



# Analysis and Something Else

**Author:** Wenxiao Yang

**Institute:** Haas School of Business, University of California Berkeley

**Date:** 2023

*All models are wrong, but some are useful.*

# Contents

<b>Chapter 1 Logic</b>	<b>1</b>
1.1 Main Methods of Proof (@ Lec 01 of ECON 204) . . . . .	1
1.1.1 Proof by Induction . . . . .	1
1.1.2 Proof by Deduction . . . . .	1
1.1.3 Proof by Contradiction . . . . .	1
1.1.4 Proof by Contraposition . . . . .	1
<b>Chapter 2 Analysis Basis</b>	<b>2</b>
2.1 Real Number $\mathbb{R}$ (@ Lec 02 of ECON 204) . . . . .	2
2.1.1 Order Axiom . . . . .	2
2.1.2 Completeness Axiom . . . . .	2
2.1.3 Supremum sup $\mathbb{X}$ , Infimum inf $\mathbb{X}$ for $\mathbb{X} \subseteq \mathbb{R}$ . . . . .	3
2.1.4 Theorem: sup / inf in set is max / min . . . . .	3
2.1.5 Transformation of sup / inf . . . . .	3
2.1.6 The Supremum Property: Nonempty set in $\mathbb{R}$ is bounded above/below $\Rightarrow \exists$ sup / inf .	3
2.1.7 Archimedean Property . . . . .	4
2.2 Metric Spaces and Normed Spaces (@ Lec 03 of ECON 204) . . . . .	4
2.2.1 Metric Space $(\mathbb{X}, d)$ and Metric $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$ . . . . .	4
2.2.2 Theorem: Subset forms a metric space . . . . .	4
2.2.3 Norm $\ \cdot\ $ and Normed Vector Space $(V, \ \cdot\ )$ . . . . .	4
2.2.4 Subset may not form a normed vector space . . . . .	5
2.2.5 Theorem: metric can be defined by norm . . . . .	5
2.2.6 Cauchy-Schwarz Inequality . . . . .	5
2.2.7 Lipschitz-equivalent Norm . . . . .	6
2.2.8 Ball, Radius, Diameter, and Distance . . . . .	6
2.3 Set Theory . . . . .	7
2.3.1 Well Defined Set . . . . .	7
2.3.2 Numerically Equivalent (@ Lec 01 of ECON 204) . . . . .	7
2.3.3 Finite, Countable Set (@ Lec 01 of ECON 204) . . . . .	7

---

2.3.4	Power Set (@ Lec 02 of ECON 204) . . . . .	7
2.3.5	Theorem (Cantor): The power set of $\mathbb{N}$ is uncountable (@ Lec 02 of ECON 204) . . . . .	7
2.3.6	Cardinalities of Sets (@ Lec 02 of ECON 204) . . . . .	8
2.3.7	Pigeonhole Principle: $ A  >  B  \Rightarrow$ no injective function $\sigma : A \rightarrow B$ . . . . .	8
2.3.8	$B^A$ : Sets of Function . . . . .	8
2.3.9	Bounded Set (@ Lec 03 of ECON 204) . . . . .	9
2.3.10	Open, Closed Set (@ Lec 04 of ECON 204) . . . . .	9
2.3.11	Theorem: Union of open sets is open, Intersection of finite open sets is open . . . . .	10
2.3.12	Interior, Exterior, Boundary, Closure (@ Lec 04 of ECON 204) . . . . .	10
2.3.13	Dense Set: $\bar{A} = X$ . . . . .	11
2.3.14	Compact Set . . . . .	12
2.3.15	Sublevel Set . . . . .	12
2.3.16	Set Operations . . . . .	12
2.3.17	Connected Sets (@ Lec 07 of ECON 204) . . . . .	12
2.3.18	Theorem: Connected Set over $\mathbb{E}^1 \Leftrightarrow$ Interval (@ Lec 07 of ECON 204) . . . . .	13
2.3.19	Theorem: Continuous Image of Connected Set is Connected . . . . .	13
2.3.20	Intermediate Value Theorem proved by connectedness . . . . .	13
2.4	Sequences . . . . .	13
2.4.1	Convergence of Sequences (@ Lec 03 of ECON 204) . . . . .	14
2.4.2	Limit Point (@ Lec 03 of ECON 204) . . . . .	14
2.4.3	Cluster Point (@ Lec 03 of ECON 204) . . . . .	14
2.4.4	Proposition: Sequences in $\mathbb{R}^n$ , Bounded above and non-decreasing sequence $\Rightarrow$ Converge	15
2.4.5	Theorem: Increasing/Decreasing Sequences in $\mathbb{R}^n$ , Limit is sup/inf (@ Lec 03 of ECON 204) . . . . .	15
2.4.6	Definition: $\limsup$ and $\liminf$ (@ Lec 03 of ECON 204) . . . . .	15
2.4.7	Exists Limit $\Leftrightarrow \lim = \limsup = \liminf$ (@ Lec 03 of ECON 204) . . . . .	16
2.4.8	Rising Sun Lemma: Sequence in $\mathbb{R}^n$ contains monotone subsequence (@ Lec 03 of ECON 204) . . . . .	16
2.4.9	Bolzano-Weierstrass: Bounded Sequence in $\mathbb{R}^n$ contains a convergent subsequence (@ Lec 03 of ECON 204) . . . . .	16
2.5	Complete Metric Spaces (@ Lec 05 of ECON 204) . . . . .	16
2.5.1	Cauchy Sequence . . . . .	16

---

2.5.2	Theorem: Convergent $\Rightarrow$ Cauchy . . . . .	17
2.5.3	Complete Metric Space and Banach Space . . . . .	17
2.5.4	Theorem: Subset of Complete Metric Space is Complete $\Leftrightarrow$ Subset is Closed . . . . .	17
2.6	Compactness over Metric Spaces (@ Lec 06 of ECON 204) . . . . .	18
2.6.1	Open Cover . . . . .	18
2.6.2	Definition of Compact Set . . . . .	18
2.6.3	Theorem: Closed Subset of Compact Set is Compact . . . . .	18
2.6.4	Theorem: Compact Set over Metric Space is Closed . . . . .	19
2.6.5	Definition of Sequentially Compact . . . . .	19
2.6.6	Theorem: Compact $\Leftrightarrow$ Sequentially Compact . . . . .	19
2.6.7	Definition of Totally Bounded . . . . .	20
2.6.8	Theorem: Compact $\Leftrightarrow$ Complete/Closed and Totally Bounded . . . . .	20
2.6.9	Heine-Borel Theorem: With Euclidean Metric, Compact $\Leftrightarrow$ Closed and Bounded . . . . .	21
2.6.10	Theorem: Continuous Images of Compact Set is Compact . . . . .	21
<b>Chapter 3 Functions</b>		<b>22</b>
3.1	Definitions of Function . . . . .	22
3.1.1	Image, Preimage, Fiber . . . . .	22
3.1.2	Function Image's Properties . . . . .	22
3.1.3	Composition of functions . . . . .	23
3.1.4	Function Composition is Associative . . . . .	23
3.1.5	Homeomorphism (@ Lec 04 of ECON 204) . . . . .	23
3.2	Injection, Surjection, Bijection . . . . .	23
3.2.1	Definitions: Injective, surjective, bijective . . . . .	23
3.2.2	Lemma 1.1.7: injective/surjective/bijective is preserved in composition . . . . .	23
3.2.3	Theorem: A function is bijective $\Leftrightarrow \exists$ inverse . . . . .	24
3.3	Function Continuity . . . . .	24
3.3.1	Continuous Function in $\mathbb{R}$ with Euclidean norm . . . . .	24
3.3.2	Continuous Function in Metric Spaces (@ Lec 04 of ECON 204) . . . . .	24
3.3.3	Theorem: Continuous $\Leftrightarrow$ Preimage of open set is open (@ Lec 04 of ECON 204) . . . . .	25
3.3.4	Theorem: Continuity is preserved in composition (@ Lec 04 of ECON 204) . . . . .	26
3.3.5	Uniform Continuity (@ Lec 04 of ECON 204) . . . . .	26

3.3.6	Theorem: Continuous function with Compact domain is Uniformly Continuous (@ Lec 06 of ECON 204) . . . . .	26
3.3.7	Lipschitz Continuous (@ Lec 04 of ECON 204) . . . . .	27
3.3.8	Contraction (from the view of Lipschitz constant) . . . . .	28
3.4	Extreme Value, Intermediate Value of Functions . . . . .	28
3.4.1	Intermediate Value Theorem (@ Lec 02 of ECON 204) . . . . .	28
3.4.2	Coercive Function . . . . .	28
3.4.3	Extreme of Functions . . . . .	29
3.4.4	Weierstrass' Theorem (Extreme value Theorem) (@ Lec 06 of ECON 204) . . . . .	29
3.4.5	Corollary: Continuous and Coercive function attains its Minimum on Closed Set . . . . .	30
3.5	Monotonic Function (@ Lec 05 of ECON 204) . . . . .	30
3.5.1	Theorem: Monotonically Increasing $\Rightarrow$ One-sided Limits exist . . . . .	30
3.5.2	Theorem: Monotonically Increasing $\Rightarrow$ Set of Points are discontinuous is finite (possibly empty) or countable . . . . .	31
3.6	Correspondence (@ Lec 07 of ECON 204) . . . . .	31
3.6.1	Definition of Correspondence . . . . .	31
3.6.2	Continuity of Correspondences . . . . .	32
3.6.3	Graph of Correspondence . . . . .	32
3.6.4	Theorem: $\Psi(x) = \{f(x)\}$ is uhc $\Leftrightarrow f$ is continuous . . . . .	34
3.6.5	Closed-valued and Compact-valued Correspondences . . . . .	34
3.7	Differentiable Function (@ Lec 11 of ECON 204) . . . . .	35
3.7.1	Definition of Differentiable in $\mathbb{R}$ . . . . .	35
3.7.2	General Definition of Differentiable . . . . .	35
3.7.3	" $O$ " and " $o$ " Notations . . . . .	36
3.7.4	Jacobian and Error Term . . . . .	36
3.7.5	Theorem: Differentiable $\Rightarrow$ Partial Derivatives (and Jacobian) exist . . . . .	37
3.7.6	Theorem: Partial Derivatives exist and Continuous $\Rightarrow$ Differentiable . . . . .	37
3.7.7	Directional Derivatives . . . . .	37
3.7.8	Chain Rule: $dF_{x_0} = dg_{f(x_0)} \circ df_{x_0}$ and $DF_{x_0} = Dg(f(x_0))Df(x_0)$ . . . . .	38
3.7.9	Mean Value Theorem: $\exists z \in l(x, y), \ f(y) - f(x)\  \leq \ df_z\  \ y - x\ $ . . . . .	38
3.7.10	Taylor's Theorem . . . . .	39
3.7.11	$C^k$ Function and Continuously Differentiable . . . . .	39

---

3.7.12 Taylor's Theorem with $C^k$	39
<b>Chapter 4 Tools for Comparative Statics</b>	<b>42</b>
4.1 Regular and Critical Points and Values	42
4.1.1 Rank of Derivatives $\text{Rank}df_x = \text{Rank}Df(x)$	42
4.1.2 Regular and Critical Points and Values	42
4.2 Inverse and Implicit Function Theorem	43
4.2.1 Inverse Function Theorem	43
4.2.2 Implicit Function Theorem	43
4.2.3 Prove Implicit Function Theorem Given Inverse Function Theorem	44
4.2.4 Prove Inverse Function Theorem Given Implicit Function Theorem	45
4.2.5 Example: Using Implicit Function Theorem in Comparative Statics	45
4.2.6 Corollary: $a \rightarrow \{x \in X : f(x, a) = 0\}$ is lhc	46
4.3 Transversality and Genericity	46
4.3.1 Lebesgue Measure Zero	46
4.3.2 Sard's Theorem	47
4.3.3 Transversality Theorem	47
<b>Chapter 5 Fixed Point Theorem</b>	<b>48</b>
5.1 Contraction Mapping Theorem (@ Lec 05 of ECON 204)	48
5.1.1 Contractions	48
5.1.2 Theorem: Contraction $\Rightarrow$ Uniformly Continuous	48
5.1.3 Theorem: Blackwell's Sufficient Conditions for Contraction	48
5.2 Fixed Point Theorem (@ Lec 05 of ECON 204)	49
5.2.1 Fixed Point	49
5.2.2 Contraction Mapping Theorem: Existence and Uniqueness of Fixed Point	50
5.2.3 Theorem: Fixed Point's Continuous Dependence on Parameters	51
5.3 Brouwer's Fixed Point Theorem (@ Lec 13 of ECON 204)	51
5.3.1 Simple One: One-dimension	51
5.3.2 Brouwer's Fixed Point Theorem	52
5.4 Fixed Points for Correspondences (@ Lec 13 of ECON 204)	52
5.4.1 Definition	52
5.4.2 Kakutani's Fixed Point Theorem	52

---

<b>Chapter 6 Differential Equations (@ Lec 14/15 of ECON 204)</b>	<b>53</b>
6.1 Initial Value Problem . . . . .	53
6.1.1 Initial Value Problem . . . . .	53
6.1.2 Theorem: Initial Value Problem, Continuous → Existence; Lipschitz → Uniqueness .	53
6.2 Autonomous Differential Equation . . . . .	54
6.2.1 Autonomous Differential Equation . . . . .	54
6.2.2 Complex Exponentials . . . . .	55
6.2.3 Solutions of Simple Case: $y'(t) = ay(t)$ . . . . .	55
6.2.4 Solutions of Special Case: Linear Differential Equations with Constant Coefficients .	55
6.2.5 The Form of Real Solutions . . . . .	58
6.3 Higher Order Differential Equations . . . . .	59
6.3.1 Second Order Linear Differential Equations . . . . .	59
6.4 Inhomogeneous Linear Differential Equations with Nonconstant Coefficients . . . . .	59
6.4.1 Theorem: General Solution of Inhomo = General Solution of Homo + Particular Solution of Inhomo . . . . .	60
6.4.2 Matrix Exponentials Notations . . . . .	60
6.4.3 Theorem: Particular Solution of Inhomogeneous Linear Differential Equations with Constant Coefficients (Matrix) . . . . .	61
<b>Chapter 7 Big <math>\mathcal{O}</math> and Small <math>o</math> Notation</b>	<b>62</b>
7.1 Definition . . . . .	62
7.1.1 Extension . . . . .	63

# Chapter 1 Logic

## 1.1 Main Methods of Proof (@ Lec 01 of ECON 204)

### 1.1.1 Proof by Induction

### 1.1.2 Proof by Deduction

### 1.1.3 Proof by Contradiction

### 1.1.4 Proof by Contraposition

- $\neg P$  ("not  $P$ ") means " $P$  is false".
- $P \wedge Q$  (" $P$  and  $Q$ ") means " $P$  is true and  $Q$  is true."
- $P \vee Q$  (" $P$  or  $Q$ ") means " $P$  is true or  $Q$  is true (or possibly both)."
- $\neg P \wedge Q$  means  $(\neg P) \wedge Q$ ;  $\neg P \vee Q$  means  $(\neg P) \vee Q$ .
- $P \Rightarrow Q$  (" $P$  implies  $Q$ ") means "whenever  $P$  is satisfied,  $Q$  is also satisfied."

**Statement:** Formally,  $P \Rightarrow Q$  is equivalent to  $\neg P \vee Q$ .

#### Definition 1.1 (Contrapositive)

The *contrapositive* of the statement  $P \Rightarrow Q$  is the statement  $\neg Q \Rightarrow \neg P$ .



#### Theorem 1.1 (Prove Contrapositive Instead)

$P \Rightarrow Q$  is true if and only if  $\neg Q \Rightarrow \neg P$  is true.



#### Proposition 1.1

$[(P \wedge Q) \Rightarrow R] \Leftrightarrow [P \Rightarrow (Q \Rightarrow R)]$ : Given  $P$ , if  $Q$  then  $R$ .

$\Leftrightarrow [(P \Rightarrow R) \vee (Q \Rightarrow R)]$ : Either  $P \Rightarrow R$  holds or  $Q \Rightarrow R$  holds



## Chapter 2 Analysis Basis

### 2.1 Real Number $\mathbb{R}$ (@ Lec 02 of ECON 204)

$\mathbb{R}$  is a field with the usual operations  $+$ ,  $\cdot$ , additive identity 0, and multiplicative identity 1.

#### 2.1.1 Order Axiom

##### Proposition 2.1 (Order Axiom)

There is a complete ordering  $\leq$ , i.e.  $\leq$  is reflexive, transitive, antisymmetric ( $\alpha \leq \beta, \beta \leq \alpha \Rightarrow \alpha = \beta$ ) with the property that  $\forall \alpha, \beta \in \mathbb{R}$  either  $\alpha \leq \beta$  or  $\beta \leq \alpha$ .

The order is compatible with  $+$  and  $\cdot$ , i.e.  $\forall \alpha, \beta, \gamma \in \mathbb{R}$

1.  $\alpha \leq \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma$ .
2.  $\alpha \leq \beta, 0 \leq \gamma \Rightarrow \alpha\gamma \leq \beta\gamma$ .



#### 2.1.2 Completeness Axiom

##### Proposition 2.2 (Completeness Axiom)

Suppose  $L, H \subseteq \mathbb{R}, L \neq \emptyset \neq H$  satisfy  $\forall l \in L, h \in H, l \leq h$ . Then,

$$\exists \alpha \in \mathbb{R} \text{ s.t. } \forall l \in L, h \in H, l \leq \alpha \leq h$$

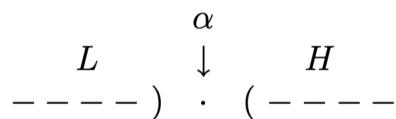


Figure 2.1: Completeness Axiom

##### Claim 2.1

The Completeness Axiom differentiates  $\mathbb{R}$  from  $\mathbb{Q}$ :

$\mathbb{Q}$  satisfies all the axioms for  $\mathbb{R}$  except the Completeness Axiom.



### 2.1.3 Supremum $\sup \mathbb{X}$ , Infimum $\inf \mathbb{X}$ for $\mathbb{X} \subseteq \mathbb{R}$

#### Definition 2.1 (Supremum and Infimum)

- (1). Suppose  $\mathbb{X}$  is bounded above. The **supremum** of  $\mathbb{X}$ , written  $\sup \mathbb{X}$ , is the least upper bound for  $\mathbb{X}$ , i.e.  $\sup \mathbb{X}$  satisfies
  - (a).  $\sup \mathbb{X} \geq x, \forall x \in \mathbb{X}$  ( $\sup \mathbb{X}$  is an upper bound).
  - (b).  $\forall y < \sup \mathbb{X}, \exists x \in \mathbb{X}$  s.t.  $x > y$  (there is no smaller upper bound).
- (2). Suppose  $\mathbb{X}$  is bounded below. The **infimum** of  $\mathbb{X}$ , written  $\inf \mathbb{X}$ , is the greatest lower bound for  $\mathbb{X}$ , i.e.  $\inf \mathbb{X}$  satisfies
  - (a).  $\inf \mathbb{X} \leq x, \forall x \in \mathbb{X}$  ( $\inf \mathbb{X}$  is a lower bound).
  - (b).  $\forall y > \inf \mathbb{X}, \exists x \in \mathbb{X}$  s.t.  $x < y$  (there is no greater lower bound).
- (3). If  $\mathbb{X}$  is not bounded above, write  $\sup \mathbb{X} = \infty$ . If  $\mathbb{X}$  is not bounded below, write  $\inf \mathbb{X} = -\infty$ . By convention,  $\sup \emptyset = -\infty$ ,  $\inf \emptyset = +\infty$ .



### 2.1.4 Theorem: sup / inf in set is max / min

#### Proposition 2.3 (sup / inf in set is max / min)

If  $\inf A = x^* \in A$  ( $\sup A = x^* \in A$ ), then  $x^* = \min A$  ( $x^* = \max A$ ).



### 2.1.5 Transformation of sup / inf

#### Theorem 2.1

Let  $A$  and  $B$  be nonempty sets of real numbers,

1. both of them bounded above,  $\sup\{a + b : a \in A, b \in B\} = \sup A + \sup B$ .
2. bounded above and below respectively,  $\sup\{a - b : a \in A, b \in B\} = \sup A - \inf B$
3.  $\inf A = -\sup(-A)$



### 2.1.6 The Supremum Property: Nonempty set in $\mathbb{R}$ is bounded above/below $\Rightarrow \exists \sup / \inf$

#### Proposition 2.4 (The Supremum Property)

Every nonempty set of real numbers that is bounded above has a supremum, which is a real number.

Every nonempty set of real numbers that is bounded below has an infimum, which is a real number.



#### Theorem 2.2

The Supremum Property (Prop 2.4) and the Completeness Axiom (Prop 2.2) are equivalent.



### 2.1.7 Archimedean Property

**Theorem 2.3 (Archimedean Property)**

$\forall x \in \mathbb{R}, y \in \mathbb{R}^+, \exists n \in \mathbb{N}$  s.t.  $ny > x$ .



## 2.2 Metric Spaces and Normed Spaces (@ Lec 03 of ECON 204)

### 2.2.1 Metric Space $(\mathbb{X}, d)$ and Metric $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$

**Definition 2.2 (Metric Space)**

A **metric space** is a pair  $(\mathbb{X}, d)$ , where  $\mathbb{X}$  is a set and  $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$  a function satisfying

1. Non-negative:  $d(x, y) \geq 0, d(x, y) = 0 \Leftrightarrow x = y, \forall x, y \in \mathbb{X}$ .
2. Symmetric:  $d(x, y) = d(y, x), \forall x, y \in \mathbb{X}$ .
3. Triangle inequality:  $d(x, z) \leq d(x, y) + d(y, z), \forall x, y, z \in \mathbb{X}$ .

A function  $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$  satisfying 1-3 is called a **metric** on  $\mathbb{X}$ .



A metric gives a notion of distance between elements of  $\mathbb{X}$ .

### 2.2.2 Theorem: Subset forms a metric space

**Theorem 2.4 (Subset forms a metric space)**

Given a metric space  $(X, d)$ ,  $(Y, d)$  for some  $Y \subseteq X$  is still a metric space.



### 2.2.3 Norm $\|\cdot\|$ and Normed Vector Space $(V, \|\cdot\|)$

**Definition 2.3 (Norm)**

Let  $V$  be a vector space over  $\mathbb{R}$ . A **norm** on  $V$  is a function  $\|\cdot\| : V \rightarrow \mathbb{R}_+$  satisfying

1. Non-negative:  $\|x\| \geq 0, \|x\| = 0 \Leftrightarrow x = 0, \forall x \in V$ .
2. Triangle inequality:  $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in V$ .
3.  $\|\alpha x\| = |\alpha| \|x\|, \forall \alpha \in \mathbb{R}, x \in V$ .



A norm gives a notion of length of a vector in  $V$ .

**Definition 2.4 (Normed Vector Space)**

A **normed vector space** is a vector space over  $\mathbb{R}$  equipped with a norm,  $(V, \|\cdot\|)$ .



### Example 2.1 Normed Vector Space

- $\mathbf{E}^n$  :  $n$ -dimensional Euclidean space.

$$V = \mathbb{R}^n, \|x\|_2 = |x| = \sqrt{\sum_{i=1}^n (x_i)^2}$$

- $V = \mathbb{R}^n, \|x\|_1 = \sum_{i=1}^n |x_i|$  (the "taxi cab" norm or  $L^1$  norm)
- $V = \mathbb{R}^n, \|x\|_\infty = \max \{|x_1|, \dots, |x_n|\}$  (the maximum norm, or sup norm, or  $L^\infty$  norm)
- $C([0, 1]), \|f\|_\infty = \sup\{|f(t)| : t \in [0, 1]\}$
- $C([0, 1]), \|f\|_2 = \sqrt{\int_0^1 (f(t))^2 dt}$
- $C([0, 1]), \|f\|_1 = \int_0^1 |f(t)| dt$

where  $C([0, 1])$  is the space of all continuous real-valued functions on  $[0, 1]$ .

#### 2.2.4 Subset may not form a normed vector space

Given a normed vector space  $(V, \|\cdot\|)$ , is  $(W, \|\cdot\|)$  for some  $W \subseteq V$ ,  $W$  may not be a vector space (e.g. it may not contain the zero vector)!

#### 2.2.5 Theorem: metric can be defined by norm

In an abstract normed vector space, the norm can be used analogously to define a notion of distance.

**Theorem 2.5 (Metric can be defined by norm)**

Let  $(V, \|\cdot\|)$  be a normed vector space. Let  $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$  be defined by  $d(v, w) = \|v - w\|$ . Then  $(V, d)$  is a metric space.



#### 2.2.6 Cauchy-Schwarz Inequality

**Theorem 2.6 (Cauchy-Schwarz Inequality)**

If  $v, w \in \mathbb{R}^n$ , then

$$\left( \sum_{i=1}^n v_i w_i \right)^2 \leq \left( \sum_{i=1}^n v_i^2 \right) \left( \sum_{i=1}^n w_i^2 \right)$$

$$\|v \cdot w\|^2 \leq \|v\|^2 \|w\|^2$$

$$\|v \cdot w\| \leq \|v\| \|w\|$$



### 2.2.7 Lipschitz-equivalent Norm

#### Definition 2.5 (Lipschitz-equivalent)

Two norms  $\|\cdot\|$  and  $\|\cdot\|^*$  on the same vector space  $V$  are said to be **Lipschitz-equivalent** (or **equivalent**) if  $\exists m, M > 0$  s.t.  $\forall x \in V$ ,

$$m\|x\| \leq \|x\|^* \leq M\|x\|$$

Equivalently,  $\exists m, M > 0$  s.t.  $\forall x \in V, x \neq 0$ ,

$$m \leq \frac{\|x\|^*}{\|x\|} \leq M$$



#### Theorem 2.7

All norms on  $\mathbb{R}^n$  are equivalent.



However, infinite-dimensional spaces support norms that are not equivalent. For example, on  $C([0, 1])$ , let  $f_n$  be the function

$$f_n(t) = \begin{cases} 1 - nt, & \text{if } t \in [0, \frac{1}{n}] \\ 0, & \text{if } t \in (\frac{1}{n}, 1] \end{cases}$$

Then  $\frac{\|f_n\|_1}{\|f_n\|_\infty} = \frac{1}{2n} \rightarrow 0$ , which means there is no lower bound  $m > 0$ .

### 2.2.8 Ball, Radius, Diameter, and Distance

In a metric space  $(X, d)$ , define

$$\begin{aligned} B_\varepsilon(x) &= \{y \in X : d(y, x) < \varepsilon\} \\ &= \text{open ball with center } x \text{ and radius } \varepsilon \\ B_\varepsilon[x] &= \{y \in X : d(y, x) \leq \varepsilon\} \\ &= \text{closed ball with center } x \text{ and radius } \varepsilon \end{aligned}$$

We can use the metric  $d$  to define a generalization of "radius". In a metric space  $(X, d)$ , define the *diameter* of a subset  $S \subseteq X$  by

$$\text{diam}(S) = \sup \{d(s, s') : s, s' \in S\}$$

Similarly, we can define the *distance from a point to a set*, and *distance between sets*, as follows:

$$\begin{aligned} d(A, x) &= \inf_{a \in A} d(a, x) \\ d(A, B) &= \inf_{a \in A} d(B, a) \\ &= \inf \{d(a, b) : a \in A, b \in B\} \end{aligned}$$

## 2.3 Set Theory

### 2.3.1 Well Defined Set

#### Definition 2.6

A set  $S$  is **well defined** if an object  $a$  is either  $a \in S$  or  $a \notin S$ .



### 2.3.2 Numerically Equivalent (@ Lec 01 of ECON 204)

#### Definition 2.7

Two sets  $A, B$  are **numerically equivalent** (or have the same cardinality) if there is a bijection  $f : A \rightarrow B$ , that is, 1-1 ( $a \neq a' \Rightarrow f(a) \neq f(a')$ ), and onto ( $\forall b \in B, \exists a \in A$  s.t.  $f(a) = b$ ).



### 2.3.3 Finite, Countable Set (@ Lec 01 of ECON 204)

#### Definition 2.8 (Finite Set)

A set is either **finite** or **infinite**. A set is **finite** if it is numerically equivalent to  $\{1, \dots, n\}$  for some  $n$ . A set that is not finite is infinite.



We give a more precise definition to classify infinite set:

#### Definition 2.9 (Countable Set)

An infinite set is **countable** if it is numerically equivalent to  $\mathbb{N}$ . An infinite set that is not countable is called **uncountable**.



#### Theorem 2.8 (Countable $\mathbb{Q}$ )

*The set of rational numbers  $\mathbb{Q}$  is countable.*



### 2.3.4 Power Set (@ Lec 02 of ECON 204)

#### Definition 2.10 (Power Set: the set of all subsets)

For any set  $A$ , we denote by  $\mathcal{P}(A)$  the collection of all subsets of  $A$ .  $\mathcal{P}(A)$  is the **power set** of  $A$ .



We may also use the notation  $2^A$  (in Berkeley ECON 204).

### 2.3.5 Theorem (Cantor): The power set of $\mathbb{N}$ is uncountable (@ Lec 02 of ECON 204)

#### Theorem 2.9 (Cantor)

$\mathcal{P}(\mathbb{N})$  (or denoted by  $2^{\mathbb{N}}$ ), the set of all subsets of  $\mathbb{N}$ , is uncountable.



### 2.3.6 Cardinalities of Sets (@ Lec 02 of ECON 204)

#### Definition 2.11 (Cardinality)

If  $A$  is a set,  $|A| = \text{cardinality of } A = \# \text{ of elements}$



$n \in \mathbb{N}, |\{1, \dots, n\}| = n; |\emptyset| = 0 (\emptyset = \text{empty set}).$

#### Proposition 2.5 (Facts about cardinality)

1. If  $A$  is numerically equivalent to  $\{1, \dots, n\}$  for some  $n \in \mathbb{N}$ , then  $|A| = n$ .
2.  $A$  and  $B$  are numerically equivalent if and only if  $|A| = |B|$ .
3. If  $|A| = n$  (finite) and  $A$  is a proper subset of  $B$  (that is,  $A \subset B$  and  $A \neq B$ ) then  $|A| < |B|$ .
4. If  $A$  is countable and  $B$  is uncountable, then  $n < |A| < |B|, \forall n \in \mathbb{N}$ .
5. If  $A \subseteq B$ , then  $|A| \leq |B|$ . (if  $B$  is countable and  $A \subseteq B$ , then  $A$  is at most countable, that is,  $A$  is either empty, finite, or countable.)
6. If there is an injection  $\sigma : A \rightarrow B$ , we can write  $|A| \leq |B|$ ;
7. If there is a surjection  $\sigma : A \rightarrow B$ , we can write  $|A| \geq |B|$ ;
8. If there is a bijection  $\sigma : A \rightarrow B$ , we can write  $|A| = |B|$ .



### 2.3.7 Pigeonhole Principle: $|A| > |B| \Rightarrow \text{no injective function } \sigma : A \rightarrow B$

#### Theorem 2.10 (Pigeonhole Principle)

If  $A$  and  $B$  are sets and  $|A| > |B|$ , then there is no injective function  $\sigma : A \rightarrow B$ .



### 2.3.8 $B^A$ : Sets of Function

If  $A, B$  are sets, then  $B^A = \{\sigma : A \rightarrow B | \sigma \text{ a function}\}$ .

**Example 2.2**  $n \in \mathbb{Z}$ , we define a function  $f : B^{\{1, \dots, n\}} \rightarrow B^n (= B \times B \times B \times \dots \times B)$  by the equation  $f(\sigma) = \{\sigma(1), \dots, \sigma(n)\}$ , where  $\sigma : \{1, \dots, n\} \rightarrow B$ . The  $f$  is a bijection.

#### Proof 2.1

1. *Injective:*

$$f(\sigma_1) = f(\sigma_2) \Rightarrow \{\sigma_1(1), \dots, \sigma_1(n)\} = \{\sigma_2(1), \dots, \sigma_2(n)\}$$

Since  $\sigma : \{1, \dots, n\} \rightarrow B$ , it is sufficient to prove  $\sigma_1 = \sigma_2$ .

2. *Surjective:*

$$\forall \{b_1, \dots, b_n\}, \text{ we have } \sigma^*(x) = b_x, x = 1, \dots, n. \text{ s.t. } f(\sigma^*) = \{b_1, \dots, b_n\}$$

**Example 2.3**

$$C(\mathbb{R}, \mathbb{R}) = \{\text{continuous functions } \sigma : \mathbb{R} \rightarrow \mathbb{R}\} \subset \mathbb{R}^{\mathbb{R}}$$

**2.3.9 Bounded Set (@ Lec 03 of ECON 204)****Definition 2.12 (Bounded Set)**

In a metric space  $(X, d)$ , a subset  $S \subseteq X$  is **bounded** if  $\exists x \in X, \beta \in \mathbb{R}$  such that  $\forall s \in S, d(s, x) \leq \beta$ .



Equivalent:  $\exists \beta > 0$  and  $s \in S$  such that  $S \subseteq B_{\beta}(s)$ .

**2.3.10 Open, Closed Set (@ Lec 04 of ECON 204)****Definition 2.13 (Open Sets)**

Let  $(X, d)$  be a metric space. A set  $\mathbb{X} \subseteq \mathbb{R}^n$  is **open** if

$\forall x \in \mathbb{X}$  we can draw a ball around  $x$  that is contained in  $\mathbb{X}$ .

i.e.  $\forall x \in \mathbb{X}, \exists \varepsilon > 0$  s.t.  $B_{\varepsilon}(x) = \{y : d(y, x) < \varepsilon\} \subseteq \mathbb{X}$



**Example 2.4** In a metric space  $(X, d)$ , if  $a$  is an isolated point of  $X$ , then the singleton set  $\{a\}$  is open: Since  $a$  is not a limit point of  $X$ , we can find a neighborhood  $B_{\varepsilon}(a)$  such that  $B_{\varepsilon}(a) \cap X = \{a\}$ . So  $B_{\varepsilon}(a) \subseteq \{a\}$  and  $x$  is an interior point. Hence, every point is an interior point and the set is open.

**Definition 2.14 (Closed Sets)**

$\mathbb{X}$  is **closed** if  $\mathbb{X}^c$  is open.

**Theorem 2.11 (Equivalent definition: Closed Sets)**

Equivalent: if  $A$  in a metric space  $(X, d)$  contains all limit points of all sequences in  $A$ ,  $A$  is closed.

$$\{x_n\} \subset A, x_n \rightarrow x \in X \Rightarrow x \in A$$

**Example 2.5 (Closed and Open Sets on  $\mathbb{E}_1$  i.e.,  $\mathbb{R}$  with the usual Euclidean metric)**

- 1)  $(1, 2) = \{x \in \mathbb{R} : 1 < x < 2\}$  - open
- 2)  $\mathbb{R}$  is both open and closed
- 3)  $(-\infty, 1) = \{x \in \mathbb{R} : x < 1\}$  - open
- 4)  $[1, \infty)$  is closed because its complement open
- 5)  $(1, 2]$  is neither open nor closed

**Example 2.6 (Closed and Open Sets on other metric space)** In the metric space  $[0, 1]$ ,  $[0, 1]$  is open. With  $[0, 1]$  as the underlying metric space,  $B_{\varepsilon}(0) = \{x \in [0, 1] : |x - 0| < \varepsilon\} = [0, \varepsilon]$ .

**Theorem 2.12 (Empty Set and Full Set are both open and closed)**

In any metric space  $(X, d)$  both  $\emptyset$  and  $X$  are open, and both  $\emptyset$  and  $X$  are closed.

**Theorem 2.13 (Every neighborhood is an open set)**

Every neighborhood  $B_\varepsilon(x)$  is an open set.

**Theorem 2.14 (Finite sets are closed in any metric space)**

Finite sets are closed in any metric space  $(X, d)$ .

**Proposition 2.6**

Any set in a metric space  $(X, d)$  can be written as the intersection of open sets.

**2.3.11 Theorem: Union of open sets is open, Intersection of finite open sets is open****Theorem 2.15 (Union of open sets is open, Intersection of finite open sets is open)**

In any metric space  $(X, d)$ ,

1. The union of an arbitrary (finite, countable, or uncountable) collection of open sets is open.
2. The intersection of a finite collection of open sets is open.

**2.3.12 Interior, Exterior, Boundary, Closure (@ Lec 04 of ECON 204)**

Given a set  $S \subseteq X$ , the **point** of  $X$  can be classified into three types relative to  $S$ :

- **Interior (points)**, denoted  $\text{int}(S)$ :  $\vec{x} \in S$  for which there exists some  $B(\vec{x}, r) \subseteq S$ , is the largest open set contained in  $S$  (the union of all open sets contained in  $S$ ).

$$x \in \text{int } A \Leftrightarrow \exists \varepsilon > 0 : B_\varepsilon(x) \subseteq A$$

- **Exterior (points)**, denoted  $\text{ext}(S)$ :  $\vec{x} \notin S$  for which there exists some  $B(\vec{x}, r)$  containing no points of  $S$ , is the largest open set contained in  $X \setminus S$ .

$$x \in \text{ext } A \Leftrightarrow \exists \varepsilon > 0 : B_\varepsilon(x) \subseteq A^C$$

- **Boundary (points)** denoted  $\partial(S)$  or  $\text{bd}(S)$ : all other points (for which any  $B(\vec{x}, r)$  will contain some points of  $S$  and some points outside  $S$ ).

$$x \in \partial A \Leftrightarrow \forall \varepsilon > 0 : B_\varepsilon(x) \cap A \neq \emptyset \text{ and } B_\varepsilon(x) \cap A^C \neq \emptyset$$

- **Closure of  $S$** , denoted  $\bar{S}$  or  $\text{cl}(S) = \text{int}(S) \cup \partial(S)$ , is the smallest closed set containing  $S$  (the intersection of all closed sets containing  $S$ ).

$$x \in \text{cl } A \Leftrightarrow \forall \varepsilon > 0 : B_\varepsilon(x) \cap A \neq \emptyset$$

- Moreover, boundary satisfies  $\partial(S) = \overline{(X \setminus S)} \cap \bar{S}$ .

### Proposition 2.7

- (1)  $\text{int}A$ ,  $\text{ext}A$  and  $\partial A$  partition  $X$ .
- (2)  $\text{int}A$  and  $\partial A$  partition  $\text{cl}A$ .
- (3) A set  $S$  is **open** iff  $S = \text{int}(S)$  - i.e.,  $S$  does not contain any of its boundary points.
- (4) A set  $S$  is **closed** iff  $S = \bar{S} = \text{int}(S) \cup \partial(S)$  - i.e.,  $S$  contains all of its boundary points.
- (5) A set  $S$  is **open** and **closed** iff  $\partial(S) = \emptyset$ .
- (6)  $\overline{A \cap B} \subseteq \bar{A} \cap \bar{B}$ .



### 2.3.13 Dense Set: $\bar{A} = X$

#### Definition 2.15 (Dense Set)

Let  $(X, d)$  be a metric space. A set  $A \subseteq X$  is **dense** in  $X$  if  $\bar{A} = X$ .



**Example 2.7**  $\mathbb{Q}$  is dense in  $\mathbb{R}$ : for all  $x \in \mathbb{Q}$ , let  $y_n = \frac{\lfloor 10^n x \rfloor}{10^n}$  with  $n \in \mathbb{N}$ . Then,  $\{y_n\} \subseteq \mathbb{Q}$ , and  $y_n \rightarrow x$  since  $|y_n - x| < 10^{-n}$ . Then  $\exists N \in \mathbb{N}$  s.t.  $\forall n > N$ ,  $|y_n - x| < \varepsilon$  (or  $y_n \in B_\varepsilon(x)$ ). This implies  $B_\varepsilon(x) \cap \mathbb{Q} \neq \emptyset$ , so  $\bar{\mathbb{Q}} = \mathbb{R}$ .

#### Corollary 2.1 ( $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$ )

$\mathbb{Q}^n$  is dense in  $\mathbb{R}^n$  for  $n \in \mathbb{N}$ .



#### Theorem 2.16

If  $A$  is dense in  $X$ , for all  $x \in X$ :  $\forall \varepsilon > 0$ ,  $B_\varepsilon(x) \cap A \neq \emptyset$ .



#### Proof 2.2

Since  $A$  is dense in  $X$ , then for any  $x \in X = \bar{A}$ ,  $\forall \varepsilon > 0$ ,  $B_\varepsilon(x) \cap A \neq \emptyset$  by the definition of closure.

**Example 2.8** If  $\mathcal{A}$  is a collection of open subsets of  $\mathbb{R}^n$ , pairwise disjoint, that is,  $A_\lambda \cap A_{\lambda'} = \emptyset$  if  $\lambda \neq \lambda'$ , then  $\mathcal{A}$  is at most countable.

Let  $\Lambda$  be the set of indices of  $\mathcal{A}$ , i.e.,  $\mathcal{A} = \{A_\lambda\}_{\lambda \in \Lambda}$ . Choose an arbitrary  $\lambda \in \Lambda$ , and choose an arbitrary point  $x \in A_\lambda$ . Since  $A_\lambda$  is open,  $\exists \varepsilon > 0$  s.t.  $B_\varepsilon(x) \subseteq A_\lambda$ . Because  $\mathbb{Q}^n$  is dense in  $\mathbb{R}^n$ ,  $\exists q_\lambda \in \mathbb{Q}^n$  s.t.  $q_\lambda \in B_\varepsilon(x) \subseteq A_\lambda$ . Therefore,  $\forall \lambda \in \Lambda$ ,  $\exists q_\lambda \in \mathbb{Q}^n$  s.t.  $q_\lambda \in B_\varepsilon(x) \subseteq A_\lambda$ . Therefore,  $\forall \lambda \in \Lambda$ ,  $\exists q_\lambda \in \mathbb{Q}^n$  s.t.  $q_\lambda \in A_\lambda$ . Note that  $\mathcal{A}$  is a collection of pairwise disjoint open sets, then  $\forall \lambda, \lambda' \in \Lambda$ ,  $\lambda \neq \lambda'$ ,  $q_\lambda \neq q'_{\lambda'}$ . Therefore, we can consider a function  $f : \mathcal{A} \rightarrow \mathbb{Q}^n$ , with  $f(A_\lambda) = q_\lambda$ ,  $\forall \lambda \in \Lambda$ . Since  $\forall \lambda, \lambda' \in \Lambda$ ,  $\lambda \neq \lambda'$ ,  $q_\lambda \neq q'_{\lambda'}$ , then  $f$  is an injection. This indicates the cardinality of  $\mathbb{Q}^n$  is weakly larger than  $\mathcal{A}$ , then  $\mathcal{A}$  is at most countable.

### 2.3.14 Compact Set

#### Definition 2.16 (Compact Set)

$\mathbb{X} \subseteq \mathbb{R}^n$  is compact if it is closed and bounded.



**Example 2.9 Compact Set**  $[1, 2] = \{x \in \mathbb{R} : 1 \leq x \leq 2\}; \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 4\}$

### 2.3.15 Sublevel Set

#### Definition 2.17 (Sublevel Set)

The sublevel set of a function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  (for some level  $c \in \mathbb{R}$ ) is the set

$$\overline{L_c} = \{x \in \mathbb{R}^n : f(x) \leq c\}$$



### 2.3.16 Set Operations

#### Definition 2.18

A binary operation on a set  $A$  is a function  $* : A \times A \rightarrow A$ .

The operation is *associative* if  $a * (b * c) = (a * b) * c, \forall a, b, c \in A$ .

The operation is *commutative* if  $a * b = b * a, \forall a, b \in A$ .



### Example 2.10

$+, \circ$  are both *associative* and *commutative* operations on  $\mathbb{Z}, \mathbb{N}, \mathbb{Q}, \mathbb{R}$ ;  $-$  is neither *associative* nor *commutative* operation on  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ , but not  $\mathbb{N}$ .

#### Definition 2.19

A subset  $H \subset S$  is closed under  $*$  if  $a * b \in H$  for all  $a, b \in H$ .



#### Definition 2.20

$*$  has identity element  $e \in S$  if  $a * e = e * a = a$  for all  $a \in S$ .



### 2.3.17 Connected Sets (@ Lec 07 of ECON 204)

#### Definition 2.21 (Separated/Connected Set)

Two sets  $A, B$  in a metric space are **separated** if

$$\bar{A} \cap B = A \cap \bar{B} = \emptyset$$

where  $\bar{A}$  is the closure of  $A$ .

A set in a metric space is **connected** if it cannot be written as the union of two nonempty separated sets.



**Equivalent Definition:** A set  $Y$  in a metric space  $X$  is connected if there do not exist *open sets*  $A$  and  $B$  such that  $A \cap B = \emptyset$ ,  $Y \subseteq A \cup B$  and  $A \cap Y \neq \emptyset$  and  $B \cap Y \neq \emptyset$ .

**Note** that separation does not require that  $\bar{A} \cap \bar{B} = \emptyset$ . For example,  $[0, 1) \cup (1, 2]$  is not connected, since  $[0, 1)$  and  $(1, 2]$  are separated.

### 2.3.18 Theorem: Connected Set over $\mathbb{E}^1 \Leftrightarrow$ Interval (@ Lec 07 of ECON 204)

#### Theorem 2.17 (Connected Set over $\mathbb{E}^1 \Leftrightarrow$ Interval)

A set  $S \subseteq \mathbb{E}^1$  of real numbers is **connected** if and only if it is an **interval**, i.e. if  $x, y \in S$  and  $z \in (x, y)$ , then  $z \in S$ .



Notice that this result is only valid in  $\mathbb{R}$ . For example, connected sets in  $\mathbb{R}^n$  need not be intervals, or even convex.

### 2.3.19 Theorem: Continuous Image of Connected Set is Connected

#### Theorem 2.18 (Continuous Image of Connected Set is Connected)

Let  $X$  be a metric space and  $f : X \rightarrow Y$  be continuous. If  $C$  is a connected subset of  $X$ , then  $f(C)$  is connected.



### 2.3.20 Intermediate Value Theorem proved by connectedness

#### Corollary 2.2 (Intermediate Value Theorem)

If  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, and  $f(a) < d < f(b)$ , then there exists  $c \in (a, b)$  such that  $f(c) = d$ .



#### Proof 2.3

Since  $[a, b]$  is an interval, it is connected. So  $f([a, b])$  is connected, hence  $f([a, b])$  is an interval.  $f(a) \in f([a, b])$ , and  $f(b) \in f([a, b])$ , and  $d \in [f(a), f(b)]$ ; since  $f([a, b])$  is an interval,  $d \in f([a, b])$ , i.e. there exists  $c \in [a, b]$  such that  $f(c) = d$ . Since  $f(a) < d < f(b)$ ,  $c \neq a$ ,  $c \neq b$ , so  $c \in (a, b)$ .

## 2.4 Sequences

Sequences  $\{x_k\}_{k=1}^{\infty}$ , ... or  $\{x_k\}$ ,  $x_k \in \mathbb{R}^n$

#### Definition 2.22 (Subsequence)

Suppose  $\{x_n\}$  is a sequence and  $n_1 < n_2 < \dots$ , then  $\{x_{n_k}\}$  is called a **subsequence**.



### 2.4.1 Convergence of Sequences (@ Lec 03 of ECON 204)

**Definition 2.23 (Convergence: note  $x_k \rightarrow x, \lim_{k \rightarrow \infty} x_k = x$ )**

Let  $(X, d)$  be a metric space. A sequence  $\{x_k\}$  converges to  $x$  (written  $x_k \rightarrow x$  or  $\lim_{k \rightarrow \infty} x_k = x$ ) if

$$\forall \varepsilon > 0, \exists N_\varepsilon \text{ s.t. } d(x_k, x) < \varepsilon, \forall k \geq N_\varepsilon$$



### 2.4.2 Limit Point (@ Lec 03 of ECON 204)

**Definition 2.24 (Limit Point)**

$x$  is a **limit point** of  $\{x_k\}$  if  $\exists$  a subsequence of  $\{x_k\}$  that converges to  $x$ .



**Definition 2.25 (Limit Point, alternative definition)**

Let  $(X, d)$  be a metric space. A point  $x \in X$  is a **limit point** of a set  $A \subseteq X$  if

$$\forall \varepsilon > 0, B_\varepsilon(x) \cap A \not\subseteq \{x\}$$



**Theorem 2.19 (Set of limit points of any set  $A \subseteq \mathbb{R}^n$  is closed.)**

The set of limit points of any set  $A \subseteq \mathbb{R}^n$  is closed.



**Proof 2.4**

Suppose  $\mathcal{S}$  is the set of limit points of  $A$ . Let  $D = \mathbb{R}^n \setminus \mathcal{S}$ . Take an arbitrary  $x \in D$ . Then  $x$  is not the limit point of  $A$ , so  $\exists \varepsilon_x > 0$  s.t.  $B_{\varepsilon_x}(x) \cap A \subseteq \{x\}$ . For any  $y \in B_{\varepsilon_x}(x)$ , since  $B_{\varepsilon_x}(x)$  is open,  $\exists \varepsilon_y > 0$  s.t.  $B_{\varepsilon_y}(y) \subseteq B_{\varepsilon_x}(x)$ . Let  $\varepsilon_m = \min\{\varepsilon_x, \varepsilon_y\}$  such that  $B_{\varepsilon_m}(y) \subseteq B_{\varepsilon_x}(x)$  and  $y \notin B_{\varepsilon_m}(y)$ . So,  $B_{\varepsilon_m}(y) \cap A = \emptyset$ . So,  $y$  is not a limit point of  $A$ , that is  $y \in \mathbb{R}^n \setminus \mathcal{S} = D$ . Therefore,  $\forall y \in B_{\varepsilon_x}(x) \Rightarrow y \in D$ . We can conclude  $B_{\varepsilon_x}(x) \subseteq D, \forall x \in D$ . So,  $D$  is open.

**Theorem 2.20 (Subsequence has the same limit as the original sequence)**

All subsequences of a convergent sequence converge to the same limit as the original sequence.



**Theorem 2.21 (Uniqueness of Limits)**

In a metric space  $(X, d)$ , if  $x_k \rightarrow x$  and  $x_k \rightarrow x'$ , then  $x = x'$ .



### 2.4.3 Cluster Point (@ Lec 03 of ECON 204)

**Definition 2.26 (Cluster Point)**

An element  $c$  is a **cluster point** of a sequence  $\{x_n\}$  in a metric space  $(X, d)$  if  $\forall \varepsilon > 0, \{n : x_n \in B_\varepsilon(c)\}$  is an infinite set. Equivalently,

$$\forall \varepsilon > 0, N \in \mathbb{N}, \exists n > N \text{ s.t. } x_n \in B_\varepsilon(c)$$



**Example 2.11**  $x_n = \begin{cases} 1 - \frac{1}{n}, & \text{if } n \text{ even} \\ \frac{1}{n}, & \text{if } n \text{ odd} \end{cases}$  has cluster points  $\{0, 1\}$ .

**Theorem 2.22 (Cluster Point  $\Leftrightarrow$  exists subsequence converges to it)**

Let  $(X, d)$  be a metric space,  $c \in X$ , and  $\{x_n\}$  a sequence in  $X$ . Then  $c$  is a cluster point of  $\{x_n\}$  if and only if there is a subsequence  $\{x_{n_k}\}$  such that  $\lim_{k \rightarrow \infty} x_{n_k} = c$ .



#### 2.4.4 Proposition: Sequences in $\mathbb{R}^n$ , Bounded above and non-decreasing sequence $\Rightarrow$ Converge

**Proposition 2.8**

If  $\{x_k\}$  is bounded above(below) and non-decreasing(non-increasing) it converges.



#### 2.4.5 Theorem: Increasing/Decreasing Sequences in $\mathbb{R}^n$ , Limit is sup/inf (@ Lec 03 of ECON 204)

**Theorem 2.23**

Let  $\{x_n\}$  be an increasing (decreasing) sequence of real numbers. Then  $\lim_{n \rightarrow \infty} x_n = \sup\{x_n : n \in \mathbb{N}\}$  ( $\lim_{n \rightarrow \infty} x_n = \inf\{x_n : n \in \mathbb{N}\}$ ). In particular, the limit exists.



#### 2.4.6 Definition: $\limsup$ and $\liminf$ (@ Lec 03 of ECON 204)

Consider a sequence  $\{x_n\}$  of real numbers. Let

$$\alpha_n = \sup\{x_k : k \geq n\}$$

and

$$\beta_n = \inf\{x_k : k \geq n\}$$

Either  $\alpha_n = +\infty$  for all  $n$ , or  $\alpha_n \in \mathbb{R}$  and  $\alpha_1 \geq \alpha_2 \geq \alpha_3 \geq \dots$ . Either  $\beta_n = -\infty$  for all  $n$ , or  $\beta_n \in \mathbb{R}$  and  $\beta_1 \leq \beta_2 \leq \beta_3 \leq \dots$

**Definition 2.27 (Lim Sups and Lim Insets)**

$$\limsup_{n \rightarrow \infty} x_n = \begin{cases} +\infty & \text{if } \alpha_n = +\infty \text{ for all } n \\ \lim \alpha_n & \text{otherwise.} \end{cases}$$

$$\liminf_{n \rightarrow \infty} x_n = \begin{cases} -\infty & \text{if } \beta_n = -\infty \text{ for all } n \\ \lim \beta_n & \text{otherwise.} \end{cases}$$



### 2.4.7 Exists Limit $\Leftrightarrow \lim = \limsup = \liminf$ (@ Lec 03 of ECON 204)

#### Theorem 2.24

Let  $\{x_n\}$  be a sequence of real numbers. Then

$$\lim_{n \rightarrow \infty} x_n = \gamma \in \mathbb{R} \cup \{-\infty, \infty\}$$

$$\Leftrightarrow \limsup_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n = \gamma$$



### 2.4.8 Rising Sun Lemma: Sequence in $\mathbb{R}^n$ contains monotone subsequence (@ Lec 03 of ECON 204)

#### Theorem 2.25 (Rising Sun Lemma)

Every sequence of real numbers contains an increasing subsequence or a decreasing subsequence or both.



### 2.4.9 Bolzano-Weierstrass: Bounded Sequence in $\mathbb{R}^n$ contains a convergent subsequence (@ Lec 03 of ECON 204)

#### Theorem 2.26 (Bolzano-Weierstrass)

Every bounded sequence of real numbers contains a convergent subsequence.



## 2.5 Complete Metric Spaces (@ Lec 05 of ECON 204)

Roughly, a metric space is complete if “every sequence that ought to converge to a limit has a limit to converge to.”

### 2.5.1 Cauchy Sequence

#### Definition 2.28 (Cauchy Sequence)

A sequence  $\{x_k\}$  in a metric space  $(X, d)$  is **Cauchy** if  $\forall \varepsilon > 0, \exists N(\varepsilon)$  s.t.

$$d(x_n, x_m) < \varepsilon, \forall n, m > N(\varepsilon)$$



**Note:** Cauchy property depends only on the sequence and the metric  $d$ , not on the ambient metric space.

**Note:** In  $\mathbb{E}^1 = (\mathbb{R}, \|\cdot\|)$ ,  $\{x_k\}$  converges  $\Leftrightarrow \{x_k\}$  is Cauchy.

A Cauchy sequence is trying really hard to converge, but there may not be anything for it to converge to. Any sequence that does converge must be Cauchy, however, by the argument above.

### 2.5.2 Theorem: Convergent $\Rightarrow$ Cauchy

**Theorem 2.27 (Convergent  $\Rightarrow$  Cauchy)**

Every convergent sequence in a metric space is Cauchy.



### 2.5.3 Complete Metric Space and Banach Space

**Definition 2.29 (Complete Metric Spaces)**

A metric space  $(X, d)$  is **complete** if every Cauchy sequence  $\{x_n\} \subseteq X$  converges to a limit  $x \in X$ .



**Example 2.12**

1. Consider the earlier example of  $X = (0, 1]$  with  $d$  the usual Euclidean metric. Since  $x_n = \frac{1}{n}$  is Cauchy but does not converge in that metric space,  $((0, 1], d)$  is not complete.
2.  $\mathbb{Q}$  is not complete in the Euclidean metric:  $x_n = \frac{\lfloor 10^n \sqrt{2} \rfloor}{10^n} \rightarrow \sqrt{2}$ .

**Definition 2.30 (Banach Space)**

A **Banach space** is a normed space that is complete in the metric generated by its norm.



**Theorem 2.28 ( $\mathbb{E}^1$  is complete)**

$\mathbb{R}$  is complete with the usual metric (so  $\mathbb{E}^1$  is a Banach space).



**Theorem 2.29 ( $\mathbb{E}^n$  is complete)**

$\mathbb{E}^n$  is complete for every  $n \in \mathbb{N}$ .



**Theorem 2.30 ( $(C(X), \|\cdot\|_\infty)$  is Complete)**

Given  $X \subseteq \mathbb{R}^n$ , let  $C(X)$  be the set of bounded continuous functions from  $X$  to  $\mathbb{R}$  with

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}$$

Then  $C(X)$  is a Banach space.



### 2.5.4 Theorem: Subset of Complete Metric Space is Complete $\Leftrightarrow$ Subset is Closed

**Theorem 2.31 (Subset of Complete Metric Space is Complete  $\Leftrightarrow$  Subset is Closed)**

Suppose  $(X, d)$  is a complete metric space and  $Y \subseteq X$ . Then  $(Y, d|_Y)$  is complete if and only if  $Y$  is a closed subset of  $X$ . (where  $d : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$ ;  $d|_Y : \mathbb{Y} \times \mathbb{Y} \rightarrow \mathbb{R}_+$ )



## 2.6 Compactness over Metric Spaces (@ Lec 06 of ECON 204)

### 2.6.1 Open Cover

#### Definition 2.31 (Open Cover)

A collection of sets  $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$  in a metric space  $(X, d)$  is an **open cover** of  $A \subseteq X$  if

1.  $U_\lambda$  is open for all  $\lambda \in \Lambda$  and
2.  $A \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda$



### 2.6.2 Definition of Compact Set

#### Definition 2.32 (Compact Set)

A set  $A$  in a metric space is **compact** if every open cover of  $A$  contains a finite subcover of  $A$ . In other words, if  $\{U_\lambda : \lambda \in \Lambda\}$  is an open cover of  $A$ , there exist  $n \in N$  and  $\lambda_1, \dots, \lambda_n \in \Lambda$  such that

$$A \subseteq \bigcup_{i=1}^n U_{\lambda_i}$$



#### Example 2.13

1.  $(0, 1]$  is not a compact set over  $\mathbb{E}^1$ : Consider the open cover  $\{(\frac{1}{m}, 2) : m \in \mathbb{N}\}$ .
2.  $[0, \infty)$  is not a compact set over  $\mathbb{E}^1$ : Consider the open cover  $\{(-1, m) : m \in \mathbb{N}\}$ .

### 2.6.3 Theorem: Closed Subset of Compact Set is Compact

#### Theorem 2.32 (Closed Subset of Compact Set is Compact)

*Every closed subset  $A$  of a compact metric space  $(X, d)$  is compact.*



#### Proof 2.5

Let  $\{U_\lambda : \lambda \in \Lambda\}$  be an open cover of  $A$ .

Let  $U'_\lambda = U_\lambda \cup (X \setminus A)$ . Since  $A$  is closed,  $X \setminus A$  is open; Since  $U_\lambda$  and  $X \setminus A$  are both open,  $U'_\lambda$  is open.

By the definition of open set, if  $a \in A$ ,  $a \in \bigcup_{\lambda \in \Lambda} U'_\lambda$ . If  $a \in X \setminus A$ ,  $a \in U'_\lambda$  for sure. So,  $X \subseteq \bigcup_{\lambda \in \Lambda} U'_\lambda$ .

Hence, we can conclude  $\{U'_\lambda : \lambda \in \Lambda\}$  is an open cover of  $X$ .

Since  $X$  is compact, there exists  $n \in N$  and  $\lambda_1, \dots, \lambda_n \in \Lambda$  such that  $X \subseteq \bigcup_{i=1}^n U'_{\lambda_i}$ .

Then,

$$a \in A \Rightarrow a \in X$$

$$\Rightarrow a \in U'_{\lambda_i} \text{ for some } i$$

$$\Rightarrow a \in U_{\lambda_i} \text{ for some } i$$

Hence, we can conclude  $A \subseteq \bigcup_{i=1}^n U_{\lambda_i}$ . Thus,  $A$  is compact.

### 2.6.4 Theorem: Compact Set over Metric Space is Closed

**Theorem 2.33 (Compact Set over Metric Space is Closed)**

If  $A$  is a compact subset of the metric space  $(X, d)$ , then  $A$  is closed.



**Proof 2.6**

Suppose  $X \setminus A$  is not open, so we can find a point  $x \in X \setminus A$  such that, for every  $\varepsilon > 0$ ,  $A \cap B_\varepsilon(x) \neq \emptyset$ , and hence  $A \cap B_\varepsilon[x] \neq \emptyset$ . For  $n \in \mathbb{N}$ , let

$$U_n = X \setminus B_{1/n}[x]$$

Each  $U_n$  is open, and

$$\cup_{n \in \mathbb{N}} U_n = X \setminus \{x\} \supseteq A$$

since  $x \notin A$ . Therefore,  $\{U_n : n \in \mathbb{N}\}$  is an open cover for  $A$ . Since  $A$  is compact, there is a finite subcover  $\{U_{n_1}, \dots, U_{n_k}\}$ . Let  $n = \max\{n_1, \dots, n_k\}$ . Then

$$\begin{aligned} U_n &= X \setminus B_{1/n}[x] \\ &\supseteq X \setminus B_{1/n_j}[x] \quad (j = 1, \dots, k) \\ U_n &\supseteq \cup_{j=1}^k U_{n_j} \\ &\supseteq A \end{aligned}$$

But  $A \cap B_{1/n}[x] \neq \emptyset$ , so  $A \not\subseteq X \setminus B_{1/n}[x] = U_n$ . This is a contradiction, which proves that  $A$  is closed.



### 2.6.5 Definition of Sequentially Compact

**Definition 2.33 (Sequentially Compact)**

A set  $A$  in a metric space  $(X, d)$  is **sequentially compact** if every sequence of elements of  $A$  contains a convergent subsequence whose limit lies in  $A$ .



### 2.6.6 Theorem: Compact $\Leftrightarrow$ Sequentially Compact

**Theorem 2.34 (Compact  $\Leftrightarrow$  Sequentially Compact)**

A set  $A$  in a metric space  $(X, d)$  is **compact** if and only if it is **sequentially compact**.



### 2.6.7 Definition of Totally Bounded

#### Definition 2.34 (Totally Bounded)

A set  $A$  in a metric space  $(X, d)$  is **totally bounded** if, for every  $\varepsilon > 0$ ,

$$\exists x_1, \dots, x_n \in A \text{ s.t. } A \subseteq \bigcup_{i=1}^n B_\varepsilon(x_i).$$



#### Claim 2.2

"Totally bounded" is stronger than "bounded".



#### Example 2.14

1.  $X = [0, 1]$  is totally bounded with Euclidean metric: Given  $\varepsilon > 0$ , let  $x_i = \frac{i}{n}, i = 1, \dots, n - 1$ .

2.  $X = [0, 1]$  is not totally bounded with discrete metric  $d(x, y) = \begin{cases} 1, & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}$ : Given  $\varepsilon = \frac{1}{2}$ ,

$B_\varepsilon(x_i) = \{x_i\}$ . Hence,  $\bigcup_i^n B_\varepsilon(x_i) = \{x_1, \dots, x_n\} \not\supseteq [0, 1]$ .

**Note:**  $X$  is bounded but not totally bounded.

### 2.6.8 Theroem: Compact $\Leftrightarrow$ Complete/Closed and Totally Bounded

#### Theorem 2.35 (Subset is Compact $\Leftrightarrow$ Complete and Totally Bounded)

Let  $A$  be a subset of a metric space  $(X, d)$ . Then  $A$  is **compact** if and only if  $A$  is **complete** and **totally bounded**.



#### Proof 2.7

$\text{Compact} \Rightarrow \text{totally bounded}$  (by definition).  $\text{Compact} \Leftrightarrow \text{sequentially compact} \Rightarrow \text{every Cauchy sequence converges to a limit in the subset (i.e., Complete)}$ .

Conversely, suppose complete and totally bounded. Because totally bounded, we can extract a Cauchy subsequence. Because complete, subsequence converges to a limit in the subset, which shows sequentially compact and hence compact.

#### Corollary 2.3 (Subset is Compact $\Leftrightarrow$ Closed and Totally Bounded)

Let  $A$  be a subset of a metric space  $(X, d)$ . Then  $A$  is **compact** if and only if  $A$  is **closed** and **totally bounded**.



#### Proof 2.8

Directly by theorem 2.31 and theorem 2.35.

### 2.6.9 Heine-Borel Theorem: With Euclidean Metric, Compact $\Leftrightarrow$ Closed and Bounded

**Claim 2.3 (With Euclidean Metric, Compact  $\Leftrightarrow$  Closed and Bounded)**

If  $A \subseteq \mathbb{E}^n$ , then  $A$  is **totally bounded** if and only if  $A$  is **bounded**.



**Theorem 2.36 (Heine-Borel Theorem: With Euclidean Metric, Compact  $\Leftrightarrow$  Closed and Bounded)**

If  $A \subseteq \mathbb{E}^n$ , then  $A$  is **compact** if and only if  $A$  is **closed and bounded**.



### 2.6.10 Theorem: Continuous Images of Compact Set is Compact

**Theorem 2.37 (Continuous Images of Compact Set is Compact)**

Let  $(X, d)$  and  $(Y, \rho)$  be metric spaces. If  $f : X \rightarrow Y$  is continuous and  $C$  is a compact subset of  $(X, d)$ , then  $f(C)$  is compact in  $(Y, \rho)$ .



# Chapter 3 Functions

## 3.1 Definitions of Function

### Definition 3.1 (Function)

*Function* is a rule  $\sigma : A \rightarrow B$  that assigns an element  $B$  to *every* element of  $A$ .  $\forall a \in A, \sigma(a) \in B$ .



### 3.1.1 Image, Preimage, Fiber

#### Definition 3.2

1.  $A$  is the domain of  $\sigma$ ,  $B$  is the range of  $\sigma$ .
2. We call  $\sigma(a) = \text{value of } \sigma \text{ at } a$  as the image of  $a$ .
3. A set  $C \subset B$ , we call  $\sigma^{-1}(C) = \{a \in A | \sigma(a) \in C\}$  as the preimage of  $C$ .
4. An element  $b \in B$ , we call  $\sigma^{-1}(b) = \{a \in A | \sigma(a) = b\}$  as the fiber of  $b$ .



### 3.1.2 Function Image's Properties

#### Theorem 3.1

Let  $X, Y$  be sets,  $f : X \rightarrow Y$  be a function,  $A, B \subseteq X$  be arbitrary subsets of  $X$ , and  $C, D \subseteq Y$  be arbitrary subsets of  $Y$ . Then

1.  $f(A \cap B) \subseteq f(A) \cap f(B)$ . (*Strict subset example:*  $y = x^2$ .  $A = (-2, 0)$  and  $B = (0, 2)$ .)
2.  $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
3.  $f(A \cup B) = f(A) \cup f(B)$
4.  $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
5.  $A \subseteq f^{-1}(f(A))$  (*Strict subset example:*  $y = x^2$ ,  $A = [0, 2]$ . Then  $f(A) = [0, 4]$ ,  $f^{-1}(f(A)) = [-2, 2]$ .)
6.  $f(f^{-1}(C)) \subseteq C$  (*Strict subset example:*  $X = [0, 1], Y = [0, 2], f(x) = x$ . Let  $C = [0.5, 1.5]$ . Then  $f^{-1}(C) = [0.5, 1]$ ,  $f(f^{-1}(C)) = [0.5, 1]$ .)



### 3.1.3 Composition of functions

#### Definition 3.3 (Function Composition)

The function composition  $\circ$  is an operation that takes two functions  $\sigma : A \rightarrow B$  and  $\tau : B \rightarrow C$ , , and produces a function  $\tau \circ \sigma : A \rightarrow C$  that fulfills  $\forall a \in A, (\tau \circ \sigma)(a) = \tau(\sigma(a))$ .



### 3.1.4 Function Composition is Associative

#### Proposition 3.1 (Associativity of Functions)

Suppose  $\sigma : A \rightarrow B, \tau : B \rightarrow C, \rho : C \rightarrow D$  are functions and  $\circ$  is the function composition, then

$$\rho \circ (\tau \circ \sigma) = (\rho \circ \tau) \circ \sigma.$$



### 3.1.5 Homeomorphism (@ Lec 04 of ECON 204)

#### Definition 3.4 (Homeomorphism)

Let  $(X, d)$  and  $(Y, \rho)$  be metric spaces. A function  $f : X \rightarrow Y$  is called a **homeomorphism** if it is one-to-one, onto, continuous, and its inverse function is continuous.



## 3.2 Injection, Surjection, Bijection

### 3.2.1 Definitions: Injective, surjective, bijective

A function  $\sigma : A \rightarrow B$  is called,

1. *Injective (1 to 1)*

$$\forall a_1, a_2 \in A, \sigma(a_1) = \sigma(a_2) \Rightarrow a_1 = a_2$$

2. *Surjective (onto)*

$$\forall b \in B, \exists a \in A, s.t. \sigma(a) = b$$

3. *Bijective* (if injective and surjective)

### 3.2.2 Lemma 1.1.7: injective/surjective/bijective is preserved in composition

#### Lemma 3.1 (Lemma 1.1.7)

Suppose  $\sigma : A \rightarrow B, \tau : B \rightarrow C$  are functions,

If  $\sigma, \tau$  are injective, then  $\tau \circ \sigma$  is injective.

If  $\sigma, \tau$  are surjective, then  $\tau \circ \sigma$  is surjective.

If  $\sigma, \tau$  are bijective, then  $\tau \circ \sigma$  is bijective.



### 3.2.3 Theorem: A function is bijective $\Leftrightarrow \exists$ inverse

#### Theorem 3.2 (Bijective $\Leftrightarrow \exists$ Inverse)

A function  $\sigma : A \rightarrow B$  is a bijection if and only if  $\exists$  a function  $\tau : B \rightarrow A$  such that

$$\sigma \circ \tau = id_B = \text{identity on } B(id_B(x) = x, \forall x \in B)$$

$$\tau \circ \sigma = id_A$$

Such  $\tau$  is unique, called inverse of  $\sigma$ ,  $\tau = \sigma^{-1}$ .



## 3.3 Function Continuity

### 3.3.1 Continuous Function in $\mathbb{R}$ with Euclidean norm

#### Definition 3.5 (Continuity at Point)

A real-valued function  $f$  is continuous at  $x$  if

"For every  $\{x_k\}$  converging to  $x$  satisfies that  $\lim_{k \rightarrow \infty} f(x_k) = f(x)$ ".

Equivalent definition:

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. } \|y - x\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$



Continuity at  $x_0$  requires:

1.  $f(x_0)$  is defined; and
2. either
  - $x_0$  is an isolated point of  $X$ , i.e.  $\exists \varepsilon > 0$  s.t.  $B_\varepsilon(x) = \{x\}$ ; or
  - $\lim_{x \rightarrow x_0} f(x)$  exists and equals  $f(x_0)$

#### Definition 3.6 (Continuous Function)

A real-valued function  $f$  is continuous if it is continuous at all points in its domain.



### 3.3.2 Continuous Function in Metric Spaces (@ Lec 04 of ECON 204)

#### Definition 3.7 (Continuity in Metric Spaces)

Let  $(X, d)$  and  $(Y, \rho)$  be metric spaces. A function  $f : X \rightarrow Y$  is **continuous** at a point  $x_0 \in X$  if

$$\forall \varepsilon > 0, \exists \delta(x_0, \varepsilon) > 0 \text{ s.t. } d(x, x_0) < \delta(x_0, \varepsilon) \Rightarrow \rho(f(x), f(x_0)) < \varepsilon.$$



### 3.3.3 Theorem: Continuous $\Leftrightarrow$ Preimage of open set is open (@ Lec 04 of ECON 204)

**Theorem 3.3 (Continuous  $\Leftrightarrow$  Preimage of open set is open)**

Let  $(X, d)$  and  $(Y, \rho)$  be metric spaces, and  $f : X \rightarrow Y$ . Then  $f$  is continuous if and only if

$$f^{-1}(A) = \{x \in X : f(x) \in A\} \text{ is open in } X \quad \forall A \subseteq Y \text{ s.t. } A \text{ is open in } Y$$



**Proposition 3.2**

Let  $A \subseteq \mathbb{R}$  be an open set and  $f : A \rightarrow \mathbb{R}$ . The two following statements are equivalent:

(a).  $f$  is continuous.

(b). for all  $c \in \mathbb{R}$ , the sets  $E[f < c] = \{x \in A \mid f(x) < c\}$  and  $E[f > c] = \{x \in A \mid f(x) > c\}$  are open.



**Proof 3.1**

1. " $(a) \Rightarrow (b)$ ": Suppose  $f$  is continuous. Consider any  $c \in \mathbb{R}$ . Since  $(-\infty, c)$  and  $(c, \infty)$  are open and  $f$  is continuous,  $E[f < c] = f^{-1}((-\infty, c))$  and  $E[f > c] = f^{-1}((c, \infty))$  are open.

2. " $(b) \Rightarrow (a)$ ": Suppose  $E[f < c] = \{x \in A \mid f(x) < c\}$  and  $E[f > c] = \{x \in A \mid f(x) > c\}$  are open for all  $c \in \mathbb{R}$ . Let  $C \subseteq \mathbb{R}$  be an arbitrary open set over  $\mathbb{R}$ . There are following cases about the  $C$ :

(1).  $C = (-\infty, k)$ , where  $k \in \mathbb{R}$ . In this case  $f^{-1}(C) = E[f < k]$ , which is open by the given condition.

(2).  $C = (k, \infty)$ , where  $k \in \mathbb{R}$ . In this case  $f^{-1}(C) = E[f > k]$ , which is open by the given condition.

(3).  $C = \emptyset$ . In this case  $f^{-1}(C) = \{x \in A \mid f(x) \in \emptyset\} = E[f > k_1] \cap E[f < k_2], \forall k_1 \geq k_2$ , which is a finite intersection of two open sets. So,  $f^{-1}(C)$  is open.

(4).  $C = (k_1, k_2)$ , where  $k_1, k_2 \in \mathbb{R}$  s.t.  $k_1 < k_2$ . In this case  $f^{-1}(C) = \{x \in A \mid f(x) \in (k_1, k_2)\} = E[f > k_1] \cap E[f < k_2]$ , which is a finite intersection of two open sets. So,  $f^{-1}(C)$  is open.

(5).  $C = (-\infty, k_2) \cup (k_1, \infty)$ . In this case  $f^{-1}(C) = \{x \in A \mid f(x) \in (-\infty, k_2) \cup (k_1, \infty)\} = E[f > k_1] \cup E[f < k_2], \forall k_1, k_2 \in \mathbb{R}$  s.t.  $k_1 > k_2$ .  $f^{-1}(C)$  is a union of two open set, which is open.

(6).  $C = \mathbb{R}$ . In this case  $f^{-1}(C) = \{x \in A \mid f(x) \in \mathbb{R}\} = \{x \in A \mid f(x) \in (-\infty, k_2) \cup (k_1, \infty)\} = E[f > k_1] \cup E[f < k_2], \forall k_1, k_2 \in \mathbb{R}$  s.t.  $k_1 < k_2$ .  $f^{-1}(C)$  is a union of two open set, which is open.

Hence,  $f^{-1}(C) = \{x \in A : f(x) \in C\}$  is open  $\forall C \subseteq \mathbb{R}$  s.t.  $C$  is open in  $\mathbb{R}$

### 3.3.4 Theorem: Continuity is preserved in composition (@ Lec 04 of ECON 204)

#### Theorem 3.4 (Continuity is preserved in composition)

Let  $(X, d_X)$ ,  $(Y, d_Y)$  and  $(Z, d_Z)$  be metric spaces. If  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are continuous, then  $g \circ f : X \rightarrow Z$  is continuous.



#### Proof 3.2

Proved by previous theorem 3.3.

### 3.3.5 Uniform Continuity (@ Lec 04 of ECON 204)

#### Definition 3.8 (Uniformly Continuous)

Suppose  $f : (X, d) \rightarrow (Y, \rho)$ .  $f$  is **uniformly continuous** if

$$\forall \varepsilon > 0, \exists \delta(\varepsilon) > 0 \text{ s.t. } \forall x_0 \in X, d(x, x_0) < \delta(\varepsilon) \Rightarrow \rho(f(x), f(x_0)) < \varepsilon$$



#### Claim 3.1

**Uniformly Continuous** implies (is stronger than) **Continuous**.

( $f$  is continuous if  $\forall x_0 \in X, \varepsilon > 0, \exists \delta(x_0, \varepsilon) > 0$  s.t.  $d(x, x_0) < \delta(x_0, \varepsilon) \Rightarrow \rho(f(x), f(x_0)) < \varepsilon$ )



Given  $\varepsilon > 0$ , "uniformly continuous" requires  $\delta(\varepsilon)$  that works for all  $x_0 \in X$ .

### 3.3.6 Theorem: Continuous function with Compact domain is Uniformly Continuous (@ Lec 06 of ECON 204)

#### Theorem 3.5 (Continuous function with Compact domain is Uniformly Continuous)

Let  $(X, d)$  and  $(Y, \rho)$  be metric spaces,  $C$  a **compact** subset of  $X$ , and  $f : C \rightarrow Y$  a **continuous** function.

Then  $f$  is **uniformly continuous** on  $C$ .



#### Proof 3.3

Fix  $\varepsilon > 0$ . Since  $f$  is continuous, given  $c \in C$ , we can find  $\delta(c) > 0$  such that

$$\forall x \in C, d(x, c) < 2\delta(c) \Rightarrow \rho(f(x), f(c)) < \frac{\varepsilon}{2}$$

Let  $U_c = B_{\delta(c)}(c)$ , then  $\{U_c : c \in C\}$  is an open cover of  $C$ . Since  $C$  is compact, there is a finite subcover  $\{U_{c_1}, \dots, U_{c_n}\} = \{B_{\delta(c_1)}(c_1), \dots, B_{\delta(c_n)}(c_n)\}$ . Let  $\delta = \min\{\delta(c_1), \dots, \delta(c_n)\}$ .

Given  $x, y \in C$  with  $d(x, y) < \delta$ , note that  $x \in U_{c_i}$  for some  $i \in \{1, \dots, n\}$ , so  $d(x, c_i) < \delta(c_i)$ .

$$\begin{aligned} d(y, c_i) &\leq d(y, x) + d(x, c_i) \\ &< \delta + \delta(c_i) \\ &\leq \delta(c_i) + \delta(c_i) = 2\delta(c_i) \end{aligned}$$

So,

$$\begin{aligned} \rho(f(x), f(y)) &\leq \rho(f(x), f(c_i)) + \rho(f(c_i), f(y)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

That is,  $d(x, y) < \delta \Rightarrow \rho(f(x), f(y)) < \varepsilon$ , which proves that  $f$  is uniformly continuous.

### 3.3.7 Lipschitz Continuous (@ Lec 04 of ECON 204)

#### Definition 3.9 (Lipschitz (Continuous) in Normed Vector Space)

Let  $X, Y$  be normed vector spaces,  $\mathbb{E} \subseteq X$ .

(1). A function  $f : X \rightarrow Y$  is **Lipschitz** on  $\mathbb{E}$  satisfies

$$\exists \gamma > 0, \|f(\mathbf{x}) - f(\mathbf{y})\|_Y \leq \gamma \|\mathbf{x} - \mathbf{y}\|_X, \forall \mathbf{x}, \mathbf{y} \in \mathbb{E}$$

or we call  $\gamma$ -**Lipschitz continuous**;

(2).  $f$  is **locally Lipschitz** on  $\mathbb{E}$  if

$$\forall x_0 \in \mathbb{E}, \exists \varepsilon > 0 \text{ s.t. } f \text{ is Lipschitz on } B_\varepsilon(x_0) \cap \mathbb{E}$$



#### Claim 3.2

- 1). If  $f$  is  $\gamma$ -Lipschitz continuous, then it is also  $(\gamma + 1)$ -Lipschitz continuous. The minimal such  $\gamma$  is called a Lipschitz constant of function  $f$
- 2). The definition can be extended to general metric spaces by replacing  $\|\cdot\|_X$  and  $\|\cdot\|_Y$  to general metrics.



#### Example 3.1 (Lipschitz)

1.  $f(x) = 2x$  is 2-Lipschitz continuous;
2. What about  $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ , where  $\mathbf{A}$  is a matrix? Spectral norm  $\|\mathbf{A}\|_2$  (for Euclidean norm).
3. What about  $f(x) = x^2$ ? Not Lipschitz continuous, or the Lipschitz constant is  $\infty$ .

**Example 3.2 (Locally Lipschitz)** Every  $C^1$  function is locally Lipschitz. (Recall that a function  $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$  is said to be  $C^1$  if all its first partial derivatives exist and are continuous.)

**Claim 3.3 (Lipschitz continuity vs. continuity vs. uniform continuity)**

*Lipschitz continuity is stronger than either continuity or uniform continuity:*

- *locally Lipschitz  $\Rightarrow$  continuous*
- *Lipschitz  $\Rightarrow$  uniformly continuous*

**3.3.8 Contraction (from the view of Lipschitz constant)**

- a). If the Lipschitz constant  $\gamma \leq 1$ , then  $f$  is called a non-expansive mapping.
- b). If the Lipschitz constant  $\gamma < 1$ , then  $f$  is called a contraction mapping.

**Example 3.3**

1.  $f(x) = 2x$  is not a contraction mapping;  $f(x) = 0.5x$  is.
2.  $f(x) = Ax$  is a contraction mapping (with respect to Euclidean norm) iff  $\|A\|_2 < 1$ .

**3.4 Extreme Value, Intermediate Value of Functions****3.4.1 Intermediate Value Theorem (@ Lec 02 of ECON 204)****Theorem 3.6 (Intermediate Value Theorem)**

*Suppose  $f : [a, b] \rightarrow \mathbb{R}$  is continuous, and  $f(a) < d < f(b)$ . Then there exists  $c \in (a, b)$  such that  $f(c) = d$ .*

**3.4.2 Coercive Function****Definition 3.10 (Coercive)**

*A real-valued function  $f : \mathbb{X} \rightarrow \mathbb{R}$  is coercive if for every  $\{x_k\} \subset \mathbb{X}$  s.t.  $\|x_k\| \rightarrow \infty$ ,  $f(x_k) \rightarrow \infty$*

**Example 3.4 Check coercive**

- 1)  $x \in \mathbb{R}^2$ ,  $f(x) = x_1^2 + x_2^2$  - coercive
- 2)  $x \in \mathbb{R}$ ,  $f(x) = 1 - e^{-|x|}$  - not coercive
- 3)  $x \in \mathbb{R}^2$ ,  $f(x) = x_1^2 + x_2^2 - 2x_1x_2$  - not coercive (we need  $f(x_k) \rightarrow \infty$  for all  $\|x_k\| \rightarrow \infty$ )

### 3.4.3 Extreme of Functions

#### Definition 3.11 (Extreme of Functions)

Let  $\mathbb{X} \subseteq \mathbb{R}^n$ ,  $f : \mathbb{X} \rightarrow \mathbb{R}$

$$\inf_{x \in \mathbb{X}} f(x) = \inf\{f(x) : x \in \mathbb{X}\}$$



If  $\exists x^* \in \mathbb{X}$  s.t.  $\inf f(x) = f(x^*)$ . Then,  $f$  achieves (attains) its minimum and  $f(x^*) = \min_{x \in \mathbb{X}} f(x)$

$x^*$  is called a **minimizer** of  $f$ , written as  $x^* \in \arg \min_{x \in \mathbb{X}} f(x)$ . If  $x^*$  is unique, we write  $x^* = \arg \min_{x \in \mathbb{X}} f(x)$

Similarly, supremum and maximum of  $f$ .

### 3.4.4 Weierstrass' Theorem (Extreme value Theorem) (@ Lec 06 of ECON 204)

#### Theorem 3.7 (Weierstrass' Theorem (Extreme value Theorem))

Given a **compact** set  $C$  in a metric space  $(X, d)$  and a **continuous** function  $f : C \rightarrow \mathbb{R}$ , then  $f$  is bounded on  $C$  and attains its minimum and maximum on  $C$  i.e.,

$$\exists x_1 \in C \text{ s.t. } f(x_1) = \inf f(C)$$

$$\exists x_2 \in C \text{ s.t. } f(x_2) = \sup f(C)$$



#### Proof 3.4

Since  $C$  is compact and  $f$  is continuous,  $f(C) \subset \mathbb{R}$  is compact by theorem 2.37. Hence,  $f(C)$  is closed and bounded.

1. *Proof of maximum:* Let  $M = \sup f(C) < \infty$ . Then  $\forall m > 0$  there exists  $y_m \in f(C)$  such that

$$M - \frac{1}{m} \leq y_m \leq M$$

So  $y_m \rightarrow M$  and  $\{y_m\} \subseteq f(C)$ . Since  $f(C)$  is closed,  $M \in f(C)$ , i.e., there exists  $c \in C$  such that  $f(c) = M = \sup f(C)$ .

2. *Proof of minimum:* Let  $M = \inf f(C) > -\infty$ . Then  $\forall m > 0$  there exists  $y_m \in f(C)$  such that

$$M \geq y_m \geq M + \frac{1}{m}$$

So  $y_m \rightarrow M$  and  $\{y_m\} \subseteq f(C)$ . Since  $f(C)$  is closed,  $M \in f(C)$ , i.e., there exists  $c \in C$  such that  $f(c) = M = \inf f(C)$ .

### 3.4.5 Corollary: Continuous and Coercive function attains its Minimum on Closed Set

**Corollary 3.1 (Continuous and Coercive function attains its Minimum on Closed Set)**

Let  $f$  be continuous and coercive on closed set  $\mathbb{X}$  (not necessarily bounded), it attains its min on  $\mathbb{X}$ .



**Proof 3.5**

Consider  $\{\sigma_k\}$  as in proof of WT.

Since  $f$  is closed,  $f(x) < \infty, \forall x \in \mathbb{X}$ . And  $f$  is coercive on  $\mathbb{X}$ , which means  $f(x) \rightarrow \infty$  if  $\|x\| \rightarrow \infty$ .

Hence,  $\{\sigma_k\} \in \mathbb{X}$  is bounded. Rest of proof same as proof of WT.

**Example 3.5**  $f(x) = f(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + e^{-x_3} + e^{2x_3}$

1) Does  $f$  achieve its min and max on  $\mathbb{X}_1 = \{x \in \mathbb{R}^3 : x_1^2 + 2x_2^2 + 3x_3^2 \leq 6\}$ ?

-  $\mathbb{X}_1$  is compact and  $f$  is continuous. Both min and max are achieved (WT).

2) Does  $f$  achieve its min and max over  $\mathbb{R}^3$ ?

-  $f \rightarrow \infty$  whenever  $\|x\| \rightarrow \infty \Rightarrow f$  is coercive.

-  $\mathbb{R}^3$  is closed.

$\Rightarrow f$  achieves its min. on  $\mathbb{R}^3$  by corollary to WT.

- max does not exist since  $f \rightarrow \infty$  as  $\|x\| \rightarrow \infty$ .

3) Does  $f$  achieve its min and max over  $\mathbb{X}_2 = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 3\}$ ?

-  $\mathbb{X}_2$  is closed, but not bounded.

- Since  $f$  is coercive, min achieved.

- max does not exist since setting  $x_1 = 0, x_2 = 3 - x_3$  and letting  $x_3 \rightarrow \infty$  makes  $f \rightarrow \infty$

## 3.5 Monotonic Function (@ Lec 05 of ECON 204)

**Definition 3.12**

A function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is **monotonically increasing** if  $y > x \Rightarrow f(y) \geq f(x)$ .



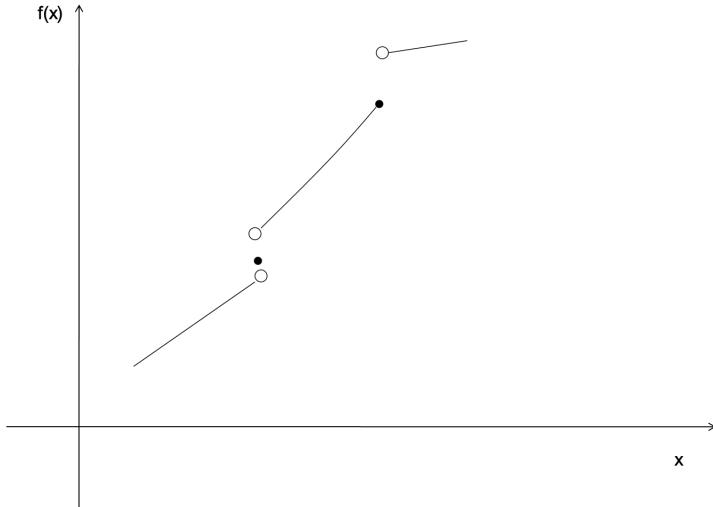
### 3.5.1 Theorem: Monotonically Increasing $\Rightarrow$ One-sided Limits exist

**Theorem 3.8**

Let  $a, b \in \mathbb{R}$  with  $a < b$ , and let  $f : (a, b) \rightarrow \mathbb{R}$  be monotonically increasing. Then the one-sided limits:

$f(t^+) = \lim_{u \rightarrow t^+} f(u)$  and  $f(t^-) = \lim_{u \rightarrow t^-} f(u)$  exist and are real numbers for all  $t \in (a, b)$ .





**Figure 3.1:** A monotonic function has only simple jump discontinuities.

### 3.5.2 Theorem: Monotonically Increasing $\Rightarrow$ Set of Points are discontinuous is finite (possibly empty) or countable

#### Theorem 3.9

Let  $a, b \in \mathbb{R}$  with  $a < b$ , and let  $f : (a, b) \rightarrow \mathbb{R}$  be monotonically increasing. Then  $D = \{t \in (a, b) : f \text{ is discontinuous at } t\}$  is finite (possibly empty) or countable.



## 3.6 Correspondence (@ Lec 07 of ECON 204)

### 3.6.1 Definition of Correspondence

#### Definition 3.13 (Correspondence)

A **correspondence**  $\Psi : X \rightarrow 2^Y$  from  $X$  to  $Y$  is a function from  $X$  to  $2^Y$ , that is,  $\Psi(x) \subseteq Y$  for every  $x \in X$ . ( $2^Y$  is the set of all subsets of  $Y$ )



**Example 3.6** Let  $u : \mathbb{R}_+^n \rightarrow \mathbb{R}$  be a continuous utility function,  $y > 0$  and  $p \in \mathbb{R}_{++}^n$ , that is,  $p_i > 0$  for each  $i$ . Define  $\Psi : \mathbb{R}_{++}^n \times \mathbb{R}_{++} \rightarrow 2^{\mathbb{R}_+^n}$  by

$$\Psi(p, y) = \operatorname{argmax} u(x)$$

$$\text{s.t. } x \geq 0$$

$$p \cdot x \leq y$$

$\Psi$  is the demand correspondence associated with the utility function  $u$ ; typically  $\Psi(p, y)$  is multi-valued.

### 3.6.2 Continuity of Correspondences

Let  $X \subseteq \mathbb{E}^n$ ,  $Y \subseteq \mathbb{E}^m$ , and  $\Psi : X \rightarrow 2^Y$ .

#### Definition 3.14 (Upper Hemicontinuous)

$\Psi$  is **upper hemicontinuous** (uhc) at  $x_0 \in X$  if, for every open set  $V$  with  $\Psi(x_0) \subseteq V$ , there is an open set  $U$  with  $x_0 \in U$  s.t.

$$x \in U \Rightarrow \Psi(x) \subseteq V$$



Upper hemicontinuity reflects the requirement that  $\Psi$  doesn't "jump down/implode in the limit" at  $x_0$ . (*A set to "jump down" at the limit  $x_0$ : It should mean the set suddenly gets smaller – it "implodes in the limit" – that is, there is a sequence  $x_n \rightarrow x_0$  and points  $y_n \in \Psi(x_n)$  that are far from every point of  $\Psi(x_0)$  as  $n \rightarrow \infty$ .*)

#### Definition 3.15 (Lower Hemicontinuous)

$\Psi$  is **lower hemicontinuous** (lhc) at  $x_0 \in X$  if, for every open set  $V$  with  $\Psi(x_0) \cap V \neq \emptyset$ , there is an open set  $U$  with  $x_0 \in U$  s.t.

$$x \in U \Rightarrow \Psi(x) \cap V \neq \emptyset$$



Lower hemicontinuity reflects the requirement that  $\Psi$  doesn't "jump up/explode in the limit" at  $x_0$ . (*A set to "jump up" at the limit  $x_0$ : It should mean that the set suddenly gets bigger – it "explodes in the limit" – that is, there is a sequence  $x_n \rightarrow x_0$  and a point  $y_0 \in \Psi(x_0)$  that is far from every point of  $\Psi(x_n)$  as  $n \rightarrow \infty$ .*)

#### Definition 3.16 (Continuous Correspondence)

$\Psi$  is **continuous** at  $x_0 \in X$  if it is both **uhc** and **lhc** at  $x_0$ .



#### Proposition 3.3

$\Psi$  is upper hemicontinuous (respectively lower hemicontinuous, continuous) if it is uhc (respectively lhc, continuous) at every  $x \in X$ .



### 3.6.3 Graph of Correspondence

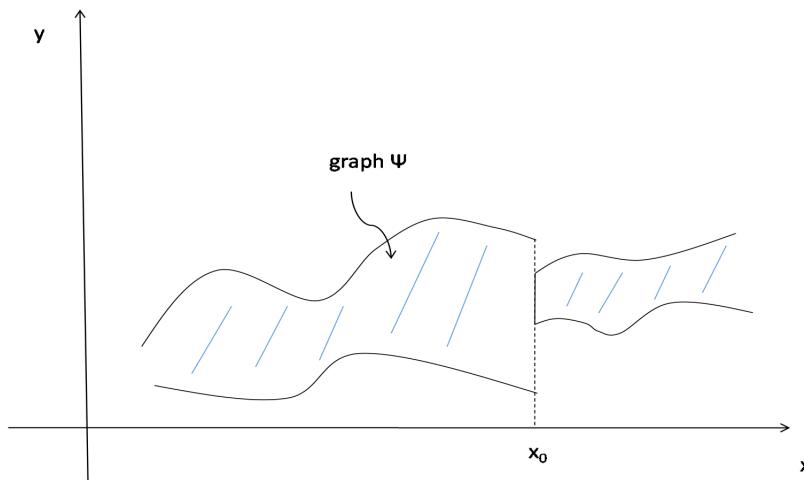
An alternative notion of continuity looks instead at properties of the graph of the correspondence.

#### Definition 3.17 (Graph of Correspondence)

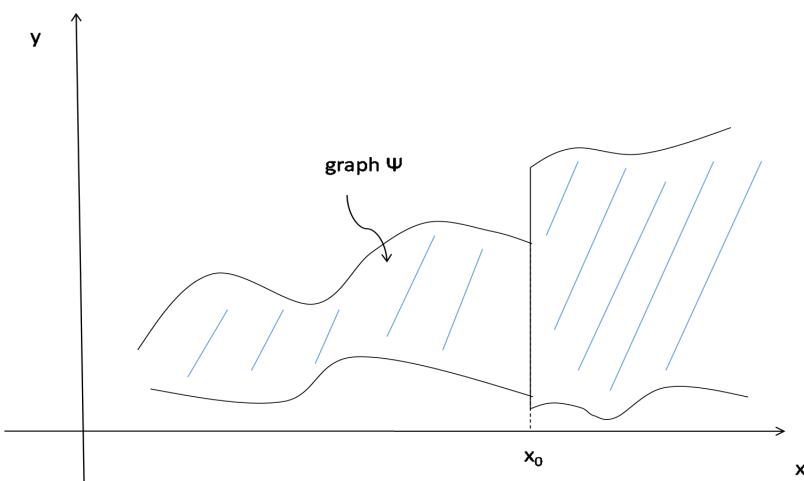
The **graph** of a correspondence  $\Psi : X \rightarrow 2^Y$  is the set

$$\text{graph } \Psi = \{(x, y) \in X \times Y : y \in \Psi(x)\}$$





**Figure 3.2:** The correspondence  $\Psi$  “implodes in the limit” at  $x_0$ .  $\Psi$  is not upper hemicontinuous at  $x_0$ .



**Figure 3.3:** The correspondence  $\Psi$  “explodes in the limit” at  $x_0$ .  $\Psi$  is not lower hemicontinuous at  $x_0$ .

By the definition of continuous function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , each convergent sequence  $\{(x_n, y_n)\}$  in graph  $f$  converges to a point  $(x, y)$  in graph  $f$ , that is, graph  $f$  is closed.

### Definition 3.18 (Closed Graph)

Let  $X \subseteq \mathbb{E}^n$ ,  $Y \subseteq \mathbb{E}^m$ . A correspondence  $\Psi : X \rightarrow 2^Y$  has closed graph if its graph is a closed subset of  $X \times Y$ , that is, if for any sequences  $\{x_n\} \subseteq X$  and  $\{y_n\} \subseteq Y$  such that  $x_n \rightarrow x \in X$ ,  $y_n \rightarrow y \in Y$  and  $y_n \in \Psi(x_n)$  for each  $n$ , then  $y \in \Psi(x)$ .



**Example 3.7** Consider the correspondence  $\Psi(x) = \begin{cases} \{\frac{1}{x}\}, & \text{if } x \in (0, 1] \\ \{0\}, & \text{if } x = 0 \end{cases}$  ("implode in the limit")

Let  $V = (-0.1, 0.1)$ . Then  $\Psi(0) = \{0\} \subseteq V$ , but no matter how close  $x$  is to 0,  $\Psi(x) = \{\frac{1}{x}\} \not\subseteq V$ , so  $\Psi$  is not uhc at 0. However, note that  $\Psi$  has closed graph.

### 3.6.4 Theorem: $\Psi(x) = \{f(x)\}$ is uhc $\Leftrightarrow f$ is continuous

#### Theorem 3.10 ( $\Psi(x) = \{f(x)\}$ is uhc $\Leftrightarrow f$ is continuous)

Let  $X \subseteq \mathbb{E}^n$ ,  $Y \subseteq \mathbb{E}^m$  and  $f : X \rightarrow Y$ . Let  $\Psi : X \rightarrow 2^Y$  be defined by  $\Psi(x) = \{f(x)\}$  for all  $x \in X$ .

Then  $\Psi$  is uhc if and only if  $f$  is continuous.



### 3.6.5 Closed-valued and Compact-valued Correspondences

#### Definition 3.19 (Closed-valued and Compact-valued Correspondences)

A correspondence  $\Psi : X \rightarrow 2^Y$  is called **closed-valued** if  $\Psi(x)$  is a closed subset of  $Y$  for all  $x$ ;  $\Psi$  is called **compact-valued** if  $\Psi(x)$  is compact for all  $x$ .



For closed-valued correspondences these concepts can be more tightly connected. A closed-valued and upper hemicontinuous correspondence must have closed graph. For a closed-valued correspondence with a compact range, upper hemicontinuity is equivalent to closed graph.

#### Theorem 3.11 (Properties)

Let  $X \subseteq \mathbb{E}^n$ ,  $Y \subseteq \mathbb{E}^m$ , and  $\Psi : X \rightarrow 2^Y$ .

- (i).  $\Psi$  is closed-valued and uhc  $\Rightarrow \Psi$  has closed graph.
- (ii).  $\Psi$  is closed-valued and uhc  $\Leftrightarrow \Psi$  has closed graph. (If  $Y$  is compact)
- (iii). If  $\Psi$  has closed graph and there is an open set  $W$  with  $x_0 \in W$  and a compact set  $Z$  such that  $x \in W \cap X \Rightarrow \Psi(x) \subseteq Z$ , then  $\Psi$  is uhc at  $x_0$ .



## 3.7 Differentiable Function (@ Lec 11 of ECON 204)

### 3.7.1 Definition of Differentiable in $\mathbb{R}$

#### Definition 3.20

Let  $f : I \rightarrow \mathbb{R}$ , where  $I \subseteq \mathbb{R}$  is an open interval.  $f$  is **differentiable** at  $x \in I$  if

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = a$$

for some  $a \in \mathbb{R}$ .



This is equivalent to  $\exists a \in \mathbb{R}$  such that:

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{f(x+h) - (f(x) + ah)}{h} = 0 \\ \Leftrightarrow & \forall \varepsilon > 0 \exists \delta > 0 \text{ s.t. } 0 < |h| < \delta \Rightarrow \left| \frac{f(x+h) - (f(x) + ah)}{h} \right| < \varepsilon \\ \Leftrightarrow & \forall \varepsilon > 0 \exists \delta > 0 \text{ s.t. } 0 < |h| < \delta \Rightarrow \frac{|f(x+h) - (f(x) + ah)|}{|h|} < \varepsilon \\ \Leftrightarrow & \lim_{h \rightarrow 0} \frac{|f(x+h) - (f(x) + ah)|}{|h|} = 0 \end{aligned}$$

Recall that the limit considers  $h$  near zero, but not  $h = 0$ .

#### Proposition 3.4

$f$  is **differentiable**  $\Rightarrow f$  is **continuous**.



### 3.7.2 General Definition of Differentiable

#### Definition 3.21

If  $X \subseteq \mathbb{R}^n$  is open,  $f : X \rightarrow \mathbb{R}^m$  is **differentiable** at  $x \in X$  if  $\exists T_x \in L(\mathbb{R}^n, \mathbb{R}^m)$  s.t.

$$\lim_{h \rightarrow 0, h \in \mathbb{R}^n} \frac{|f(x+h) - (f(x) + T_x(h))|}{|h|} = 0$$

$f$  is differentiable if it is differentiable at all  $x \in X$ .



Note that  $T_x$  is uniquely determined by Equation (1).  $h$  is a small, nonzero element of  $\mathbb{R}^n$ ;  $h \rightarrow 0$  from any direction, from above, below, along a spiral, etc. The definition requires that one linear operator  $T_x$  works no matter how  $h$  approaches zero. In this case,  $f(x) + T_x(h)$  is the best linear approximation to  $f(x+h)$  for small  $h$ .

### 3.7.3 "O" and "o" Notations

#### Definition 3.22 ("O" and "o")

1.  $y = O(|h|^n)$  as  $h \rightarrow 0$  – read “ $y$  is big-Oh of  $|h|^n$ ” – means  $\exists K, \delta > 0$  s.t.  $|h| < \delta \Rightarrow |y| \leq K|h|^n$
2.  $y = o(|h|^n)$  as  $h \rightarrow 0$  – read “ $y$  is little-Oh of  $|h|^n$ ” – means  $\lim_{h \rightarrow 0} \frac{|y|}{|h|^n} = 0$

 **Note** Note that the statement  $y = O(|h|^{n+1})$  as  $h \rightarrow 0$  implies  $y = o(|h|^n)$  as  $h \rightarrow 0$ .

In both cases,  $y \rightarrow 0$  as  $h \rightarrow 0$ , but the *rate* is different.

#### Definition 3.23 (New Definition)

Using this notation, note that  $f$  is **differentiable** at  $x \Leftrightarrow$

$$\exists T_x \in L(\mathbb{R}^n, \mathbb{R}^m) \text{ such that } f(x + h) = f(x) + T_x(h) + o(|h|) \text{ as } h \rightarrow 0$$

### 3.7.4 Jacobian and Error Term

#### Definition 3.24 (Jacobian, Error Term)

- $df_x$  is the linear transformation  $T_x$
- $Df(x)$  is the matrix of  $df_x$  with respect to the standard basis. This is called the **Jacobian** or Jacobian matrix of  $f$  at  $x$
- $E_f(h) = f(x + h) - (f(x) + df_x(h))$  is the **error term**

Using this notation,  $f$  is differentiable at  $x \Leftrightarrow E_f(h) = o(h)$  as  $h \rightarrow 0$

Now compute  $Df(x) = (a_{ij})$ . Let  $\{e_1, \dots, e_n\}$  be the standard basis of  $\mathbb{R}^n$ . Look in direction  $e_j$  (note that  $|\gamma e_j| = |\gamma|$ ).

$$o(\gamma) = f(x + \gamma e_j) - (f(x) + T_x(\gamma e_j))$$

$$\begin{aligned} &= f(x + \gamma e_j) - \left( f(x) + \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \gamma \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right) \\ &= f(x + \gamma e_j) - \left( f(x) + \begin{pmatrix} \gamma a_{1j} \\ \vdots \\ \gamma a_{mj} \end{pmatrix} \right) \end{aligned}$$

For  $i = 1, \dots, m$ , let  $f^i$  denote the  $i^{\text{th}}$  component of the function  $f$ :

$$f^i(x + \gamma e_j) - (f^i(x) + \gamma a_{ij}) = o(\gamma)$$

$$\text{so, } a_{ij} = \frac{\partial f^i}{\partial x_j}(x)$$

### 3.7.5 Theorem: Differentiable $\Rightarrow$ Partial Derivatives (and Jacobian) exist

#### Theorem 3.12 (Differentiable $\Rightarrow$ Partial Derivatives (and Jacobian) exist)

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $f : X \rightarrow \mathbb{R}^m$  is differentiable at  $x \in X$ . Then  $\frac{\partial f^i}{\partial x_j}$  exists at  $x$  for

$1 \leq i \leq m, 1 \leq j \leq n$ , and

$$Df(x) = \begin{pmatrix} \frac{\partial f^1}{\partial x_1}(x) & \cdots & \frac{\partial f^1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^m}{\partial x_1}(x) & \cdots & \frac{\partial f^m}{\partial x_n}(x) \end{pmatrix}$$

i.e. the Jacobian at  $x$  is the matrix of partial derivatives at  $x$ .



**Note** If  $f$  is differentiable at  $x$ , then all first-order partial derivatives  $\frac{\partial f^i}{\partial x_j}$  exist at  $x$ . However, the converse is false: existence of all the first-order partial derivatives does not imply that  $f$  is differentiable.

**Example 3.8**  $f(X) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2}, & X \neq (0, 0) \\ 0, & X = (0, 0) \end{cases}$

### 3.7.6 Theorem: Partial Derivatives exist and Continuous $\Rightarrow$ Differentiable

The missing piece is continuity of the partial derivatives:

#### Theorem 3.13 (Partial Derivatives exist and Continuous $\Rightarrow$ Differentiable)

If all the first-order partial derivatives  $\frac{\partial f^i}{\partial x_j}$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ) exist and are continuous at  $x$ , then  $f$  is differentiable at  $x$ .



### 3.7.7 Directional Derivatives

Suppose  $X \subseteq \mathbb{R}^n$  open,  $f : X \rightarrow \mathbb{R}^m$  is differentiable at  $x$ , and  $|u| = 1$ .

$$f(x + \gamma u) - (f(x) + T_x(\gamma u)) = o(\gamma) \text{ as } \gamma \rightarrow 0$$

$$\Rightarrow f(x + \gamma u) - (f(x) + \gamma T_x(u)) = o(\gamma) \text{ as } \gamma \rightarrow 0$$

$$\Rightarrow \lim_{\gamma \rightarrow 0} \frac{f(x + \gamma u) - f(x)}{\gamma} = T_x(u) = Df(x)u$$

i.e. the **directional derivative** in the direction  $u$  (with  $|u| = 1$ ) is

$$Df(x)u \in \mathbb{R}^m$$

**3.7.8 Chain Rule:**  $dF_{x_0} = dg_{f(x_0)} \circ df_{x_0}$  and  $DF_{x_0} = Dg(f(x_0))Df(x_0)$ **Theorem 3.14 (Chain Rule)**

Let  $X \subseteq \mathbb{R}^n$ ,  $Y \subseteq \mathbb{R}^m$  be open,  $f : X \rightarrow Y$ ,  $g : Y \rightarrow \mathbb{R}^p$ . Let  $x_0 \in X$  and  $F = g \circ f$ . If  $f$  is differentiable at  $x_0$  and  $g$  is differentiable at  $f(x_0)$ , then  $F = g \circ f$  is differentiable at  $x_0$  and

$$dF_{x_0} = dg_{f(x_0)} \circ df_{x_0} \text{ (composition of linear transformations)}$$

$$DF_{x_0} = Dg(f(x_0))Df(x_0) \text{ (matrix multiplication)}$$

**3.7.9 Mean Value Theorem:**  $\exists z \in l(x, y), \|f(y) - f(x)\| \leq \|df_z\| \|y - x\|$ **Theorem 3.15 (Mean Value Theorem, Univariate Case:  $\mathbb{R} \rightarrow \mathbb{R}$ )**

Let  $a, b \in \mathbb{R}$ . Suppose  $f : [a, b] \rightarrow \mathbb{R}$  is continuous on  $[a, b]$  and differentiable on  $(a, b)$ . Then there exists  $c \in (a, b)$  such that

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

that is

$$f(b) - f(a) = f'(c)(b - a)$$



**Note** The Mean Value Theorem is useful for estimating bounds on functions and error terms in approximation of functions.

**Definition 3.25 (Line Segment)**

$l(x, y) = \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\}$  is the **line segment** from  $x$  to  $y$ .

**Theorem 3.16 (Mean Value Theorem,  $\mathbb{R}^n \rightarrow \mathbb{R}$ )**

Suppose  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable on an open set  $X \subseteq \mathbb{R}^n$ ,  $x, y \in X$  and  $l(x, y) \subseteq X$ . Then there exists  $z \in l(x, y)$  such that

$$f(y) - f(x) = Df(z)(y - x)$$

**Theorem 3.17 (Mean Value Theorem, General Case:  $\mathbb{R}^n \rightarrow \mathbb{R}^m$ )**

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $f : X \rightarrow \mathbb{R}^m$  is differentiable. If  $x, y \in X$  and  $l(x, y) \subseteq X$ , then there exists  $z \in l(x, y)$  such that

$$\|f(y) - f(x)\| \leq \|df_z(y - x)\|$$

$$\leq \|df_z\| \|y - x\|$$



### 3.7.10 Taylor's Theorem

#### Theorem 3.18 (Taylor's Theorem)

Let  $f : I \rightarrow \mathbb{R}$  be  $n$ -times differentiable, where  $I \subseteq \mathbb{R}$  is an open interval. If  $x, x + h \in I$ , then

$$f(x + h) = f(x) + \sum_{k=1}^{n-1} \frac{f^{(k)}(x)h^k}{k!} + E_n$$

where  $f^{(k)}$  is the  $k^{\text{th}}$  derivative of  $f$  and

$$E_n = \frac{f^{(n)}(x + \lambda h)h^n}{n!} \text{ for some } \lambda \in (0, 1)$$



#### Theorem 3.19 (Taylor's Theorem in $\mathbb{R}^1$ )

Let  $f : I \rightarrow \mathbb{R}$  be  $n$  times differentiable, where  $I \subseteq \mathbb{R}$  is an open interval and  $x \in I$ . Then,

$$f(x + h) = f(x) + \sum_{k=1}^n \frac{f^{(k)}(x)h^k}{k!} + o(h^n) \text{ as } h \rightarrow 0$$

If  $f$  is  $(n + 1)$  times continuously differentiable (i.e. all derivatives up to order  $n + 1$  exist and are continuous), then

$$f(x + h) = f(x) + \sum_{k=1}^n \frac{f^{(k)}(x)h^k}{k!} + o(h^{n+1}) \text{ as } h \rightarrow 0$$



### 3.7.11 $C^k$ Function and Continuously Differentiable

#### Definition 3.26 ( $C^k$ Function)

$f$  is  $C^k$  if all partial derivatives of order less than or equal to  $k$  exist and are continuous in  $X$ .



#### Definition 3.27 (Continuously Differentiable)

Let  $X \subseteq \mathbb{R}^n$  be open. A function  $f : X \rightarrow \mathbb{R}^m$  is **continuously differentiable** on  $X$  if

1.  $f$  is differentiable on  $X$  and
2.  $df_x$  is a continuous function of  $x$  from  $X$  to  $L(\mathbb{R}^n, \mathbb{R}^m)$ , with operator norm  $\|df_x\|$ .



#### Theorem 3.20 (Continuously Differentiable $\Leftrightarrow C^1$ )

Suppose  $X \subseteq \mathbb{R}^n$  is open and function  $f : X \rightarrow \mathbb{R}^m$  is **continuously differentiable** if and only if  $f$  is  $C^1$ .



### 3.7.12 Taylor's Theorem with $C^k$

## Linear Terms

### Theorem 3.21

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $x \in X$ . If  $f : X \rightarrow \mathbb{R}^m$  is differentiable, then

$$f(x + h) = f(x) + Df(x)h + o(h) \text{ as } h \rightarrow 0$$



The previous theorem is essentially a restatement of the definition of differentiability.

### Theorem 3.22

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $x \in X$ . If  $f : X \rightarrow \mathbb{R}^m$  is  $C^2$ , then

$$f(x + h) = f(x) + Df(x)h + O(|h|^2) \text{ as } h \rightarrow 0$$



## Quadratic Terms

We treat each component of the function separately, so consider  $f : X \rightarrow \mathbb{R}$ ,  $X \subseteq \mathbb{R}^n$  an open set. Let

$$D^2f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \frac{\partial^2 f}{\partial x_2^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \cdots & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{pmatrix}$$

$$f \in C^2 \Rightarrow \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$$

$\Rightarrow D^2f(x)$  is symmetric

$\Rightarrow D^2f(x)$  has an orthonormal basis of eigenvectors and thus can be diagonalized

### Theorem 3.23 (Stronger Version of Taylor's Theorem)

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $x \in X$ . If  $f : X \rightarrow \mathbb{R}$  is  $C^2$ , then

$$f(x + h) = f(x) + Df(x)h + \frac{1}{2}h^T(D^2f(x))h + O(|h|^2) \text{ as } h \rightarrow 0$$

If  $f \in C^3$ ,

$$f(x + h) = f(x) + Df(x)h + \frac{1}{2}h^T(D^2f(x))h + O(|h|^3) \text{ as } h \rightarrow 0$$



### Definition 3.28

We say  $f$  has a **saddle** at  $x$  if  $Df(x) = 0$  but  $f$  has neither a local maximum nor a local minimum at  $x$ .



**Corollary 3.2**

Suppose  $X \subseteq \mathbb{R}^n$  is open and  $x \in X$ . If  $f : X \rightarrow \mathbb{R}$  is  $C^2$ , then there is an orthonormal basis  $\{v_1, \dots, v_n\}$  and corresponding eigenvalues  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  of  $D^2f(x)$  such that

$$\begin{aligned} f(x + h) &= f(x + \gamma_1 v_1 + \dots + \gamma_n v_n) \\ &= f(x) + \sum_{i=1}^n (Df(x)v_i) \gamma_i + \frac{1}{2} \sum_{i=1}^n \lambda_i \gamma_i^2 + o(|\gamma|^2) \end{aligned}$$

where  $\gamma_i = h \cdot v_i$ .

1. If  $f \in C^3$ , we may strengthen  $o(|\gamma|^2)$  to  $O(|\gamma|^3)$ .
2. If  $f$  has a local maximum or local minimum at  $x$ , then

$$Df(x) = 0$$

3. If  $Df(x) = 0$ , then

$$\lambda_1, \dots, \lambda_n > 0 \Rightarrow f \text{ has a local minimum at } x$$

$$\lambda_1, \dots, \lambda_n < 0 \Rightarrow f \text{ has a local maximum at } x$$

$$\lambda_i < 0 \text{ for some } i, \lambda_j > 0 \text{ for some } j \Rightarrow f \text{ has a saddle at } x$$

$$\lambda_1, \dots, \lambda_n \geq 0, \lambda_i > 0 \text{ for some } i \Rightarrow f \text{ has a local minimum or a saddle at } x$$

$$\lambda_1, \dots, \lambda_n \leq 0, \lambda_i < 0 \text{ for some } i \Rightarrow f \text{ has a local maximum or a saddle at } x$$

$$\lambda_1 = \dots = \lambda_n = 0 \quad \text{gives no information.}$$



# Chapter 4 Tools for Comparative Statics

Consider the function  $f : (0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}$  s.t.

$$f(x, a) = \sin x + a$$

Let  $X = (0, 2\pi)$  and let  $f_a(x) = f(x, a) = \sin x + a$  denote the perturbed function for fixed  $a$ .

## 4.1 Regular and Critical Points and Values

### 4.1.1 Rank of Derivatives $\text{Rank } df_x = \text{Rank } Df(x)$

Suppose  $X \subseteq \mathbb{R}^n$  is open. Suppose  $f : X \rightarrow \mathbb{R}^m$  is differentiable at  $x \in X$ , and let  $W = \{e_1, \dots, e_n\}$  denote the standard basis of  $\mathbb{R}^n$ . Then  $df_x \in L(\mathbb{R}^n, \mathbb{R}^m)$ , and

$$\begin{aligned}\text{Rank } df_x &= \dim \text{Im}(df_x) \\ &= \dim \text{span}\{df_x(e_1), \dots, df_x(e_n)\} \\ &= \dim \text{span}\{Df(x)e_1, \dots, Df(x)e_n\} \\ &= \dim \text{span}\{\text{column 1 of } Df(x), \dots, \text{column n of } Df(x)\} \\ &= \text{Rank } Df(x)\end{aligned}$$

Thus,

$$\text{Rank } df_x \leq \min\{m, n\}$$

$df_x$  has **full rank** if  $\text{Rank } df_x = \min\{m, n\}$ , that is, is  $df_x$  has the maximum possible rank.

### 4.1.2 Regular and Critical Points and Values

#### Definition 4.1 (Regular and Critical Points and Values)

Suppose  $X \subseteq \mathbb{R}^n$  is open. Suppose  $f : X \rightarrow \mathbb{R}^m$  is differentiable at  $x \in X$ .

1.  $x$  is a **regular point** of  $f$  if  $\text{Rank } df_x = \min\{m, n\}$ .
2.  $x$  is a **critical point** of  $f$  if  $\text{Rank } df_x < \min\{m, n\}$ .
3.  $y$  is a **critical value** of  $f$  if there exists  $x \in f^{-1}(y)$  such that  $x$  is a critical point of  $f$ .
4.  $y$  is a **regular value** of  $f$  if  $y$  is not a critical value of  $f$ .



 **Note** Notice that if  $y \notin f(X)$ , so  $f^{-1}(y) = \emptyset$ , then  $y$  is automatically a regular value of  $f$ .

**Example 4.1** Suppose  $f(x, y) = (\sin x, \cos y)$ ,  $Df(x, y) = \begin{bmatrix} \cos x & 0 \\ 0 & -\sin y \end{bmatrix}$ . Critical point:  $\{(\frac{k\pi}{2}, \mathbb{R}) : k \in 2\mathbb{Z} + 1\} \cup \{(\mathbb{R}, k\pi) : k \in \mathbb{Z}\}$ ; Critical values:  $\{(x, y) : x = 1 \text{ or } x = -1 \text{ or } y = 1 \text{ or } y = -1\}$

## 4.2 Inverse and Implicit Function Theorem

### 4.2.1 Inverse Function Theorem

Using Taylor's theorem to approximate

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + o(x - x_0)$$

The requirement of "regular point" is necessary for the  $Df(x_0)$  being invertible.

#### Theorem 4.1 (Inverse Function Theorem)

Suppose  $X \subseteq \mathbb{R}^n$  is open. Suppose  $f : X \rightarrow \mathbb{R}^n$  is  $C^1$  on  $X$ , and  $x_0 \in X$ . If  $\det Df(x_0) \neq 0$  (i.e.,  $x_0$  is a regular point of  $f$ ), then there are open neighborhoods  $U$  of  $x_0$  and  $V$  of  $f(x_0)$  s.t.

$f : U \rightarrow V$  is bijective (on-to-one and onto)

$\exists f^{-1} : V \rightarrow U$  is  $C^1$

$$Df^{-1}(f(x_0)) = [Df(x_0)]^{-1}$$

$$(\text{In } \mathbb{R}, (f^{-1})'(f(x_0)) = (f'(x_0))^{-1})$$

If in addition  $f \in C^k$ , then  $f^{-1} \in C^k$ .



### 4.2.2 Implicit Function Theorem

Using Taylor's theorem to approximate

$$f(x, a) = f(x_0, a_0) + Df(x_0, a_0)(x - x_0) + Df(x_0, a_0)(a - a_0) + \text{remainder}$$

The requirement of "regular point" is necessary for the  $Df(x_0, a_0)$  being invertible.

We want to know how the function  $x^*(a)$  changes with keeping  $f(x^*, a) = 0$ .

#### Theorem 4.2 (Implicit Function Theorem)

Suppose  $X \subseteq \mathbb{R}^n$  and  $A \subseteq \mathbb{R}^p$  are open and  $f : X \times A \rightarrow \mathbb{R}^n$  is  $C^1$ . Suppose  $f(x_0, a_0) = 0$  and  $\det(D_x f(x_0, a_0)) \neq 0$ , i.e.  $x_0$  is a regular point of  $f(\cdot, a_0)$ . Then there are open neighborhoods  $U$  of  $x_0$  ( $U \subseteq X$ ) and  $W$  of  $a_0$  such that

$$\forall a \in W, \exists! x \in U \text{ s.t. } f(x, a) = 0$$

For each  $a \in W$  let  $g(a)$  be that unique  $x$ . Then  $g : W \rightarrow U$  is  $C^1$  and

$$Dg(a_0) = -[D_x f(x_0, a_0)]^{-1} [D_a f(x_0, a_0)]$$

If in addition  $f \in C^k$ , then  $g \in C^k$ .



### 4.2.3 Prove Implicit Function Theorem Given Inverse Function Theorem

#### Proof 4.1

1. Firstly, we prove "g is differentiable": The "change of a" incurs the value change:

$$\begin{aligned} f(x_0, a_0 + h) &= f(x_0, a_0) + D_a f(x_0, a_0)h + o(h) \\ &= D_a f(x_0, a_0)h + o(h) \end{aligned}$$

Find a  $\Delta x$  such that the new  $x$  can let the value go back to 0, i.e.,  $f(x_0 + \Delta x, a_0 + h) = 0$ . That is,

$$g(a_0 + h) = x_0 + \Delta x$$

To prove "g is differentiable", we want to prove " $\exists T \in L(A, X)$  s.t.  $\Delta x = T(h) + o(h)$ "

$$\begin{aligned} 0 &= f(x_0 + \Delta x, a_0 + h) \\ &= f(x_0, a_0) + D_x f(x_0, a_0 + h)\Delta x + D_a f(x_0, a_0)h + o(\Delta x) + o(h) \\ &= D_x f(x_0, a_0 + h)\Delta x + D_a f(x_0, a_0)h + o(\Delta x) + o(h) \end{aligned}$$

$$D_x f(x_0, a_0 + h)\Delta x = -D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

Because  $f$  is  $C^1$  and the determinant is a continuous function of the entries of the matrix,  $\det D_x f(x_0, a_0 + h) \neq 0$  for  $h$  sufficiently small, so

$$\Delta x = -[D_x f(x_0, a_0 + h)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$$

Since  $f \in C^1$ ,  $\Delta x = -[D_x f(x_0, a_0) + o(1)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$

Since  $f \in C^1$ ,  $\Delta x = -[D_x f(x_0, a_0)]^{-1} D_a f(x_0, a_0)h + o(\Delta x) + o(h)$

Hence, "g is differentiable" is proved and the derivative of g is  $Dg(a_0) = -[D_x f(x_0, a_0)]^{-1} [D_a f(x_0, a_0)]$ .

2. Secondly, given the "g is differentiable", we can also compute the derivative by

$$Df(g(a), a)(a_0) = 0$$

$$D_x f(x_0, a_0)Dg(a_0) + D_a f(x_0, a_0) = 0$$

$$Dg(a_0) = -[D_x f(x_0, a_0)]^{-1} D_a f(x_0, a_0)$$

**Example 4.2**  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ ,  $f((3, -1, 2)) = (0, 0)$ ,  $Df(3, -1, 2) = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ . Then, let  $(x_0, a_0) = (3, -1, 2)$ , where  $x_0 = 3$  and  $a_0 = (-1, 2)$ . Or, we can let  $(x_0, a_0) = (3, -1, 2)$ , where  $x_0 = (3, -1)$  and  $a_0 = 2$ .

#### 4.2.4 Prove Inverse Function Theorem Given Implicit Function Theorem

##### Proof 4.2 (Prove Inverse Function Theorem Given Implicit Function Theorem)

Define  $F : X \times \mathbb{R}^n$  s.t.  $F(x, y) = y - f(x)$ . Let  $y_0 = f(x_0)$ .

$$D_x F(x, y) = -Df(x), D_y F(x, y) = I_{n \times n}$$

According to the implicit function theorem, there are open sets  $U \subseteq X$  and  $V \subseteq \mathbb{R}^n$  such that  $x_0 \in U$ ,  $y_0 \in V$  and a function  $g : V \rightarrow U$  differentiable at  $y_0$  such that  $F(g(y), y) = 0$  for all  $y \in V$ . So,  $0 = F(g(y), y) = y - f(g(y))$ , we have  $f(g(y)) = y$ , that is  $g = f^{-1}$ .  $f : U \rightarrow V$  is bijective because it has inverse  $g : V \rightarrow U$ .

By the implicit function theorem,  $g(y)$  is differentiable and

$$Df^{-1}(y_0) = Dg(y_0) = -[D_x F(x_0, y_0)]^{-1}[D_y F(x_0, y_0)] = [Df(x_0)]^{-1}$$

where  $y_0 = f(x_0)$ .

By the implicit function theorem, the  $g = f^{-1}$  is  $C^k$  if  $f$  is  $C^k$ .

All in all, the inverse function theorem is proved.

#### 4.2.5 Example: Using Implicit Function Theorem in Comparative Statics

**Example 4.3** Let us consider a firm that produces a good  $y$ ; it uses two inputs  $x_1$  and  $x_2$ . The firm sells the output and acquires the inputs in competitive markets: The market price of  $y$  is  $p$ , and the cost of each unit of  $x_1$  and  $x_2$  are  $w_1$  and  $w_2$  respectively. Its technology is given by  $f : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ , where  $f(x_1, x_2) = x_1^a x_2^b$ ,  $a + b < 1$ . Its profits take the form

$$\pi(x_1, x_2; p, w_1, w_2) = px_1^a x_2^b - w_1 x_1 - w_2 x_2$$

The firm selects  $x_1$  and  $x_2$  in order to maximize profits. **We aim to know how its choice of  $x_1$  and  $x_2$  is affected by a change in  $w_1$ .**

Assuming an interior solution, the first-order conditions of this optimization problem are

$$\begin{aligned} \frac{\partial \pi}{\partial x_1}(x_1^*, x_2^*; p, w_1, w_2) &= pa(x_1^*)^{a-1}(x_2^*)^b - w_1 = 0 \\ \frac{\partial \pi}{\partial x_2}(x_1^*, x_2^*; p, w_1, w_2) &= pb(x_1^*)^a(x_2^*)^{b-1} - w_2 = 0 \end{aligned}$$

for some  $(x_1, x_2) = (x_1^*, x_2^*)$ .

Let us define

$$F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} pa(x_1^*)^{a-1}(x_2^*)^b - w_1 \\ pb(x_1^*)^a(x_2^*)^{b-1} - w_2 \end{bmatrix}$$

Jacobian matrices are

$$D_{(x_1, x_2)} F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} pa(a-1)(x_1^*)^{a-2}(x_2^*)^b & pab(x_1^*)^{a-1}(x_2^*)^{b-1} \\ pab(x_1^*)^{a-1}(x_2^*)^{b-1} & pb(b-1)(x_1^*)^a(x_2^*)^{b-2} \end{bmatrix}$$

$$D_{w_1} F(x_1^*, x_2^*; p, w_1, w_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

By the implicit function theorem, we can get

$$\begin{bmatrix} \frac{\partial x_1^*}{\partial w_1} \\ \frac{\partial x_2^*}{\partial w_1} \end{bmatrix} = -[D_{(x_1, x_2)} F(x_1^*, x_2^*; p, w_1, w_2)]^{-1} [D_{w_1} F(x_1^*, x_2^*; p, w_1, w_2)]$$

$$= [D_{(x_1, x_2)} F(x_1^*, x_2^*; p, w_1, w_2)]^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

#### 4.2.6 Corollary: $a \rightarrow \{x \in X : f(x, a) = 0\}$ is lhc

##### Corollary 4.1

Suppose  $X \subseteq \mathbb{R}^n$  and  $A \subseteq \mathbb{R}^p$  are open and  $f : X \times A \rightarrow \mathbb{R}^n$  is  $C^1$ . If 0 is a regular value of  $f(\cdot, a_0)$ , then the correspondence

$$a \rightarrow \{x \in X : f(x, a) = 0\}$$

is **lower hemicontinuous** at  $a_0$ .



## 4.3 Transversality and Genericity

### 4.3.1 Lebesgue Measure Zero

#### Definition 4.2 (Lebesgue Measure Zero)

Suppose  $A \subseteq \mathbb{R}^n$ .  $A$  has **Lebesgue measure zero** if for every  $\varepsilon > 0$  there is a countable collection of rectangles  $I_1, I_2, \dots$  such that

$$\sum_{k=1}^{\infty} \text{Vol}(I_k) < \varepsilon \text{ and } A \subseteq \bigcup_{k=1}^{\infty} I_k$$

Here by a rectangle we mean  $I_k = \times_{j=1}^n (a_j^k, b_j^k) = \{x \in \mathbb{R}^n : x_j \in (a_j^k, b_j^k), \forall j\}$  for some  $a_j^k < b_j^k \in \mathbb{R}$ ,

and

$$\text{Vol}(I_k) = \prod_{j=1}^n |b_j^k - a_j^k|$$



### Example 4.4

1. “Lower-dimensional” sets have Lebesgue measure zero. For example,  $A = \{x \in \mathbb{R}^2 : x_2 = 0\}$
2. Any **finite** set has Lebesgue measure zero in  $\mathbb{R}^n$ .
3. **Finite Union** of sets that have Lebesgue measure zero has Lebesgue measure zero: If  $A_n$  has Lebesgue measure zero  $\forall n$  then  $\bigcup_{n \in N} A_n$  has Lebesgue measure zero.
4. Every **countable** set (e.g.  $\mathbb{Q}$ ) has Lebesgue measure zero.
5. No open set in  $\mathbb{R}^n$  has Lebesgue measure zero.

### 4.3.2 Sard’s Theorem

#### Theorem 4.3 (Sard’s Theorem)

Let  $X \subseteq \mathbb{R}^n$  be open, and  $f : X \rightarrow \mathbb{R}^m$  be  $C^r$  with  $r \geq 1 + \max\{0, n - m\}$ . Then the set of all critical values of  $f$  has Lebesgue measure zero.



### 4.3.3 Transversality Theorem

#### Theorem 4.4 (Transversality Theorem)

Let  $X \subseteq \mathbb{R}^n$  and  $A \subseteq \mathbb{R}^p$  be open, and  $f : X \times A \rightarrow \mathbb{R}^m$  be  $C^r$  with  $r \geq 1 + \max\{0, n - m\}$ . Suppose that 0 is a regular value of  $f$  (that is all  $(x, a)$  such that  $f(x, a) = 0$  are regular points). Then,

1.  $\exists A_0 \subseteq A$  such that  $A \setminus A_0$  has Lebesgue measure zero.
2.  $\forall a \in A_0$ , 0 is a regular value of  $f_a = f(\cdot, a)$ .



**Example 4.5**  $f : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  s.t.  $f(x, y, z, w) = (g(x) + y, z^3 + 1, w + x + y^2)$

# Chapter 5 Fixed Point Theorem

## 5.1 Contraction Mapping Theorem (@ Lec 05 of ECON 204)

### 5.1.1 Contractions

#### Definition 5.1

Let  $(X, d)$  be a nonempty complete metric space. An operator is a function  $T : X \rightarrow X$ . An operator  $T$  is a **contraction of modulus  $\beta$**  if  $\beta < 1$  and

$$d(T(x), T(y)) \leq \beta d(x, y), \forall x, y \in X$$



A contraction shrinks distances by a *uniform* factor  $\beta < 1$ .

### 5.1.2 Theorem: Contraction $\Rightarrow$ Uniformly Continuous

#### Theorem 5.1 (Contraction $\Rightarrow$ Uniformly Continuous)

*Every contraction is uniformly continuous.*



#### Proof 5.1

Let  $\delta = \frac{\varepsilon}{\beta}$ .

### 5.1.3 Theorem: Blackwell's Sufficient Conditions for Contraction

Let  $X$  be a set, and let  $B(X)$  be the set of all bounded functions from  $X$  to  $\mathbb{R}$ . Then  $(B(X), \|\cdot\|_\infty)$  is a normed vector space.

(Notice that below we use shorthand notation that identifies a constant function with its constant value in  $\mathbb{R}$ , that is, we write interchangeably  $a \in \mathbb{R}$  and  $a : X \rightarrow \mathbb{R}$  to denote the function such that  $a(x) = a, \forall x \in X$ .)

#### Theorem 5.2 (Blackwell's Sufficient Conditions)

Consider  $B(X)$  with the sup norm  $\|\cdot\|_\infty$ . Let  $T : B(X) \rightarrow B(X)$  be an operator satisfying

1. (monotonicity)  $f(x) \leq g(x), \forall x \in X \Rightarrow (Tf)(x) \leq (Tg)(x), \forall x \in X$
2. (discounting)  $\exists \beta \in (0, 1)$  such that for every  $a \geq 0$  and  $x \in X$ ,

$$(T(f + a))(x) \leq (Tf)(x) + \beta a$$

Then  $T$  is a contraction with modulus  $\beta$ .



**Proof 5.2**

Fix  $f, g \in B(X)$ . By the definition of the sup norm,

$$f(x) \leq g(x) + \|f - g\|_{\infty} \forall x \in X$$

Then

$$(Tf)(x) \leq (T(g + \|f - g\|_{\infty})) (x) \leq (Tg)(x) + \beta \|f - g\|_{\infty} \quad \forall x \in X$$

where the first inequality above follows from monotonicity, and the second from discounting. Thus

$$(Tf)(x) - (Tg)(x) \leq \beta \|f - g\|_{\infty} \quad \forall x \in X$$

Reversing the roles of  $f$  and  $g$  above gives

$$(Tg)(x) - (Tf)(x) \leq \beta \|f - g\|_{\infty} \quad \forall x \in X$$

Thus

$$\|T(f) - T(g)\|_{\infty} \leq \beta \|f - g\|_{\infty}$$

Thus  $T$  is a contraction with modulus  $\beta$

## 5.2 Fixed Point Theorem (@ Lec 05 of ECON 204)

### 5.2.1 Fixed Point

**Definition 5.2 (Fixed Point)**

A **fixed point** of an operator  $T$  is element  $x^* \in X$  such that  $T(x^*) = x^*$ .

**Definition 5.3 (Fixed Point of Function)**

Let  $X$  be a nonempty set and  $f : X \rightarrow X$ . A point  $x^* \in X$  is a **fixed point** of  $f$  if  $f(x^*) = x^*$ .



**Example 5.1** Let  $X = \mathbb{R}$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$

1.  $f(x) = 2x$  has fixed point:  $x = 0$ .
2.  $f(x) = x$  has fixed points:  $x \in \mathbb{R}$ .
3.  $f(x) = x + 1$  doesn't have fixed points.

## 5.2.2 Contraction Mapping Theorem: Existence and Uniqueness of Fixed Point

### Theorem 5.3 (Contraction Mapping Theorem)

Let  $(X, d)$  be a nonempty complete metric space and  $T : X \rightarrow X$  a contraction with modulus  $\beta < 1$ .

Then

1.  $T$  has a unique fixed point  $x^*$ .
2. For every  $x_0 \in X$ , the sequence defined by

$$\begin{aligned}x_1 &= T(x_0) \\x_2 &= T(x_1) = T(T(x_0)) = T^2(x_0) \\&\vdots \\x_{n+1} &= T(x_n) = T^{n+1}(x_0)\end{aligned}$$

converges to  $x^*$ .



Note that the theorem asserts both the **existence** and **uniqueness** of the fixed point, as well as giving an **algorithm** to find the fixed point of a contraction.

### Proof 5.3

Define the sequence  $\{x_n\}$  as above. Then,

$$\begin{aligned}d(x_{n+1}, x_n) &= d(T(x_n), T(x_{n-1})) \\&\leq \beta d(x_n, x_{n-1}) \\&\leq \beta^n d(x_1, x_0)\end{aligned}$$

Then for any  $n > m$ ,

$$\begin{aligned}d(x_n, x_m) &\leq d(x_1, x_0) \sum_{i=m}^{n-1} \beta^i \\&< d(x_1, x_0) \sum_{i=m}^{\infty} \beta^i \\&= \frac{\beta^m}{1-\beta} d(x_1, x_0) \rightarrow 0 \text{ as } m \rightarrow \infty\end{aligned}$$

Fixed  $\varepsilon > 0$ , we can choose  $N(\varepsilon)$  such that  $\forall n, m > N(\varepsilon)$ ,

$$d(x_n, x_m) < \frac{\beta^m}{1-\beta} d(x_1, x_0) < \varepsilon$$

Therefore,  $\{x_n\}$  is Cauchy. Since  $(X, d)$  is complete,  $x_n \rightarrow x^*$  for some  $x^* \in X$ .

Next we show that  $x^*$  is a fixed point of  $T$ .

$$\begin{aligned} T(x^*) &= T\left(\lim_{n \rightarrow \infty} x_n\right) \\ &= \lim_{n \rightarrow \infty} T(x_n) \text{ since } T \text{ is continuous} \\ &= \lim_{n \rightarrow \infty} x_{n+1} \\ &= x^* \end{aligned}$$

so  $x^*$  is a fixed point of  $T$ .

Finally, we show that there is at most one fixed point. Suppose  $x^*$  and  $y^*$  are both fixed points of  $T$ , so  $T(x^*) = x^*$  and  $T(y^*) = y^*$ . Then

$$\begin{aligned} d(x^*, y^*) &= d(T(x^*), T(y^*)) \\ &\leq \beta d(x^*, y^*) \\ \Rightarrow (1 - \beta)d(x^*, y^*) &\leq 0 \\ \Rightarrow d(x^*, y^*) &\leq 0 \end{aligned}$$

So  $d(x^*, y^*) = 0$ , which implies  $x^* = y^*$ .

### 5.2.3 Theorem: Fixed Point's Continuous Dependence on Parameters

#### Theorem 5.4 (Continuous Dependence on Parameters)

Let  $(X, d)$  and  $(\Omega, \rho)$  be two metric spaces and  $T : X \times \Omega \rightarrow X$ . For each parameter  $\omega \in \Omega$  let  $T_\omega : X \rightarrow X$  be defined by  $T_\omega(x) = T(x, \omega)$ .

Suppose (1).  $(X, d)$  is complete, (2).  $T$  is continuous in  $\omega$  (that is  $T(x, \cdot) : \Omega \rightarrow X$  is continuous for each  $x \in X$ ), and (3).  $\exists \beta < 1$  such that  $T_\omega$  is a contraction of modulus  $\beta \forall \omega \in \Omega$ .

Then the fixed point function (about parameter  $\omega$ )  $x^* : \Omega \rightarrow X$  defined by  $x^*(\omega) = T_\omega(x^*(\omega))$  is continuous.



## 5.3 Brouwer's Fixed Point Theorem (@ Lec 13 of ECON 204)

### 5.3.1 Simple One: One-dimension

#### Theorem 5.5

Let  $X = [a, b]$  for  $a, b \in \mathbb{R}$  with  $a < b$  and let  $f : X \rightarrow X$  be continuous. Then  $f$  has a fixed point.



#### Proof 5.4

Easily proved by Intermediate Value Theorem.

### 5.3.2 Brouwer's Fixed Point Theorem

#### Theorem 5.6 (Brouwer's Fixed Point Theorem)

Let  $X \subseteq \mathbb{R}^n$  be nonempty, **compact**, and **convex**, and let  $f : X \rightarrow X$  be continuous. Then  $f$  has a fixed point.



#### Proof 5.5

Consider the case when the set  $X$  is the unit ball in  $\mathbb{R}^n$ .

Using a fact that "Let  $B$  be the unit ball in  $\mathbb{R}^n$ . Then there is no continuous function  $h : B \rightarrow \partial B$  such that  $h(x_0) = x_0$  for every  $x_0 \in \partial B$ ", which is intuitive but hard to prove. (See *J. Franklin, Methods of Mathematical Economics*, for an elementary (but long) proof.)

Then prove by contradiction: suppose  $f$  has no fixed points in  $B$ . That is,  $\forall x \in B, x \neq f(x)$ . Since  $x$  and its image  $f(x)$  are distinct points in  $B$  for every  $x$ , we can carry out the following construction. For each  $x \in B$ , construct the line segment originating at  $f(x)$  and going through  $x$ . Let  $g(x)$  denote the intersection of this line segment with  $\partial B$ . This construction gives a continuous function  $g : B \rightarrow \partial B$ . Furthermore, notice that if  $x_0 \in \partial B$ , then  $x_0 = g(x_0)$ . Then,  $g$  gives  $g(x) = x, \forall x \in \partial B$ . Since there are no such functions by the fact above, we have a contradiction.

## 5.4 Fixed Points for Correspondences (@ Lec 13 of ECON 204)

### 5.4.1 Definition

#### Definition 5.4 (Fixed Points for Correspondences)

Let  $X$  be nonempty and  $\psi : X \rightarrow 2^X$  be a correspondence. A point  $x^* \in X$  is a fixed point of  $\psi$  if  $x^* \in \psi(x^*)$ .



**Note** We only need  $x^*$  to be in  $\psi(x^*)$ , not  $\{x^*\} = \psi(x^*)$ . That is,  $\psi$  need not be single-valued at  $x^*$ . So  $x^*$  can be a fixed point of  $\psi$  but there may be other elements of  $\psi(x^*)$  different from  $x^*$ .

### 5.4.2 Kakutani's Fixed Point Theorem

#### Theorem 5.7 (Kakutani's Fixed Point Theorem)

Let  $X \subseteq \mathbb{R}^n$  be a non-empty, **compact**, **convex** set and  $\psi : X \rightarrow 2^X$  be an **upper hemi-continuous** correspondence with non-empty, **compact**, **convex** values. Then  $\psi$  has a fixed point in  $X$ .



# Chapter 6 Differential Equations (@ Lec 14/15 of ECON 204)

## 6.1 Initial Value Problem

### 6.1.1 Initial Value Problem

#### Definition 6.1 (Differential Equation)

A **differential equation** is an equation of the form  $y'(t) = F(y(t), t)$ .



where  $F : U \rightarrow \mathbb{R}^n$  and  $U$  is an open subset of  $\mathbb{R}^n \times \mathbb{R}$ .

#### Definition 6.2 (Initial Value Problem)

1. An initial value problem is a differential equation combined with an initial condition

$$y(t_0) = y_0$$

where  $(y_0, t_0) \in U$

2. A solution of the initial value problem is a differentiable function  $y : (a, b) \rightarrow \mathbb{R}^n$  such that  $t_0 \in (a, b)$ ,  $y(t_0) = y_0$  and, for all  $t \in (a, b)$ ,  $y'(t) = F(y(t), t)$ .
3. The general solution of the differential equation is the family of all solutions for all initial values  $(y_0, t_0) \in U$ .



### 6.1.2 Theorem: Initial Value Problem, Continuous $\rightarrow$ Existence; Lipschitz $\rightarrow$ Uniqueness

#### Theorem 6.1 (Initial Value Problem, Continuous $\rightarrow$ Existence; Lipschitz $\rightarrow$ Uniqueness)

Consider the initial value problem

$$y'(t) = F(y(t), t), \quad y(t_0) = y_0$$

$U$  is an open subset of  $\mathbb{R}^n \times \mathbb{R}$  containing  $(y_0, t_0)$ .

1. Suppose  $F : U \rightarrow \mathbb{R}^n$  is continuous. Then the initial value problem has a solution.
2. If, in addition,  $F$  is Lipschitz in  $y$  on  $U$   
(i.e. there is a constant  $K$  such that for all  $(y, t), (\hat{y}, t) \in U$ ,  $|F(y, t) - F(\hat{y}, t)| \leq K|y - \hat{y}|$ ),  
then there is an interval  $(a, b)$  containing  $t_0$  such that the solution is unique on  $(a, b)$ .



**Example 6.1** Consider the initial value problem

$$y'(t) = 1 + y^2(t), \quad y(0) = 0$$

Here, we have  $F(y, t) = 1 + y^2$  which is Lipschitz in  $y$  over  $U = V \times \mathbb{R}$ , provided that  $V$  is bounded, but not

over all of  $\mathbb{R} \times \mathbb{R}$ . The theorem tells us that the initial value problem has a unique solution over some interval of  $t$ ,  $(a, b)$ , with  $0 \in (a, b)$ .

We can find  $y(t) = \tan t$  is a unique solution in  $(a, b) \subset (-\frac{\pi}{2}, \frac{\pi}{2})$  with  $a < 0, b > 0$ .

**Example 6.2** Consider the initial value problem

$$y'(t) = 2\sqrt{|y|}, y(0) = 0$$

As  $F(y) = 2\sqrt{|y|}$  is not locally Lipschitz in  $y$  at  $y = 0$ , we can find for all  $\alpha \geq 0$ ,

$$y_\alpha(t) = \begin{cases} 0, & t \leq \alpha \\ (t - \alpha)^2, & t \geq \alpha \end{cases}$$

is a solution of the initial value problem, which is not unique.

## 6.2 Autonomous Differential Equation

### 6.2.1 Autonomous Differential Equation

#### Definition 6.3 (Autonomous Differential Equation)

An **autonomous differential equation** is a differential equation of the form  $y'(t) = F(y(t))$  where  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  depends on  $t$  only through the value of  $y(t)$ .

A **stationary point** of an autonomous differential equation is a point  $y_s \in \mathbb{R}^n$  such that  $F(y_s) = 0$ .



We study the qualitative properties of autonomous differential equations by looking for stationary points. The constant function

$$y(t) = y_s$$

is a solution (and the unique solution when  $F$  is Lipschitz) of the initial value problem

$$y' = F(y), y(t_0) = y_s$$

If  $F$  is  $C^2$ , then Taylor's Theorem implies that near a stationary point  $y_s$ ,

$$\begin{aligned} F(y_s + h) &= F(y_s) + DF(y_s)h + O(|h|^2) \\ &= DF(y_s)h + O(|h|^2) \end{aligned}$$

Thus, when we are sufficiently close to the stationary point, the solutions of the autonomous differential equation are closely approximated by the solutions of the linear differential equation

$$y' = (y - y_s)' = DF(y_s)(y - y_s)$$

Thus, we study solutions of linear differential equations, using linear algebra. For this we first recall the formulation of complex exponentials, which are central in the general solution of linear differential equations.

### 6.2.2 Complex Exponentials

By the Taylor series,  $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ . If  $x = a + ib \in \mathbb{C}$ ,  $a, b \in \mathbb{R}$ , we have

$$\begin{aligned} e^x &= e^{a+ib} = e^a e^{ib} \\ &= e^a \left( \sum_{k=0}^{\infty} (-1)^k \frac{b^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{b^{2k+1}}{(2k+1)!} \right) \\ &= e^a (\cos b + i \sin b) \end{aligned}$$

Hence,

1. If  $a < 0$ , then  $e^{tx} \rightarrow 0$  as  $t \rightarrow \infty$ .
2. If  $a > 0$ , then  $|e^{tx}| \rightarrow \infty$  as  $t \rightarrow \infty$ .
3. If  $a = 0$ , then  $|e^{tx}| \rightarrow 1$  for all  $t \in \mathbb{R}$ .

### 6.2.3 Solutions of Simple Case: $y'(t) = ay(t)$

#### Proposition 6.1 (Simple Case)

Given  $y'(t) = ay(t)$ , we have

$$y(t) = ce^{at}, c \in \mathbb{R}^*$$



#### Proof 6.1

$$\begin{aligned} y'(t) &= \frac{dy}{dt} = ay \\ \frac{dy}{y} &= adt \\ \int \frac{1}{y} dy &= \int adt \\ \ln y &= at + c', c' \in \mathbb{R} \\ y(t) &= e^{c'} e^{at} = ce^{at}, c \in \mathbb{R}^* \end{aligned}$$

#### Proposition 6.2 (Simple Case with Initial Value)

Given initial value  $(t_0, y_0)$ ,  $y'(t) = ay(t)$ , we have

$$y(t) = y_0 e^{a(t-t_0)}$$



### 6.2.4 Solutions of Special Case: Linear Differential Equations with Constant Coefficients

Let  $M \in \mathbb{R}^{n \times n}$ . The linear differential equation

$$y' = (y - y_s)' = M(y - y_s)$$

has a complete solution in closed form.

The matrix representation

$$M = DF(y_s)$$

need not be symmetric, hence may not be diagonalizable. If  $M$  is diagonalizable over  $\mathbb{C}$ , the complete solution takes the following simple form:

**Theorem 6.2 (Solutions of Linear Differential Equations with Constant Coefficients)**

Consider the linear differential equation

$$y' = (y - y_s)' = M(y - y_s)$$

where  $M$  is a real  $n \times n$  matrix. Suppose that  $M$  can be diagonalized over the complex field  $\mathbb{C}$ . Let  $U$  be the standard basis of  $\mathbb{R}^n$  and  $V = \{v_1, \dots, v_n\}$  be a basis of (complex) eigenvectors corresponding to the eigenvalues  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  of  $M$ . Then the solution of the initial value problem is given by

$$y(t) = y_s + P^{-1} \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & 0 & 0 \\ 0 & e^{\lambda_2(t-t_0)} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_n(t-t_0)} \end{bmatrix} P(y(t_0) - y_s) \quad (6.1)$$

where  $P = \text{Mtx}_{V,U}(id)$ , and the general complex solution is obtained by allowing  $y(t_0)$  to vary over  $\mathbb{C}^n$ ; it has  $n$  complex degrees of freedom. The general real solution is obtained by allowing  $y(t_0)$  to vary over  $\mathbb{R}^n$ ; it has  $n$  real degrees of freedom. Every real solution is a linear combination of the real and imaginary parts of a complex solution. In particular,

1. If the real part of each eigenvalue is less than zero, all solutions converge to  $y_s$ .
2. If the real part of each eigenvalue is greater than zero, all solutions diverge from  $y_s$  and tend to infinity.
3. If the real parts of some eigenvalues are less than zero and the real parts of other eigenvalues are greater than zero, solutions follow roughly hyperbolic paths.
4. If the real parts of all eigenvalues are zero, all solutions follow closed cycles around  $y_s$ .



**Note** If one or more of the eigenvalues are complex, each of the three matrices in Equation 6.1 will contain complex entries, but the product of the three matrices is real. Thus, if the initial condition  $y_0$  is real, Equation 6.1 gives us a real solution; indeed, it gives us the unique solution of the initial value problem.



**Note** Given a fixed time  $t_0$ , the general real solution is obtained by varying the initial values of  $y(t_0)$  over  $\mathbb{R}^n$ , which provides  $n$  real degrees of freedom. You might think that varying  $t_0$  provides one additional degree of freedom, but it doesn't. Given any solution satisfying the initial condition  $y(t_0) = y_0$ , the solution is defined on

some interval  $(t_0 - \delta, t_0 + \delta)$ ; given  $t_1 \in (t_0 - \delta, t_0 + \delta)$ , let  $y_1 = y(t_1)$ ; then the solution with initial condition  $y_1 = y(t_1)$  is the same as the solution with initial condition  $y_0 = y(t_0)$ . The same holds true for the general complex solution.

**Proof 6.2**

Let  $P = \text{Mtx}_{V,U}(id)$ . Rewrite the differential equation in terms of a new variable

$$z = Py$$

Let  $z_s = Py_s$ , then we have

$$\begin{aligned} z - z_s &= P(y - y_s) \\ (z - z_s)' &= z' = Py' \\ &= PM(y - y_s) \\ &= PMP^{-1}(z - z_s) \\ &= B(z - z_s) \end{aligned}$$

where  $B = \text{diag}(\lambda_1, \dots, \lambda_n)$ .

Thus, the  $i^{\text{th}}$  component of  $(z(t) - z_s)$  satisfies

$$(z(t) - z_s)_i' = \lambda_i(z(t) - z_s)_i$$

by the proposition 6.2, we have

$$(z(t) - z_s)_i = e^{\lambda_i(t-t_0)}(z(t_0) - z_s)_i$$

so,

$$z(t) - z_s = \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & 0 & 0 \\ 0 & e^{\lambda_2(t-t_0)} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_n(t-t_0)} \end{bmatrix} (z(t_0) - z_s)$$

$$\begin{aligned}
y(t) - y_s &= P^{-1} \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & 0 & 0 \\ 0 & e^{\lambda_2(t-t_0)} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_n(t-t_0)} \end{bmatrix} (z(t_0) - z_s) \\
&= P^{-1} \begin{bmatrix} e^{\lambda_1(t-t_0)} & 0 & 0 & 0 \\ 0 & e^{\lambda_2(t-t_0)} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & e^{\lambda_n(t-t_0)} \end{bmatrix} P(y(t_0) - y_s)
\end{aligned}$$

## 6.2.5 The Form of Real Solutions

We can determine the form of the real solutions once we know the eigenvalues. In an important special case, we can solve for the solution of the initial value problem without calculating the diagonalization, as in Equation 6.1.

### Theorem 6.3

*Consider the differential equation*

$$y' = (y - y_s)' = M(y - y_s)$$

*Suppose that the matrix  $M$  can be diagonalized over  $\mathbb{C}$ . Let the eigenvalues of  $M$  with the correct multiplicity be*

$$a_1 + ib_1, a_1 - ib_1, \dots, a_m + ib_m, a_m - ib_m, a_{m+1}, \dots, a_{n-m}$$

*Then for each fixed  $i = 1, \dots, n$ , every real solution is of the form*

$$(y(t) - y_s)_i = \sum_{j=1}^m e^{a_j(t-t_0)} (C_{ij} \cos b_j(t-t_0) + D_{ij} \sin b_j(t-t_0)) + \sum_{j=m+1}^{n-m} C_{ij} e^{a_j(t-t_0)}$$

*The  $n^2$  parameters*

$$\{C_{ij} : i = 1, \dots, n; j = 1, \dots, n-m\} \cup \{D_{ij} : i = 1, \dots, n; j = 1, \dots, m\}$$

*have  $n$  real degrees of freedom. The parameters are uniquely determined from the  $n$  real initial conditions of an initial value problem.*



## 6.3 Higher Order Differential Equations

When we consider a differential equation of order  $n$  is an equation of the form

$$y^{(n)}(t) = F(y(t), y'(t), \dots, y^{n-1}(t), t)$$

We can always rewrite

$$\bar{y}(t) = \begin{bmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{n-1}(t) \end{bmatrix}$$

### 6.3.1 Second Order Linear Differential Equations

Consider the second order differential equation  $y'' = cy + by'$  with  $b, c \in \mathbb{R}$ .

Rewrite  $\bar{y}(t) = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix}$ , then

$$\bar{y}'(t) = \begin{bmatrix} y'(t) \\ y''(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ c & b \end{bmatrix} \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ c & b \end{bmatrix} \bar{y}$$

By the Theorem 6.2, we can now solve it. Specifically, the eigenvalues of  $\begin{bmatrix} 0 & 1 \\ c & b \end{bmatrix}$  are  $\frac{b \pm \sqrt{b^2 + 4c}}{2}$ .

**Example 6.3** Consider the second order linear differential equation  $y'' = 2y + y'$ , we can solve the eigenvalues  $\lambda_1 = 2, \lambda_2 = -1$  and eigenvectors  $v_1 = (1, 2), v_2 = (1, -1)$ . The general solution of  $\bar{y}(t) = \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix}$  can be solved by the Theorem 6.2.

## 6.4 Inhomogeneous Linear Differential Equations with Nonconstant Coefficients

Consider the inhomogeneous linear differential equation

$$y' = M(t)y + H(t) \quad (1)$$

where  $M$  is continuous function from  $t$  to the set of  $n \times n$  matrices; and  $H$  is continuous function from  $t$  to  $\mathbb{R}^n$ .

### 6.4.1 Theorem: General Solution of Inhomo = General Solution of Homo + Particular Solution of Inhomo

There is a close relationship between solutions of the *inhomogeneous* linear differential equation 1 and the associated *homogeneous* linear differential equation

$$y' = M(t)y \quad (2)$$

#### Theorem 6.4 (General Solution of Inhomo=General Solution of Homo+Particular Solution of Inhomo)

*The general solution of the inhomogeneous linear differential equation 1 is*

$$y_h + y_p$$

*where  $y_h$  is the general solution of the homogeneous linear differential equation 2 and  $y_p$  is any particular solution of the inhomogeneous linear differential equation 1.*



Hence, to find general solution of inhomogeneous equation:

1. Find general solution of homogeneous equation;
2. Find a particular solution of inhomogeneous equation;
3. Add these to get general solution of inhomogeneous equation

### 6.4.2 Matrix Exponentials Notations

In analogy with how we define  $e^x$  for  $x \in \mathbb{R}$ , for an  $n \times n$  matrix  $M$  we define

$$e^M = \sum_{k=0}^{\infty} \frac{M^k}{k!} = I + M + \frac{M^2}{2} + \dots$$

and

$$e^{tM} = \sum_{k=0}^{\infty} \frac{t^k M^k}{k!}$$

The matrix exponential satisfies many properties analogous to the exponential function in  $\mathbb{R}$  and  $\mathbb{C}$ . Here are a few of the most important properties, which can be established fairly directly from the definitions above.

- If  $D$  is a diagonal matrix with diagonal  $d_1, \dots, d_n$ ,

$$e^D = \begin{pmatrix} e^{d_1} & 0 & \cdots & 0 \\ 0 & e^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{d_n} \end{pmatrix}$$

- $e^{A+B} = e^A e^B$  if  $AB = BA$  (this is not necessarily valid for matrices  $A$  and  $B$  that do not commute)
- $e^{P^{-1}AP} = P^{-1}e^A P$

Then, if  $M$  is diagonalizable, so  $M = P^{-1}DP$  for a diagonal matrix  $D$ , then

$$e^M = P^{-1}e^D P = P^{-1} \begin{pmatrix} e^{d_1} & 0 & \cdots & 0 \\ 0 & e^{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{d_n} \end{pmatrix} P$$

- $g(t) = e^{tM}$  is differentiable (in fact  $C^\infty$ ) and  $g'(t) = M e^{tM}$

### 6.4.3 Theorem: Particular Solution of Inhomogeneous Linear Differential Equations with Constant Coefficients (Matrix)

#### Theorem 6.5 (Particular Solution with Constant Coefficients (Matrix))

Consider the inhomogeneous linear differential equation 1, and suppose that  $M(t)$  is a constant matrix  $M$ , independent of  $t$ . A particular solution of the inhomogeneous linear differential equation 1, satisfying the initial condition  $y_p(t_0) = y_0$ , is given by

$$y_p(t_0) = e^{(t-t_0)M} y_0 + \int_{t_0}^t e^{(t-s)M} H(s) ds \quad (3)$$



# Chapter 7 Big $\mathcal{O}$ and Small $o$ Notation

## 7.1 Definition

### Definition 7.1 ("O" and "o")

1.  $y = O(|h|^n)$  as  $h \rightarrow 0$  – read “ $y$  is big-Oh of  $|h|^n$ ” – means  $\exists K, \delta > 0$  s.t.  $|h| < \delta \Rightarrow |y| \leq K|h|^n$
2.  $y = o(|h|^n)$  as  $h \rightarrow 0$  – read “ $y$  is little-Oh of  $|h|^n$ ” – means  $\lim_{h \rightarrow 0} \frac{|y|}{|h|^n} = 0$



### Complexity:

#### Definition 7.2

A sequence  $f(n)$  is  $O(1)$  if  $\lim_{n \rightarrow \infty} f(n) < \infty$ .



#### Definition 7.3

A sequence  $f(n)$  is  $O(g(n))$  if  $\frac{f(n)}{g(n)}$  is  $O(1)$ .



#### Definition 7.4

A sequence  $f(n)$  is  $o(1)$  if  $\lim_{n \rightarrow \infty} \sup f(n) = 0$ .



#### Definition 7.5

A sequence  $f(n)$  is  $o(g(n))$  if  $\lim_{n \rightarrow \infty} \sup \frac{f(n)}{g(n)} = 0$ .



#### Definition 7.6

A sequence  $f(n)$  is asymptotic to  $g(n)$  if  $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = 1$ . (This is denoted by  $f(n) \sim g(n)$  as  $a \rightarrow \infty$ )



For two scalar functions  $f(x) \in \mathbb{R}, g(x) \in \mathbb{R}_+$ , where  $x \in \mathbb{R}$ , we write:

1.  $f(x) = \mathcal{O}(g(x))$  if  $\limsup_{x \rightarrow \infty} \frac{|f(x)|}{g(x)} < \infty$ ; we say  $f$  is dominated by  $g$  asymptotically.
2.  $f(x) = \Omega(g(x))$  if  $\liminf_{x \rightarrow \infty} \frac{|f(x)|}{g(x)} > 0$ .
3.  $f(x) = \Theta(g(x))$  if  $f(x) = \mathcal{O}(g(x))$  and  $f(x) = \Omega(g(x))$  both hold.
4.  $f(x) = o(g(x))$  if  $\liminf_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 0$ .

### Example 7.1

$$n^3 + n + 2 = \Omega(1), n^3 + n + 2 = \Omega(n^2)$$

$$n^3 + n + 2 = \Theta(n^3)$$

$$n^3 + n + 2 = o(n^4)$$

### 7.1.1 Extension

$f(x) = \mathcal{O}(g(x))$  as  $x \rightarrow a$  if  $\limsup_{x \rightarrow a} \frac{|f(x)|}{g(x)} < \infty$ .

**Example 7.2**  $\varepsilon^2 + \varepsilon^3 = \mathcal{O}(\varepsilon^2)$  as  $\varepsilon \rightarrow 0$