in [0, 1]$. By the Leray-Schauder principle, Φ has a fixed point. This fixed point satisfies

$$v = \nu^{-1} B(v, v) + \nu^{-1} \widetilde{K},$$

or equivalently

$$\nu v - B(v, v) = \widetilde{K}.$$

Testing against $w \in X$ gives (21.26). □

21.3 Uniqueness for small data

Theorem 21.5 (Small-data uniqueness)**.** *Assume, in addition, that the force is sufficiently small in the sense that*

$$2\alpha\beta^2\|\widetilde{K}\|_X < \nu^2. \tag{21.34}$$

Then the weak solution obtained above is unique.

Proof. Let $v_1, v_2 \in X$ be two weak solutions. Thus

$$\nu v_i - B(v_i, v_i) = \widetilde{K}, \qquad i = 1, 2. \tag{21.35}$$

Set $u = v_1 - v_2$. Subtracting the two equations gives

$$\nu u = B(v_1, v_1) - B(v_2, v_2). \tag{21.36}$$

Taking the inner product with u yields

$$\nu\|u\|_X^2 = \langle B(v_1, v_1) - B(v_2, v_2), u\rangle_X = a(v_1, v_1, u) - a(v_2, v_2, u). \tag{21.37}$$

Using bilinearity in the first two variables,

$$a(v_1, v_1, u) - a(v_2, v_2, u) = a(u, v_1, u) + a(v_2, u, u). \tag{21.38}$$

The second term vanishes, indeed,

$$a(v_2, u, u) = \int_U (v_2)_k u_j \partial_k u_j \,\mathrm{d}x = \frac{1}{2}\int_U (v_2)_k \partial_k(|u(x)|^2)\,\mathrm{d}x = -\frac{1}{2}\int_U \operatorname{div}(v_2)|u(x)|^2\,\mathrm{d}x = 0. \tag{21.39}$$

Therefore

$$\nu\|u\|_X^2 = a(u, v_1, u). \tag{21.40}$$

Using (21.4),

$$\nu\|u\|_X^2 \le \alpha\beta^2\|u\|_X\,\|v_1\|_X\,\|u\|_X = \alpha\beta^2\|v_1\|_X\,\|u\|_X^2. \tag{21.41}$$

By the a priori estimate (21.33), applied to the solution v_1 with $t = 1$, $\|v_1\|_X \le \nu^{-1}\|\widetilde{K}\|_X$. Thus

$$\nu\|u\|_X^2 \le \alpha\beta^2\nu^{-1}\|\widetilde{K}\|_X\,\|u\|_X^2. \tag{21.42}$$

In particular, uniqueness follows under the stronger-than-needed condition $\alpha\beta^2\|\widetilde{K}\|_X < \nu^2$. Alternatively, without using the cancellation in the mixed term, one obtains directly from the local Lipschitz estimate

$$\|B(v_1, v_1) - B(v_2, v_2)\|_X \le \alpha\beta^2(\|v_1\|_X + \|v_2\|_X)\|v_1 - v_2\|_X. \tag{21.43}$$

Using $\|v_i\|_X \le \nu^{-1}\|\widetilde{K}\|_X$, this gives

$$\nu\|u\|_X \le 2\alpha\beta^2\nu^{-1}\|\widetilde{K}\|_X\,\|u\|_X. \tag{21.44}$$

Hence, if $2\alpha\beta^2\|\widetilde{K}\|_X < \nu^2$, then $u = 0$. Therefore $v_1 = v_2$. □

Corollary 21.6 (Smallness condition in terms of the force)**.** *If*

$$2\alpha\beta^2 C_P\,\|K\|_{L^2(U)} < \nu^2, \tag{21.45}$$

then the weak stationary Navier-Stokes solution is unique.

Proof. This follows immediately from (21.11), namely $\|\widetilde{K}\|_X \le C_P\,\|K\|_{L^2(U)}$. □