The Catalogue

Contents

Chapter 1

Mathematical Logic I

1.1 Propositional Logic

List of preliminary references:

1. Two Proofs of Completeness Theorem
2. Hilbert Proof Systems, Formal Proofs,Deduction Theorem

1.1.1 Language

The symbols Σ of propositional logic are divided into 3 type

1. The collection of countably many letters Σ_A: say p_1, p_2, p_3 and so on.
2. The collection of parentheses Σ_O: for example, "(" and ")", or "[" and "]".
3. The collection of connectives Σ_C, which consists of $\neg$ and $\rightarrow$.

Definition 1.1. The language L of propositional logic, consisting of *well-formed formulas* (or wffs), is recursively constructed as follows

1. The letters in Σ_A are wffs. Specifically, these are said to be *atomic*.
2. Given a wff α, then $\neg(\alpha)$ is a wff.
3. Given 2 wffs α and β, then $(\alpha) \rightarrow (\beta)$ is a wff.

The wffs that are not atomic is then called *compound* formula.

Note: parentheses can sometimes be removed for clarity (especially if the formula inside is atomic), given that we usually evaluate negation ($\neg$) first, then implication ($\rightarrow$) from left to right.

1.1.2 Hilbert system

The Hilbert system is a deductive system (over the language of propositional logic) such that

1. The followings are *axioms* for given wffs α, β, γ

   (IP1) $\alpha \to (\beta \to \alpha)$

   (IP2) $[\alpha \to (\beta \to \gamma)] \to [(\alpha \to \beta) \to (\alpha \to \gamma)]$

   (PNC) $(\neg\alpha \to \neg\beta) \to [(\neg\alpha \to \beta) \to \alpha]$

2. The following is an *inference rule* for given wffs α and β

   $$\alpha, \alpha \to \beta \vdash \beta$$

   That is, from α and $\alpha \to \beta$, deduce β. The schema of these rules is termed *modus ponens*.

1.1.3 Proof

Let Γ be a collection of wffs, called *hypothesis*. A *proof* (or *derivation*) of a wff φ from Γ is a finite sequence of wffs, or *steps*, $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ (with $\lambda_n = \varphi$) , where each is either an axiom, belongs to Γ, or a result of MP from previous steps (i.e. $\lambda_i, \lambda_j \vdash_{MP} \lambda_k$ for some $i, j < k$).

If there is such proof, we say Γ *syntactically entails* φ, and denote $\Gamma \vdash \varphi$. If Γ is empty, then we called φ a *theorem* (of propositional logic), and denote $\vdash \varphi$.

Remark 1.2.

1. If $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a proof of φ from Γ, then $(\lambda_1, \lambda_2, \ldots, \lambda_{i-1})$ is a proof of λ_i (albeit there can be shorter one).

2. If Δ contains every wff in Γ (i.e. $\Delta \supseteq \Gamma$), then any proof of φ from Γ is also a proof of Δ. In other words, $\Delta \vdash \varphi$ whenever $\Gamma \vdash \varphi$.

   Thus, in particular, any theorem of propositional logic is syntactically entailed by arbitrary collection of hypothesis Γ.

3. Suppose $(\lambda_1, \lambda_2, \ldots, \lambda_n = \varphi)$ is a sequence of steps where each is either provable from Γ, or an application of MP to some previous steps, then we can concatenate the proofs of λ_i into a proof of φ! This allows us to use theorems to prove other theorems, rather than starting from scratch (i.e. axioms).

Lemma 1.3 (Reflexivity)**.** For each φ, we have $\varphi \to \varphi$.

Proof.

1. Substitute $\alpha, \gamma := \varphi$ into IP2
$$[\varphi \to (\beta \to \varphi)] \to [(\varphi \to \beta) \to (\varphi \to \varphi)]$$
2. Apply MP to IP1 and step (1),
$$(\varphi \to \beta) \to (\varphi \to \varphi)$$
3. Substitute $\beta := \theta \to \varphi$ (say for an atomic formula $\theta := p$) into step (2)
$$[\varphi \to (\theta \to \varphi)] \to (\varphi \to \varphi)$$
4. Apply MP to IP1 and step (3), we finally get $\varphi \to \varphi$.

□

Theorem 1.4 (Deduction Metatheorem)**.** $\Gamma, \varphi \vdash \psi$ if and only if $\Gamma \vdash \varphi \to \psi$. In fact, we can construct a proof of $\varphi \to \psi$ from Γ based on a proof of ψ from Γ and φ.

Proof. If $\Gamma \vdash \varphi \to \psi$, then by MP, we get $\Gamma, \varphi \vdash \psi$. More precisely, given a proof of $\varphi \to \psi$ from Γ, then add φ and ψ to it to get a proof of ψ from Γ and φ, with ψ being the result of MP applying to $\varphi \to \psi$ and φ.

Vice versa, suppose $(\lambda_1, \lambda_2, \ldots, \lambda_n = \psi)$ is a proof of ψ from Γ and φ. We will convert each λ_i to $\varphi \to \lambda_i$, with some steps in between so that $\varphi \to \lambda_i$ is provable from Γ.

Inductively, for any given i, suppose we already construct $\theta_1, \theta_2, \ldots, \theta_{s_i}$ where some θ_t is of the form $\varphi \to \lambda_j$ (for $j < i$) and vice versa, then

1. If λ_i is an axiom, or belongs to Γ:
   (a) Set $\theta_{s_i+1} := \lambda_i$ as first step.
   (b) Next, from IP1, denote $\theta_{s_i+2} := \lambda_i \to (\varphi \to \lambda_i)$ as the next step.
   (c) Finally, apply MP to previous ones to get $\theta_{s_i+3} := \varphi \to \lambda_i$ as the resultant step.
2. If λ_i is φ: since $\varphi \to \varphi$ is a theorem, we let $\theta_{s_i+1} := \varphi \to \lambda_i$.
3. If λ_i is an employment of MP to λ_j and λ_k (for some $j, k < i$): without loss of generality (WLOG), λ_k must be of the form $\lambda_j \to \lambda_i$. So we already get $\theta_{t_j} := \varphi \to \lambda_j$ and $\theta_{t_k} := \varphi \to (\lambda_j \to \lambda_i)$ from previous steps (by induction hypothesis), let the following steps be a proof of $\varphi \to \lambda_i$ based on those.
   (a) From IP2, we set $\theta_{s_i+1} := [\varphi \to (\lambda_j \to \lambda_i)] \to [(\varphi \to \lambda_j) \to (\varphi \to \lambda_i)]$.

(b) Apply MP to previous step θ_{s_i+1} and θ_{t_k}, we get the next step $\theta_{s_i+2} := (\varphi \to \lambda_j) \to (\varphi \to \lambda_i)$.

(c) Apply MP again, but this time to θ_{s_i+2} and θ_{t_j} to get $\theta_{s_i+3} := \varphi \to \lambda_i$.

After the final step $\theta_m := \varphi \to \psi$, we get a sequence $(\theta_1, \theta_2, \ldots, \theta_m)$ where each θ_i is either

1. An axiom, or belongs to Γ: this occurs in (1a), (1b), and (3a).
2. An application of MP to previous steps: this occurs in (1c), (3b), and (3c).
3. $\varphi \to \varphi$: this occurs in (2).

Since each θ_i does not depend on the hypothesis φ but only Γ, by Remark 1.2 and Lemma 1.3, such sequence can serve as a proof of $\varphi \to \psi$ from Γ only. □

Remark 1.5. If there are multiple usages of deduction metatheorem, we convert the proof one at a time, the former to the latter. For example, if $\Gamma, \varphi, \psi \vdash \xi$ and we want to show that $\Gamma \vdash \varphi \to (\psi \to \xi)$, first convert the proof of ξ from Γ, φ and ψ, to a proof of $\psi \to \xi$ from Γ and φ. Then convert such proof to a proof of $\varphi \to (\psi \to \xi)$ from Γ.

Proposition 1.6 (Hypothetical Syllogism - HS)**.** Given φ, ψ, ξ, then the followings hold

1. $(\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$
2. $(\psi \to \xi) \to [(\varphi \to \psi) \to (\varphi \to \xi)]$

Proof. Applying MP repeatedly to $\varphi, \varphi \to \psi$ and $\psi \to \xi$, we get $\varphi, \varphi \to \psi, \psi \to \xi \vdash \xi$. So by deduction theorem, we get

$$\begin{aligned} \varphi \to \psi, \psi \to \xi &\vdash \varphi \to \xi \\ \varphi \to \psi &\vdash (\psi \to \xi) \to (\varphi \to \xi) \\ &\vdash (\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)] \end{aligned}$$

Similarly, $(\psi \to \xi) \to [(\varphi \to \psi) \to (\varphi \to \xi)]$.

*To produce a proof of $(\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$, first note that the following sequence is a proof for $\varphi, \varphi \to \psi, \psi \to \xi \vdash \xi$.

1. From the hypothesis, φ.
2. From the hypothesis, $\varphi \to \psi$.
3. Apply MP to steps (1) and (2), we get ψ.
4. From the hypothesis, $\psi \to \xi$.

5. Finally, apply MP to steps (3) and (4), we get ξ.

*Hence, the following is a proof for $\varphi \to \psi, \psi \to \xi \vdash \varphi \to \xi$

1. By RX, $\varphi \to \varphi$.

2. From the hypothesis, $\varphi \to \psi$.
3. By IP1, $(\varphi \to \psi) \to [\varphi \to (\varphi \to \psi)]$.
4. Apply MP to steps (2) and (3), $\varphi \to (\varphi \to \psi)$.

5. By IP2, $[\varphi \to (\varphi \to \psi)] \to [(\varphi \to \varphi) \to (\varphi \to \psi)]$.
6. Apply MP to steps (4) and (5), $(\varphi \to \varphi) \to (\varphi \to \psi)$.
7. Apply MP to steps (1) and (6), $\varphi \to \psi$.

8. From the hypothesis, $\psi \to \xi$.
9. By IP1, $(\psi \to \xi) \to [\varphi \to (\psi \to \xi)]$.
10. Apply MP to steps (8) and (9), $\varphi \to (\psi \to \xi)$.

11. By IP2, $[\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$.
12. Apply MP to steps (10) and (11), $(\varphi \to \psi) \to (\varphi \to \xi)$.
13. Apply MP to steps (7) and (12), $\varphi \to \xi$.

Removing some redundancy (steps 1 to 7), we get another proof for $\varphi \to \psi, \psi \to \xi \vdash \varphi \to \xi$

1. From the hypothesis, $\lambda_1 := \varphi \to \psi$.
2. From the hypothesis, $\lambda_2 := \psi \to \xi$.
3. By IP1, $\lambda_3 := \lambda_2 \to (\varphi \to \lambda_2)$.
4. Apply MP to steps (2) and (3), $\lambda_4 := \varphi \to (\psi \to \xi)$.
5. By IP2, $\lambda_4 \to [\lambda_1 \to (\varphi \to \xi)]$.
6. Apply MP to steps (4) and (5), $\lambda_1 \to (\varphi \to \xi)$.
7. Apply MP to steps (1) and (6), $\varphi \to \xi$.

*From the proof above, we get a proof for $\varphi \to \psi \vdash (\psi \to \xi) \to (\varphi \to \xi)$

1. From the hypothesis, $\varphi \to \psi$.

2. By IP1, $(\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \psi)]$.
3. Apply MP to steps (1) and (2), $(\psi \to \xi) \to (\varphi \to \psi)$.
4. By RX, $(\psi \to \xi) \to (\psi \to \xi)$.
5. By IP1, $\lambda_5 := (\psi \to \xi) \to [\varphi \to (\psi \to \xi)]$.
6. By IP1, $\lambda_5 \to [(\psi \to \xi) \to \lambda_5]$.
7. Apply MP to steps (5) and (6), $\lambda_7 := (\psi \to \xi) \to \lambda_5$.
8. By IP2, $\lambda_7 \to [[(\psi \to \xi) \to (\psi \to \xi)] \to \lambda_5]$.
9. Apply MP to steps (7) and (8), $[(\psi \to \xi) \to (\psi \to \xi)] \to \lambda_5$.
10. Apply MP to steps (4) and (9), $\lambda_5 = (\psi \to \xi) \to (\varphi \to (\psi \to \xi))$.
11. By IP2, $\lambda_{11} := [\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$.
12. By IP1, $\lambda_{11} \to [(\psi \to \xi) \to \lambda_{11}]$.
13. Apply MP to steps (11) and (12), $\lambda_{13} := (\psi \to \xi) \to \lambda_{11}$.
14. By IP2, $\lambda_{13} \to [\lambda_5 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]$.
15. Apply MP to steps (13) and (14), $[(\psi \to \xi) \to (\varphi \to (\psi \to \xi))] \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]$.
16. Apply MP to steps (10) and (15), $\lambda_{16} := (\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))$.
17. By IP2, $\lambda_{16} \to [((\psi \to \xi) \to (\varphi \to \psi)) \to ((\psi \to \xi) \to (\varphi \to \xi))]$.
18. Apply MP to steps (16) and (17), $[(\psi \to \xi) \to (\varphi \to \psi)] \to [(\psi \to \xi) \to (\varphi \to \xi)]$.
19. Apply MP to steps (3) and (18), $(\psi \to \xi) \to (\varphi \to \xi)$.

After removing some redundancy (steps 4 to 10),

1. From the hypothesis, $\lambda_1 := \varphi \to \psi$.
2. By IP1, $\lambda_1 \to [(\psi \to \xi) \to \lambda_1]$.
3. Apply MP to steps (1) and (2), $\lambda_3 := (\psi \to \xi) \to \lambda_1$.

4. By IP1, $\lambda_4 := (\psi \to \xi) \to (\varphi \to (\psi \to \xi))$.

5. By IP2, $\lambda_5 := [\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$.
6. By IP1, $\lambda_5 \to [(\psi \to \xi) \to \lambda_5]$.
7. Apply MP to steps (5) and (6), $\lambda_7 := (\psi \to \xi) \to \lambda_5$.

8. By IP2, $\lambda_7 \to [\lambda_4 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]$.
9. Apply MP to steps (7) and (8), $\lambda_4 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]$.
10. Apply MP to steps (4) and (9), $\lambda_{10} := (\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))$.

11. By IP2, $\lambda_{10} \to [\lambda_3 \to ((\psi \to \xi) \to (\varphi \to \xi))]$.
12. Apply MP to steps (10) and (11), $\lambda_3 \to [(\psi \to \xi) \to (\varphi \to \xi)]$.
13. Apply MP to steps (3) and (12), $(\psi \to \xi) \to (\varphi \to \xi)$.

*Finally, we obtain a proof for $\vdash (\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$ (with some redundancy being removed)

1. By IP1, $\lambda_1 := (\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \psi)]$.

2. By IP1, $\lambda_2 := (\psi \to \xi) \to (\varphi \to (\psi \to \xi))$.
3. By IP1, $\lambda_2 \to [(\varphi \to \psi) \to \lambda_2]$.
4. Apply MP to steps (2) and (3), $\lambda_4 := (\varphi \to \psi) \to \lambda_2$.

5. By IP2, $\lambda_5 := [\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$.
6. By IP1, $\lambda_5 \to [(\varphi \to \psi) \to \lambda_5]$.
7. Apply MP to steps (5) and (6), $\lambda_7 := (\varphi \to \psi) \to \lambda_5$.

8. By IP1, $\lambda_8 := \lambda_5 \to [(\psi \to \xi) \to \lambda_5]$.
9. By IP1, $\lambda_8 \to [(\varphi \to \psi) \to \lambda_8]$.
10. Apply MP to steps (8) and (9), $\lambda_{10} := (\varphi \to \psi) \to \lambda_8$

11. By IP2, $\lambda_{10} \to [\lambda_7 \to [(\varphi \to \psi) \to ((\psi \to \xi) \to \lambda_5)]]$.
12. Apply MP to steps (10) and (11), $\lambda_7 \to [(\varphi \to \psi) \to ((\psi \to \xi) \to \lambda_5)]$.

13. Apply MP to steps (7) and (12), $\lambda_{13} := (\varphi \to \psi) \to [(\psi \to \xi) \to \lambda_5]$.

14. By IP2, $\lambda_{14} := [(\psi \to \xi) \to \lambda_5] \to [\lambda_2 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]$.

15. By IP1, $\lambda_{14} \to [(\varphi \to \psi) \to \lambda_{14}]$.

16. Apply MP to steps (14) and (15), $\lambda_{16} := (\varphi \to \psi) \to \lambda_{14}$.

17. By IP2, $\lambda_{16} \to [\lambda_{13} \to [(\varphi \to \psi) \to [[\lambda_2 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]]]]$.

18. Apply MP to steps (16) and (17), $\lambda_{13} \to [(\varphi \to \psi) \to [[\lambda_2 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]]]$.

19. Apply MP to steps (13) and (18), $\lambda_{19} := (\varphi \to \psi) \to [[\lambda_2 \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]]$.

20. By IP2, $\lambda_{19} \to [\lambda_4 \to [(\varphi \to \psi) \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]]$.

21. Apply MP to steps (19) and (20), $\lambda_4 \to [(\varphi \to \psi) \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]]$.

22. Apply MP to steps (4) and (21), $\lambda_{22} := (\varphi \to \psi) \to [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))]$.

23. By IP2, $\lambda_{23} := [(\psi \to \xi) \to ((\varphi \to \psi) \to (\varphi \to \xi))] \to [[(\psi \to \xi) \to (\varphi \to \psi)] \to ((\psi \to \xi) \to (\varphi \to \xi))]$.

24. By IP1, $\lambda_{23} \to [(\varphi \to \psi) \to \lambda_{23}]$.

25. Apply MP to steps (23) and (24), $\lambda_{25} := (\varphi \to \psi) \to \lambda_{23}$.

26. By IP2, $\lambda_{25} \to [\lambda_{22} \to [(\varphi \to \psi) \to [[(\psi \to \xi) \to (\varphi \to \psi)] \to ((\psi \to \xi) \to (\varphi \to \xi))]]]$.

27. Apply MP to steps (25) and (26), $\lambda_{22} \to [(\varphi \to \psi) \to [[(\psi \to \xi) \to (\varphi \to \psi)] \to ((\psi \to \xi) \to (\varphi \to \xi))]]$.

28. Apply MP to steps (22) and (27), $\lambda_{28} := (\varphi \to \psi) \to [[(\psi \to \xi) \to (\varphi \to \psi)] \to ((\psi \to \xi) \to (\varphi \to \xi))]$.

29. By IP2, $\lambda_{28} \to [\lambda_1 \to [(\varphi \to \psi) \to ((\psi \to \xi) \to (\varphi \to \xi))]]$.

30. Apply MP to steps (28) and (29), $\lambda_1 \to [(\varphi \to \psi) \to ((\psi \to \xi) \to (\varphi \to \xi))]$.

31. Apply MP to steps (1) and (30), $(\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$.

*Similar proof for $(\psi \to \xi) \to [(\varphi \to \psi) \to (\varphi \to \xi)]$. □

Proposition 1.7 (Antecedent Switch - AS)**.** Given wff φ, ψ, ξ, then $[\varphi \to (\psi \to \xi)] \to [\psi \to (\varphi \to \xi)]$ holds.

Proof. Suppose $\varphi \to (\psi \to \xi)$ and ψ, we will show that $\varphi \to \xi$. Indeed,

1. Substitute $\alpha := \psi$ and $\beta := \varphi$ into IP1
$$\psi \to (\varphi \to \psi)$$
2. Apply MP to step (1) and the hypothesis ψ
$$\varphi \to \psi$$
3. Substitute $\alpha := \varphi$, $\beta := \psi$, and $\gamma := \xi$ into IP2
$$[\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$$
4. Apply MP to step (3) and the hypothesis $\varphi \to (\psi \to \xi)$.
$$(\varphi \to \psi) \to (\varphi \to \xi)$$
5. Apply MP to steps (2) and (4)
$$\varphi \to \xi$$

Thus, $\varphi \to (\psi \to \xi), \psi \vdash \varphi \to \xi$. But by deduction metatheorem, we get

$$\begin{aligned} \varphi \to (\psi \to \xi) &\vdash \psi \to (\varphi \to \xi) \\ &\vdash [\varphi \to (\psi \to \xi)] \to [\psi \to (\varphi \to \xi)] \end{aligned}$$

□

Proposition 1.8. Let φ, ψ, ξ be wffs. The followings are theorems of propositional logic

1. [DN] Double negation: $\varphi \to \neg\neg\varphi$ and $\neg\neg\varphi \to \varphi$
2. [EXP] Principle of explosion: $\neg\varphi \to (\varphi \to \psi)$
3. [CTP] Contraposition: $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$ and $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$.
4. [MI] Material implication: $\varphi \to [\neg\psi \to \neg(\varphi \to \psi)]$
5. [EXH] Proof by exhaustion: $(\varphi \to \psi) \to [(\neg\varphi \to \psi) \to \psi]$

Proof.

1. For $\neg\neg\varphi \to \varphi$

   (a) Substitute $\alpha := \varphi$ and $\beta := \neg\varphi$ into PNC

   $$(\neg\varphi \to \neg\neg\varphi) \to [(\neg\varphi \to \neg\varphi) \to \varphi]$$

   (b) Substitute $\varphi := \neg\varphi \to \neg\neg\varphi$, $\psi := \neg\varphi \to \neg\varphi$ and $\xi := \varphi$ into AS

   $$[(\neg\varphi \to \neg\neg\varphi) \to ((\neg\varphi \to \neg\varphi) \to \varphi)] \to [(\neg\varphi \to \neg\varphi) \to ((\neg\varphi \to \neg\neg\varphi) \to \varphi)]$$

   (c) Apply MP to steps (a) and (b)

   $$(\neg\varphi \to \neg\varphi) \to [(\neg\varphi \to \neg\neg\varphi) \to \varphi]$$

   (d) Substitute $\varphi := \neg\varphi$ into RX

   $$\neg\varphi \to \neg\varphi$$

   (e) Apply MP to steps (c) and (d)

   $$(\neg\varphi \to \neg\neg\varphi) \to \varphi$$

   (f) Substitute $\alpha := \neg\neg\varphi$ and $\beta := \neg\varphi$ into IP1

   $$\neg\neg\varphi \to (\neg\varphi \to \neg\neg\varphi)$$

   (g) Substitute $\varphi := \neg\neg\varphi$, $\psi := \neg\varphi \to \neg\neg\varphi$ and $\xi := \varphi$ into HS

   $$[\neg\neg\varphi \to (\neg\varphi \to \neg\neg\varphi)] \to [[(\neg\varphi \to \neg\neg\varphi) \to \varphi] \to (\neg\neg\varphi \to \varphi)]$$

   (h) Apply MP to steps (f) and (g)

   $$[(\neg\varphi \to \neg\neg\varphi) \to \varphi] \to (\neg\neg\varphi \to \varphi)$$

   (i) Apply MP to steps (e) and (h)

   $$\neg\neg\varphi \to \varphi$$

   For $\varphi \to \neg\neg\varphi$

   (a) Substitute $\alpha := \neg\neg\varphi$ and $\beta := \varphi$ into PNC

   $$(\neg\neg\neg\varphi \to \neg\varphi) \to [(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi]$$

   (b) Substitute $\varphi := \neg\varphi$ into the previous result

   $$\neg\neg\neg\varphi \to \neg\varphi$$

   (c) Apply MP to steps (a) and (b)

   $$(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi$$

(d) Substitute $\alpha := \varphi$ and $\beta := \neg\neg\neg\varphi$ into IP1

$$\varphi \to (\neg\neg\neg\varphi \to \varphi)$$

(e) Substitute $\varphi := \varphi$, $\psi := \neg\neg\neg\varphi \to \varphi$, and $\xi := \neg\neg\varphi$ into HS

$$[\varphi \to (\neg\neg\neg\varphi \to \varphi)] \to [[(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi] \to (\varphi \to \neg\neg\varphi)]$$

(f) Apply MP to steps (d) and (e)

$$[(\neg\neg\neg\varphi \to \varphi) \to \neg\neg\varphi] \to (\varphi \to \neg\neg\varphi)$$

(g) Apply MP to steps (c) and (f)

$$\varphi \to \neg\neg\varphi$$

2. Suppose $\neg\varphi$ and φ, we will show that ψ

(a) Substitute $\alpha := \varphi$ and $\beta := \neg\psi$ into IP1

$$\varphi \to (\neg\psi \to \varphi)$$

(b) Apply MP to step (a) and the hypothesis φ

$$\neg\psi \to \varphi$$

(c) Substitute $\varphi := \neg\varphi$ into previous result

$$\neg\psi \to \neg\varphi$$

(d) Substitute $\alpha := \psi$ and $\beta := \varphi$ into PNC

$$(\neg\psi \to \neg\varphi) \to [(\neg\psi \to \varphi) \to \psi]$$

(e) Apply MP to steps (c) and (d)

$$(\neg\psi \to \varphi) \to \psi$$

(f) Apply MP to steps (b) and (e)

$$\psi$$

Thus, $\neg\varphi, \varphi \vdash \psi$. Applying deduction metatheorem twice, we get $\vdash \neg\varphi \to (\varphi \to \psi)$.

3. For $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$: suppose $\neg\psi \to \neg\varphi$

(a) Substitute $\alpha := \psi$ and $\beta := \varphi$ into PNC

$$(\neg\psi \to \neg\varphi) \to [(\neg\psi \to \varphi) \to \psi]$$

(b) Apply MP to step (a) and hypothesis $\neg\psi \to \neg\varphi$

$$(\neg\psi \to \varphi) \to \psi$$

(c) Substitute $\alpha := \varphi$ and $\beta := \neg\psi$ into IP1

$$\varphi \to (\neg\psi \to \varphi)$$

(d) Substitute $\varphi := \varphi$, $\psi := \neg\psi \to \varphi$ and $\xi := \psi$ into HS

$$[\varphi \to (\neg\psi \to \varphi)] \to [[(\neg\psi \to \varphi) \to \psi] \to (\varphi \to \psi)]$$

(e) Apply MP to steps (c) and (d)

$$[(\neg\psi \to \varphi) \to \psi] \to (\varphi \to \psi)$$

(f) Apply MP to steps (b) and (e)

$$\varphi \to \psi$$

For $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$: suppose $\varphi \to \psi$

(a) Substitute $\varphi := \neg\neg\varphi$, $\psi := \varphi$, and $\xi := \psi$ into HS

$$(\neg\neg\varphi \to \varphi) \to [(\varphi \to \psi) \to (\neg\neg\varphi \to \psi)]$$

(b) Apply MP to step (a) and DN

$$(\varphi \to \psi) \to (\neg\neg\varphi \to \psi)$$

(c) Apply MP to step (b) and the hypothesis $\varphi \to \psi$

$$\neg\neg\varphi \to \psi$$

(d) Substitute $\varphi := \neg\neg\varphi$, $\psi := \psi$, and $\xi := \neg\neg\psi$ into HS

$$(\neg\neg\varphi \to \psi) \to [(\psi \to \neg\neg\psi) \to (\neg\neg\varphi \to \neg\neg\psi)]$$

(e) Apply MP to steps (c) and (d)

$$(\psi \to \neg\neg\psi) \to (\neg\neg\varphi \to \neg\neg\psi)$$

(f) Apply MP to step (e) and DN

$$\neg\neg\varphi \to \neg\neg\psi$$

(g) Substitute $\varphi := \neg\psi$ and $\psi := \neg\varphi$ into previous result $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$

$$(\neg\neg\varphi \to \neg\neg\psi) \to (\neg\psi \to \neg\varphi)$$

(h) Apply MP to step (f) and (g)

$$\neg\psi \to \neg\varphi$$

4. (a) Apply deduction metatheorem twice to MP

$$\varphi \to [(\varphi \to \psi) \to \psi]$$

(b) Substitute $\varphi := \varphi \to \psi$ and $\psi := \psi$ to CTP

$$[(\varphi \to \psi) \to \psi] \to [\neg\psi \to \neg(\varphi \to \psi)]$$

(c) Substitute $\varphi := \varphi$, $\psi := (\varphi \to \psi) \to \psi$ and $\xi := \neg\psi \to \neg(\varphi \to \psi)$ to HS

$$[\varphi \to ((\varphi \to \psi) \to \psi)] \to [[((\varphi \to \psi) \to \psi) \to (\neg\psi \to \neg(\varphi \to \psi))] \\ \to [\varphi \to (\neg\psi \to \neg(\varphi \to \psi))]]$$

(d) Apply MP to step (a) and (c)

$$[((\varphi \to \psi) \to \psi) \to (\neg\psi \to \neg(\varphi \to \psi))] \to [\varphi \to (\neg\psi \to \neg(\varphi \to \psi))]$$

(e) Apply MP to step (b) and (d)

$$\varphi \to [\neg\psi \to \neg(\varphi \to \psi)]$$

5. Suppose $\varphi \to \psi$ and $\neg\varphi \to \psi$:

(a) Apply MP to CTP and the hypothesis $\varphi \to \psi$

$$\neg\psi \to \neg\varphi$$

(b) Apply MP to CTP and the hypothesis $\neg\varphi \to \psi$

$$\neg\psi \to \neg\neg\varphi$$

(c) Substitute $\alpha := \psi$ and $\beta := \neg\varphi$ into PNC

$$(\neg\psi \to \neg\neg\varphi) \to [(\neg\psi \to \neg\varphi) \to \psi]$$

(d) Apply MP to steps (b) and (c)

$$(\neg\psi \to \neg\varphi) \to \psi$$

(e) Apply MP to steps (a) and (d)

$$\psi$$

We thus get $\varphi \to \psi, \neg\varphi \to \psi \vdash \psi$. Apply the deduction metatheorem twice, we should have $(\varphi \to \psi) \to [(\neg\varphi \to \psi) \to \psi]$.

6. (a) Substitute $\alpha, \beta := \varphi$ into PNC

$$(\neg\varphi \to \neg\varphi) \to [(\neg\varphi \to \varphi) \to \varphi]$$

(b) Apply MP to step (a) and RX

$$(\neg\varphi \to \varphi) \to \varphi$$

□

1.1.4 Truth values

Without a set-theoretic framework, this section is very informal. However, it is crucial in translating the *validity* of a formula into a formal proof of it.

Let F_0 denote the collection of all well-formed formulas of propositional logic. A *valuation* on F_0 is a function $\mathfrak{v} : F_0 \to \{0, 1\}$ which assigns each formula to either 0 (false) or 1 (true).

If Γ is a collection of wffs (i.e. $\Gamma \subseteq F_0$), we then say that $\mathfrak{v}$ *interprets*, or *satisfies*, Γ (denoted as $\mathfrak{v} \vDash \Gamma$) if

1. $\mathfrak{v}(\varphi \to \psi) = 0$ if and only if $\mathfrak{v}(\varphi) = 1$ and $\mathfrak{v}(\psi) = 0$.
2. $\mathfrak{v}(\varphi) = 1 - \mathfrak{v}(\neg\varphi)$. Equivalently,
   (a) $\mathfrak{v}(\varphi) = 0$ if and only if $\mathfrak{v}(\neg\varphi) = 1$.
   (b) $\mathfrak{v}(\varphi) = 1$ if and only if $\mathfrak{v}(\neg\varphi) = 0$.
3. $\mathfrak{v}(\varphi) = 1$ for all φ in Γ.

Remark 1.9. Such $\mathfrak{v}$ may not exist for arbitrary Γ. However, when Γ is empty, we do have a way to construct such valuation. Let A_0 be the collection of atomic formulas (i.e. $A_0 = \Sigma_A \cap F_0$). For each map $v : A_0 \to \{0, 1\}$, we can inductively extend it to another map $\overline{v} : F_0 \to \{0, 1\}$ as follows

1. $\overline{v}(p) = v(p)$ whenever $p \in A_0$.
2. $\overline{v}(\varphi \to \psi) = 0$ if and only if $\overline{v}(\varphi) = 1$ and $\overline{v}(\psi) = 0$.
3. $\overline{v}(\neg\varphi) = 0$ if and only if $\overline{v}(\varphi) = 1$, and vice versa.

In fact, by the principle of induction, if $\mathfrak{v}, \mathfrak{w}$ are valuations which interpret propositional logic (i.e. when Γ is empty) such that they agrees on atomic formulas, then they are the same

$$\mathfrak{v} = \mathfrak{w} \text{ whenever } \mathfrak{v} \mid_{A_0} = \mathfrak{w} \mid_{A_0}$$

If $\mathfrak{v} \vDash \varphi$ whenever $\mathfrak{v} \vDash \Gamma$, then we say that Γ *semantically entails* φ, and write $\Gamma \vDash \varphi$.

Remark 1.10. If there is no such $\mathfrak{v}$ (that interprets Γ), then by convention, we vacuously have $\Gamma \vDash \varphi$ for arbitrary φ. This can be explained intuitively as follows: by an abuse of notation, $\Gamma \vDash \varphi \leftrightarrow (\forall \mathfrak{v} : \mathfrak{v} \vDash \Gamma \to \mathfrak{v} \vDash \varphi)$. So if there is no $\mathfrak{v}$ satisfying Γ, then the precedent is always false, which means the whole implication is vacuously true. By equivalence, $\Gamma \vDash \varphi$.

Lemma 1.11. If $\mathfrak{v}$ interprets Γ, then

1. $\mathfrak{v}(\alpha) = 1$ whenever α is an axiom.
2. If $\mathfrak{v}(\alpha) = \mathfrak{v}(\alpha \to \beta) = 1$, then $\mathfrak{v}(\beta) = 1$.

Proof. Part 1: the following truth tables should verify $\mathfrak{v}(\alpha) = 1$ no matter what values $\mathfrak{v}$ assigns to wff inside each axiom.

α	0	0	1	1
β	0	1	0	1
$\beta \to \alpha$	1	0	1	1
$\alpha \to (\beta \to \alpha)$	1	1	1	1

α	0	0	0	0	1	1	1	1
β	0	0	1	1	0	0	1	1
γ	0	1	0	1	0	1	0	1
$\beta \to \gamma$	1	1	0	1	1	1	0	1
$\alpha \to (\beta \to \gamma)$	1	1	1	1	1	1	0	1
$\alpha \to \beta$	1	1	1	1	0	0	1	1
$\alpha \to \gamma$	1	1	1	1	0	1	0	1
$(\alpha \to \beta) \to (\alpha \to \gamma)$	1	1	1	1	1	1	0	1
$[\alpha \to (\beta \to \gamma)] \to [(\alpha \to \beta) \to (\alpha \to \gamma)]$	1	1	1	1	1	1	1	1

α	0	0	1	1
β	0	1	0	1
$\neg\alpha$	1	1	0	0
$\neg\beta$	1	0	1	0
$\neg\alpha \to \neg\beta$	1	0	1	1
$\neg\alpha \to \beta$	0	1	1	1
$(\neg\alpha \to \beta) \to \alpha$	1	0	1	1
$(\neg\alpha \to \neg\beta) \to [(\neg\alpha \to \beta) \to \alpha]$	1	1	1	1

Part 2: suppose $\mathfrak{v}(\alpha) = \mathfrak{v}(\alpha \to \beta) = 1$, if $\mathfrak{v}(\beta) = 0$ then $\mathfrak{v}(\alpha \to \beta) = 0$ by definition. Therefore, $\mathfrak{v}(\beta) = 1$. □

Theorem 1.12 (Soundness theorem)**.** If $\Gamma \vdash \varphi$, then $\Gamma \vDash \varphi$.

Proof. We do induction on the length of proof of φ from Γ.

1. The base case: when the proof of φ is only φ. Then φ must be either an axiom, or belongs to Γ.
   (a) If φ is an axiom, then $\mathfrak{v}(\varphi) = 1$ by previous Lemma 1.11.
   (b) If $\varphi \in \Gamma$, then $\mathfrak{v}(\varphi) = 1$ also since $\mathfrak{v} \vDash \Gamma$.
2. The inductive case, suppose $(\lambda_1, \ldots, \lambda_n = \varphi)$ is a proof of φ from Γ such that $\mathfrak{v}(\lambda_i) = 1$ for $1 \leq i < n$.
   (a) If φ is an axiom, or belongs to Γ: we get $\mathfrak{v}(\varphi) = 1$ from base case.
   (b) If φ is an application of MP to some λ_i and λ_j for $i, j < n$. WLOG, we suppose λ_j is of the form $\lambda_i \to \varphi$. Then by induction hypothesis, we have $\mathfrak{v}(\lambda_i) = \mathfrak{v}(\lambda_i \to \varphi) = 1$. By Lemma 1.11, we should get $\mathfrak{v}(\varphi) = 1$.

Thus, any $\mathfrak{v}$ that inteprets Γ must also satisfy φ. By definition $\Gamma \vDash \varphi$. $\square$

Remark 1.13.

1. Equivalently (to the soundness theorem), if there exists an interpretation $\mathfrak{v}$ that satisfies Γ yet not φ (i.e. $\mathfrak{v}$ satisfies $\neg\varphi$), then φ is not derivable from Γ.

2. Moreover, if there also exists interpretation $\mathfrak{v}$ that satisfies Γ and φ, then φ is said to be *independent* from Γ (as Γ cannot refute φ).

3. As for the case where every $\mathfrak{v}$ that satisfies Γ does not satisfy φ, then Γ actually refutes φ (i.e. it proves the negation of φ). This is called completeness, and we will prove it when Γ is empty, which is sufficient enough as a foundation of everything else.

1.1.5 Propositional substitution

Let $p_1, p_2, \ldots, p_n$ be the only atomic formulas (or *propositional variables*) that involved in a wff φ. For wff $\theta_1, \ldots, \theta_n$ (which matches the number of variables in φ), we can then define the *substitution* $\varphi[p_1 := \theta_1, \ldots, p_n := \theta_n]$ where we replace every occurrence of p_i in φ by respective θ_i. The *schema* of φ is defined as the collection of all substitutions into φ.

Remark 1.14 (Principle of Uniform Substitution)**.** A recursive definition is given as follows

1. $p[p := \theta]$ is the same as φ.

2. $\varphi[p := \theta] \to \psi[p := \theta]$ is the same as $(\varphi \to \psi)[p := \theta]$.

3. $\neg(\varphi[p := \theta])$ is the same as $(\neg\varphi)[p := \theta]$.

Example 1.15.

1. The schema of $p \to (q \to p)$ is all $\alpha \to (\beta \to \alpha)$ for arbitrary wff α, β.

2. The schema of $[p \to (q \to r)] \to [(p \to q) \to (p \to r)]$ is all $[\alpha \to (\beta \to \gamma)] \to [(\alpha \to \beta) \to (\alpha \to \gamma)]$ for arbitrary wff α, β, γ.

3. The schema of $(\neg p \to \neg q) \to [(\neg p \to q) \to p]$ is all $(\neg\alpha \to \neg\beta) \to [(\neg\alpha \to \beta) \to \alpha]$ for arbitrary wff α, β.

Theorem 1.16 (Uniformness Theorem)**.** Suppose φ is a theorem (of propositional logic). Then each formula in the schema of φ is also a theorem.

Proof. Assume p_i $(1 \leq i \leq n)$ are propositional variables in φ. Let $\theta_1, \theta_2, \ldots, \theta_n$ be wffs that are to substitute into φ. Since φ is a theorem, there is a proof $(\lambda_1, \lambda_2, \ldots, \lambda_m = \varphi)$ of it. We will show that the sequence $(\lambda_j[p_i := \theta_i])_{1 \leq j \leq m}$ is a proof for $\varphi[p_i := \theta_i]$. By definition, each λ_j $(j = \overline{1, m})$ is either

1. An axiom: we will show that $\lambda_j[p_i := \theta_i]$ is also an axiom if λ_j is of the form $\alpha \to (\beta \to \alpha)$, and the other 2 cases follows similarly. By Remark 1.14, $\lambda_j[p_i := \theta_i]$ is the same as $(\alpha \to (\beta \to \alpha))[p_i := \theta_i]$, or $\mu \to (\nu \to \mu)$ where $\mu := \alpha[p_i := \theta_i]$ and $\nu := \beta[p_i := \theta_i]$. Note that $\mu \to (\nu \to \mu)$ is itself an axiom, so we conclude that $\lambda_j[p_i := \theta_i]$ is an axiom.

2. An application of MP to λ_k and $\lambda_l := \lambda_k \to \lambda_j$: then $\lambda_j[p_i := \theta_i]$ is an application of MP to $\lambda_k[p_i := \theta_i]$ and $\lambda_k[p_i := \theta_i] \to \lambda_j[p_i := \theta_i]$, where the latter is the same as $\lambda_l[p_i := \theta_i]$ (due to Remark 1.14). In other words, $\lambda_j[p_i := \theta_i]$ is an application of MP to $\lambda_k[p_i := \theta_i]$ and $\lambda_l[p_i := \theta_i]$

In summary, each $\eta_j := \lambda_j[p_i := \theta_i]$ is either

1. An axiom of propositional logic

2. An application of MP to some η_k and η_l $(k, l < j)$.

Note that η_m is the same as $\varphi[p_i := \theta_i]$, so by definition, $(\eta_1, \ldots, \eta_m)$ is a proof of $\varphi[p_i := \theta_i]$ □

Remark 1.17. This provides a uniform proof for each formula in the schema of φ, assuming φ is a theorem (i.e. there exists a proof of it).

1.1.6 Completeness

The converse of *soundness*, i.e. $\Gamma \vdash \varphi$ whenever $\Gamma \vDash \varphi$, is called the *completeness* property, and in propositional logic, this does hold. In fact, this holds in arbitrary logic (with the same language as propositional logic) whenenever

1. The inference rule modus ponens holds

2. The followings statements (or wffs) are provable, given wffs φ, ψ, ξ:

   (a) $\varphi \to (\psi \to \varphi)$

   (b) $[\varphi \to (\psi \to \xi)] \to [(\varphi \to \psi) \to (\varphi \to \xi)]$

   (c) $\varphi \to \varphi$

   (d) $\varphi \to \neg\neg\varphi$

   (e) $\neg\varphi \to (\varphi \to \psi)$

   (f) $\varphi \to [\neg\psi \to \neg(\varphi \to \psi)]$

   (g) $(\varphi \to \psi) \to [(\neg\varphi \to \psi) \to \psi]$

However, *constructively*, it is only possible to construct a proof for $\Gamma \vdash \varphi$ from $\Gamma \vDash \varphi$ when there are only finitely many atomic formulas (or *propositional variables*) involved in formulas belong to Γ. We will only deal with the case when Γ is empty, as the remaining ones will be discussed in more detail much later.

Remark 1.18. Those statements are the result of

1. The axioms IP1, IP2.
2. Reflexivity RX
3. Proposition 1.8

Lemma 1.19 (Kalmár lemma)**.** Suppose $\mathfrak{v} : F_0 \to \{0,1\}$ is an interpretion of propositional logic, and $p_1, p_2, \ldots, p_n$ are atomic wffs involved in a wff φ. We then construct the following formulas $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n}, \overline{\varphi}$

1. For each p_i, if $\mathfrak{v}(p_i) = 1$ then $\overline{p_i} := p_i$. Otherwise, $\overline{p_i} := \neg p_i$.
2. If $\mathfrak{v}(\varphi) = 1$, then $\overline{\varphi} := \varphi$. Otherwise, $\overline{\varphi} := \neg\varphi$.

It is then true that $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$.

Proof. We do an induction on the number of connectives in φ

1. Base case: if φ does not have any connective, then it is just an atomic formula by (inductive) definition. Thus $\overline{\varphi}$ is the same as φ, and so $\overline{\varphi} \vdash \varphi$ trivially (as the proof only has one step $\overline{\varphi}$).
2. Inductive case, when $\varphi := \neg\psi$: then ψ and φ have the same propositional variables (which are $p_1, p_2, \ldots, p_n$). But ψ has one less connectives, so by induction

   $$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi}$$

   We divide into 2 cases

   (a) If $\mathfrak{v}(\psi) = 1$, then $\overline{\psi} = \psi$, $\mathfrak{v}(\varphi) = \mathfrak{v}(\neg\psi) = 0$, and $\overline{\varphi} = \neg\varphi = \neg\neg\psi$. By DN, $\vdash \psi \to \neg\neg\psi$, so we should get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi} \to \overline{\varphi}$.

   (b) If $\mathfrak{v}(\psi) = 0$, then $\overline{\psi} = \neg\psi$, $\mathfrak{v}(\varphi) = \mathfrak{v}(\neg\psi) = 1$, and $\overline{\varphi} = \varphi = \neg\psi$. By RX, $\vdash \neg\psi \to \neg\psi$, so we should get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi} \to \overline{\varphi}$.

   In either case, $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi} \to \overline{\varphi}$. So by applying MP, we obtain $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$.
3. Inductive case, when $\varphi := \psi \to \xi$: let $q_1, q_2, \ldots, q_m$ be the variables involved in ψ, and $r_1, r_2, \ldots, r_p$ be the variables involved in ξ. By assumption, these variables must be from $p_1, p_2, \ldots, p_n$. But since

   $$\overline{q_1}, \overline{q_2}, \ldots, \overline{q_m} \vdash \overline{\psi}$$
   $$\overline{r_1}, \overline{r_2}, \ldots, \overline{r_p} \vdash \overline{\xi}$$

   by induction hypothesis (as the connectives of ψ and ξ together are still one less than those of φ), we get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi}, \overline{\xi}$ by monotonicity.

   To show this, as a result, proves $\overline{\varphi}$, we divide into the following cases

(a) If $\mathfrak{v}(\psi) = 0$: then $\overline{\psi} = \neg\psi$, and $\mathfrak{v}(\varphi) = 1$, $\overline{\varphi} = \varphi = \psi \to \xi$. From EXP, we get

$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \neg\psi \to (\psi \to \xi) \text{ or}$$
$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi} \to \overline{\varphi}$$

So we get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$ with an application of MP.

(b) If $\mathfrak{v}(\psi) = 1$, $\mathfrak{v}(\xi) = 0$: then $\overline{\psi} = \psi$, $\overline{\xi} = \neg\xi$, and $\mathfrak{v}(\varphi) = 0$, $\overline{\varphi} = \neg\varphi = \neg(\psi \to \xi)$. But by MI,

$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \psi \to [\neg\xi \to \neg(\psi \to \xi)] \text{ or}$$
$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\psi} \to (\overline{\xi} \to \overline{\varphi})$$

So applying MP twice to $\overline{\psi} \to (\overline{\xi} \to \overline{\varphi})$ and $\overline{\psi}, \overline{\xi}$ (from induction hypothesis), and we will get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$.

(c) If $\mathfrak{v}(\psi) = 1$, $\mathfrak{v}(\xi) = 1$: then $\overline{\psi} = \psi$, $\overline{\xi} = \xi$, and $\mathfrak{v}(\varphi) = 1$, $\overline{\varphi} = \varphi = \psi \to \xi$. But by IP1,

$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \xi \to (\psi \to \xi) \text{ or}$$
$$\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\xi} \to \overline{\varphi}$$

So by applying MP once, we should get $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$.

□

Remark 1.20. To illustrate the process of proof construction, we take $(\neg\varphi \to \varphi) \to \varphi$ (which holds in every interpretation) as an example to show that

$$\varphi \vdash (\neg\varphi \to \varphi) \to \varphi \text{ (when } \mathfrak{v}(\varphi) = 1)$$

Indeed

(a) By DN

$$\varphi \to \neg\neg\varphi$$

(b) Apply MP to step (a) and the hypothesis φ

$$\neg\neg\varphi$$

(c) Since $\mathfrak{v}(\varphi) = 1$, $\mathfrak{v}(\neg\varphi) = 0$. So by EXP

$$\neg\neg\varphi \to (\neg\varphi \to \varphi)$$

(d) Apply MP to steps (b) and (c)

$$\neg\varphi \to \varphi$$

(e) Since $\mathfrak{v}(\neg\varphi \to \varphi) = 1$, we now use IP1 to get

$$\varphi \to [(\neg\varphi \to \varphi) \to \varphi]$$

(f) We then apply MP to step (e) and the hypothesis φ

$$(\neg\varphi \to \varphi) \to \varphi$$

Note that there may be redundancy (e.g. steps (a) to (d)), but this is how the proof is constructed based on the lemma.

Theorem 1.21 (Completeness theorem)**.** Given a wff φ, then $\vDash \varphi$ implies $\vdash \varphi$.

Proof. Let $p_1, p_2, \ldots, p_n$ be propositional variables in φ, $\mathfrak{v}$ be any valuation. We will denote $\overline{\theta}^{\mathfrak{v}}$ to be either θ or $\neg\theta$ based on the value of θ under $\mathfrak{v}$. Since $\vDash \varphi$, we get $\overline{\varphi}^{\mathfrak{v}} = \varphi$. By Lemma 1.19

$$\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_n}^{\mathfrak{v}} \vdash \varphi$$

We will show that $\vdash \varphi$ by induction on the number of variables in φ.

1. Base case: $\overline{p_1}^{\mathfrak{v}} \vdash \varphi$ (for any $\mathfrak{v}$). Since a valuation is uniquely specified by its value on the atomic formulas (Remark 1.9), let there exists 2 valuation $\mathfrak{v}_0$ and $\mathfrak{v}_1$ such that $\mathfrak{v}_0(p_1) = 0$ and $\mathfrak{v}_1(p_1) = 1$. Hence,

$$p_1 \vdash \varphi$$
$$\neg p_1 \vdash \varphi$$

   So by deduction metatheorem, $\vdash p_1 \to \varphi$ and $\vdash \neg p_1 \to \varphi$. By EXH

$$\vdash (p_1 \to \varphi) \to [(\neg p_1 \to \varphi) \to \varphi]$$

   We then apply MP twice to get $\vdash \varphi$.

2. Induction case: suppose we can construct a proof φ whenever $\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_n}^{\mathfrak{v}} \vdash \varphi$ for all valuation $\mathfrak{v}$. We will show that there is also a proof of φ whenever $\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_{n+1}}^{\mathfrak{v}} \vdash \varphi$ for all valuation $\mathfrak{v}$.

   This is similar to the base case. Let $\mathfrak{v}$ be any valuation, and denote $\mathfrak{w}_0$ and $\mathfrak{w}_1$ such that $\mathfrak{v}(p_i) = \mathfrak{w}_0(p_i) = \mathfrak{w}_1(p_i)$ for $1 \leq i \leq n$, yet $\mathfrak{w}_0(p_{n+1}) = 0$ and $\mathfrak{w}_1(p_{n+1}) = 1$. We thus have

$$\overline{p_1}^{\mathfrak{w}_0}, \overline{p_2}^{\mathfrak{w}_0}, \ldots, \overline{p_n}^{\mathfrak{w}_0}, p_{n+1} \vdash \varphi$$
$$\overline{p_1}^{\mathfrak{w}_1}, \overline{p_2}^{\mathfrak{w}_1}, \ldots, \overline{p_n}^{\mathfrak{w}_1}, \neg p_{n+1} \vdash \varphi$$

   By deduction theorem, we get $\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_n}^{\mathfrak{v}} \vdash p_{n+1} \to \varphi, \neg p_{n+1} \to \varphi$ (note that $\mathfrak{v}, \mathfrak{w}_0, \mathfrak{w}_1$ all agree on p_i for $i = \overline{1, n}$). Yet

$$\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_n}^{\mathfrak{v}} \vdash (p_{n+1} \to \varphi) \to [(\neg p_{n+1} \to \varphi) \to \varphi]$$

   So we can deduce $\overline{p_1}^{\mathfrak{v}}, \overline{p_2}^{\mathfrak{v}}, \ldots, \overline{p_n}^{\mathfrak{v}} \vdash \varphi$ by applying MP. As $\mathfrak{v}$ varies, we conclude that $\vdash \varphi$ by induction hypothesis.

□

Remark 1.22. The completeness property, with deduction metatheorem, can be extended with finitely many hypotheses, i.e.

$$\text{If } \lambda_1, \lambda_2, \ldots, \lambda_n \vDash \varphi \text{ then } \lambda_1, \lambda_2, \ldots, \lambda_n \vdash \varphi.$$

Beyond that, for arbitrary collection of hypothesis Γ, then a proof of φ will not be constructive (even when $\Gamma \vDash \varphi$).

Example 1.23. Note that the proof of completeness theorem does not use CTP, but only

1. RX, DN, EXP, MI, EXH, IP1.

2. MP as an inference rule.

3. Deduction metatheorem.

So we will illustrate the proof of CTP using just these basic principles. In particular, we shall prove $\theta_0 := [(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)]$.

1. By Lemma 1.19

$$\begin{aligned} \varphi, \psi &\vdash \theta_0 \\ \varphi, \neg\psi &\vdash \theta_0 \\ \neg\varphi, \psi &\vdash \theta_0 \\ \neg\varphi, \neg\psi &\vdash \theta_0 \end{aligned}$$

2. By deduction metatheorem

$$\begin{aligned} \varphi &\vdash \psi \to \theta_0 \\ \varphi &\vdash \neg\psi \to \theta_0 \end{aligned}$$

   But since $\vdash (\psi \to \theta_0) \to [(\neg\psi \to \theta_0) \to \theta_0]$ (EXH), we get $\varphi \vdash \theta_0$ by applying MP twice. Similarly, $\neg\varphi \vdash \theta_0$.

3. Again, by deduction metatheorem

$$\begin{aligned} &\vdash \varphi \to \theta_0 \\ &\vdash \neg\varphi \to \theta_0 \end{aligned}$$

   With the same argument as in (2), we finally obtain $\vdash \theta_0$.

1.1.7 Consistency

Given a collection Γ of wffs, we say that Γ is *(syntactically) inconsistent* if there exists a wff δ such that $\Gamma \vdash \delta$ and $\Gamma \vdash \neg\delta$. More specifically, there exists a wff δ, a proof of it, and a proof of the negation of it from Γ.

Theorem 1.24.

1. Γ is inconsistent if and only if $\Gamma \vdash \varphi$ for all wff φ (i.e. there is a proof of each wff).
2. $\Gamma \vdash \varphi$ if and only if $\Gamma \cup \{\neg\varphi\}$ is inconsistent.

Proof.
Part 1: suppose Γ is inconsistent, and let δ be a wff such that $\Gamma \vdash \delta$ and $\Gamma \vdash \neg\delta$. Then apply MP twice to EXP (and the previous assumptions), we get $\Gamma \vdash \varphi$ for any wff φ. Vice versa, if $\Gamma \vdash \varphi$ for any wff φ, then fix a formula δ, and we will get both $\Gamma \vdash \delta$ and $\Gamma \vdash \neg\delta$.
Part 2: suppose $\Gamma \vdash \varphi$, then $\Gamma \cup \{\neg\varphi\}$ is inconsistent by definition. Vice versa, if $\Gamma \cup \{\neg\varphi\}$, then there exists some formula δ such that $\Gamma, \neg\varphi \vdash \delta$ and $\Gamma, \neg\varphi \vdash \neg\delta$.

1. By deduction theorem, we get $\Gamma \vdash \neg\varphi \to \delta$ and $\Gamma \vdash \neg\varphi \to \neg\delta$.
2. Using truth table, we get a schema $\vdash (\neg\alpha \to \beta) \to [(\neg\alpha \to \neg\beta) \to \beta]$ (i.e. reductio ad impossibile).
3. Finally, we apply MP to arrive at $\Gamma \vdash \varphi$.

□

Remark 1.25. The above theorem provides a general way to produce a proof of φ from Γ whenever $\Gamma \cup \{\neg\varphi\}$ is inconsistent. More precisely, from a proof of $\Gamma, \neg\varphi \vdash \delta$ and a proof of $\Gamma, \neg\varphi \vdash \neg\delta$ (for some fixed δ), we can algorithmically get a proof of $\Gamma \vdash \varphi$ (regardless of whether Γ is inconsistent or not).

1.1.8 Equivalence - Leibniz rule

Given 2 wffs φ and ψ, if $\Gamma \vdash \varphi \to \psi$ and $\Gamma \vdash \psi \to \varphi$, then we say that φ and ψ are *propositionally equivalent* under Γ, and write $\varphi \equiv \psi \pmod{\Gamma}$. If Γ is empty (i.e. $\varphi \to \psi$ and $\psi \to \varphi$ will be theorems of propositional logic), then we can simply write $\varphi \equiv \psi$.

Remark 1.26.

1. If $\varphi \equiv \psi \pmod{\Gamma}$ and $\Gamma \subseteq \Delta$, then $\varphi \equiv \psi \pmod{\Delta}$. In particular, $\varphi \equiv \varphi \pmod{\Gamma}$ for any collection Γ of formulas.
2. If $\Gamma \vdash \varphi$, and $\varphi \equiv \psi \pmod{\Gamma}$, then $\Gamma \vdash \psi$.
3. If $\Gamma \vdash \varphi$ and $\Gamma \vdash \psi$, then $\varphi \equiv \psi \pmod{\Gamma}$.
4. If $\mathfrak{v} \vDash \Gamma$ and $\varphi \equiv \psi \pmod{\Gamma}$, then $\mathfrak{v}(\varphi) = \mathfrak{v}(\psi)$.

Proof. Part 1: since $\Gamma \subseteq \Delta$ and $\Gamma \vdash \varphi \to \psi$, we get $\Delta \vdash \varphi \to \psi$ (Remark 1.2). Similarly, $\Delta \vdash \psi \to \varphi$, so by definition $\varphi \equiv \psi \pmod{\Delta}$.
Part 2: since $\Gamma \vdash \varphi$ and $\Gamma\varphi \to \psi$, we get $\Gamma \vdash \psi$ by MP.
Part 3: from IP1, we get $\Gamma\psi \to (\varphi \to \psi)$, so apply MP to it and $\Gamma \vdash \psi$, we get $\Gamma \vdash \varphi \to \psi$. Similarly, $\Gamma \vdash \psi \to \varphi$, so by definition, $\varphi \equiv \psi \pmod{\Gamma}$.
Part 4: since $\Gamma \vdash \varphi \to \psi$, $\Gamma \vDash \varphi \to \psi$ by soundness theorem. Thus, $\mathfrak{v}(\varphi \to \psi) = 1$ by definition (whenever $\mathfrak{v}$ interprets Γ). Similarly, $\mathfrak{v}(\psi \to \varphi) = 1$. Now, suppose $\mathfrak{v}(\varphi) \neq \mathfrak{v}(\psi)$. WLOG, set $\mathfrak{v}(\varphi) = 1$ and $\mathfrak{v}(\psi) = 0$ (since there are only 2 values we can assign each formula to). But then $\mathfrak{v}(\varphi \to \psi) = 0$, a contradiction. We conclude that $\mathfrak{v}(\varphi) = \mathfrak{v}(\psi)$. □

Example 1.27.

1. Double Negation: $\varphi \equiv \neg\neg\varphi$
2. Contraposition: $\varphi \to \psi \equiv \neg\psi \to \neg\varphi$
3. Antecedent Swap: $\varphi \to (\psi \to \xi) \equiv \psi \to (\varphi \to \xi)$

Proof.
Part 1: a result of DN.
Part 2: a result of CTP.
Part 3: a result of AS. □

Theorem 1.28 (Leibniz rule of substitution - LS)**.** Let φ, ψ be wffs. We then have

1. $\varphi \equiv \psi \pmod{\Gamma}$ if and only if $\varphi \to \xi \equiv \psi \to \xi \pmod{\Gamma}$ for any wff ξ.
2. $\varphi \equiv \psi \pmod{\Gamma}$ if and only if $\xi \to \varphi \equiv \xi \to \psi \pmod{\Gamma}$ for any wff ξ.
3. $\varphi \equiv \psi \pmod{\Gamma}$ if and only if $\neg\varphi \equiv \neg\psi \pmod{\Gamma}$.

Proof. Part 1: suppose $\varphi \equiv \psi \pmod{\Gamma}$. In particular, $\Gamma \vdash \varphi \to \psi$. By HS and Remark 1.2, we have $\Gamma \vdash (\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$. Apply MP to the previous 2 results, we get $\Gamma \vdash (\psi \to \xi) \to (\varphi \to \xi)$. Similarly, $\Gamma \vdash (\varphi \to \xi) \to (\psi \to \xi)$, and so $\varphi \to \xi \equiv \psi \to \xi \pmod{\Gamma}$.

Vice versa, if $\varphi \to \xi \equiv \psi \to \xi \pmod{\Gamma}$ for any wff ξ. Then we let $\xi := \psi$ to get $\varphi \to \psi \equiv \psi \to \psi \pmod{\Gamma}$. But $\Gamma \vdash \psi \to \psi$ by RX and Remark 1.2, so we must have $\Gamma \vdash \varphi \to \psi$ (Theorem 1.28). Similarly, $\Gamma \vdash \psi \to \varphi$ (by setting $\xi := \varphi$), so we must have $\varphi \equiv \psi \pmod{\Gamma}$.
Part 2: same as part (1).
Part 3: suppose $\varphi \equiv \psi \pmod{\Gamma}$. By DN, we get $\Gamma \vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$ and $\Gamma \vdash (\psi \to \varphi) \to (\neg\varphi \to \neg\psi)$. Then apply MP, we should get $\Gamma \vdash \neg\psi \to \neg\varphi$ and $\Gamma \vdash \neg\varphi \to \neg\psi$, or equivalently, $\neg\varphi \equiv \neg\psi \pmod{\Gamma}$.

Vice versa, suppose $\neg\varphi \equiv \neg\psi \pmod{\Gamma}$. With the same DN (and some slice of MP), it is then true that $\varphi \equiv \psi \pmod{\Gamma}$. □

Remark 1.29. A significance of this rule is that we can replace φ with its double negation $\neg\neg\varphi$ at any point in the formula by repeated application of LS, without changing syntactical meaning because of MP.

Proposition 1.30 (Properties of Equivalence)**.**

1. Reflexivity: $\varphi \equiv \varphi \pmod{\Gamma}$
2. Symmetry: if $\varphi \equiv \psi \pmod{\Gamma}$, then $\psi \equiv \varphi \pmod{\Gamma}$.
3. Transitivity: if $\varphi \equiv \psi \pmod{\Gamma}$ and $\psi \equiv \xi \pmod{\Gamma}$, then $\varphi \equiv \xi \pmod{\Gamma}$.

Proof. Part 1: this follows from RX and Remark 1.2.
Part 2: this follows vacuously from the symmetry in the definition of $\varphi \equiv \psi \pmod{\Gamma}$.
Part 3: suppose $\varphi \equiv \psi \pmod{\Gamma}$ and $\psi \equiv \xi \pmod{\Gamma}$. In particular, $\Gamma \vdash \varphi \to \psi$ and $\Gamma \vdash \psi \to \xi$. But by HS, $\Gamma \vdash (\varphi \to \psi) \to [(\psi \to \xi) \to (\varphi \to \xi)]$. We then apply MP twice in order to get $\Gamma \vdash \varphi \to \xi$. By symmetry, we also get $\Gamma \vdash \xi \to \varphi$ similarly, so $\varphi \equiv \xi \pmod{\Gamma}$. □

1.1.9 Additional Connectives

Given wffs φ and ψ, we can define a few more logical connectives

1. Disjunction: $\varphi \vee \psi := \neg\varphi \to \psi$
2. Conjunction: $\varphi \wedge \psi := \neg(\varphi \to \neg\psi)$
3. Biconditional: $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi) \equiv \neg[(\varphi \to \psi) \to \neg(\psi \to \varphi)]$

The following propositions are left as an exercise (since we will not even refer to them explicitly most of the time, as long as they are tautological).

Proposition 1.31 (Properties of logical connectives)**.** Let 1 denote any theorem in (propositional logic), and $0 \equiv \neg 1$ (note this does not depends on the choice of theorem because of Example 1.27)

1. De Morgan law
   (a) $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi$
   (b) $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi$
2. Commutativity
$$\varphi \wedge \psi \equiv \psi \wedge \varphi$$
$$\varphi \vee \psi \equiv \psi \vee \varphi$$
3. Associativity
$$(\varphi \wedge \psi) \wedge \xi \equiv \varphi \wedge (\psi \wedge \xi)$$
$$(\varphi \vee \psi) \vee \xi \equiv \varphi \vee (\psi \vee \xi)$$

4. Identity: $\varphi \wedge 1 \equiv \varphi \vee 0 \equiv 1 \to \varphi \equiv \varphi$
5. Idempotent: $\varphi \wedge \varphi \equiv \varphi \vee \varphi \equiv \varphi$
6. Domination

$$\varphi \wedge 0 \equiv 0$$
$$\varphi \vee 1 \equiv 1$$
$$\varphi \to 1 \equiv 0 \to \varphi \equiv 1$$

7. Distributivity

$$(\varphi \wedge \psi) \vee \xi \equiv (\varphi \vee \xi) \wedge (\psi \vee \xi)$$
$$(\varphi \vee \psi) \wedge \xi \equiv (\varphi \wedge \xi) \vee (\psi \wedge \xi)$$

8. Absorption: $\varphi \wedge (\varphi \vee \psi) \equiv \varphi \vee (\varphi \wedge \psi) \equiv \varphi$
9. Complementary

$$\varphi \wedge \neg\varphi \equiv 0$$
$$\varphi \vee \neg\varphi \equiv 1$$

10. Material implication: $\varphi \to \psi \equiv \neg\varphi \vee \psi$
11. Biconditionality: $\varphi \leftrightarrow \psi \equiv (\varphi \to \psi) \wedge (\psi \to \varphi) \equiv (\varphi \wedge \psi) \vee (\neg\varphi \wedge \neg\psi)$
12. Implicative Negation: $\varphi \to 0 \equiv \neg\varphi$
13. Left self-distributivity: $\varphi \to (\psi \to \xi) \equiv (\varphi \to \psi) \to (\varphi \to \xi)$
14. Conditional distributivity

$$(\varphi \wedge \psi) \to \xi \equiv (\varphi \to \xi) \vee (\psi \to \xi)$$
$$(\varphi \vee \psi) \to \xi \equiv (\varphi \to \xi) \wedge (\psi \to \xi)$$
$$\varphi \to (\psi \wedge \xi) \equiv (\varphi \to \psi) \wedge (\varphi \to \xi)$$
$$\varphi \to (\psi \vee \xi) \equiv (\varphi \to \psi) \vee (\varphi \to \xi)$$

Note: many authors, especially logicians, use $\bot$ for 0 and $\top$ for 1 instead.

Remark 1.32. Since associativity holds for both conjunction and disjunction, we usually remove parentheses and simply write $\varphi \wedge \psi \wedge \xi$ (and $\varphi \vee \psi \vee \xi$) without any ambiguity. There is also a general rule (albeit we will not use it liberally) for implication: we interpret $\varphi_1 \to \varphi_2 \to \cdots \to \varphi_n$ as $\varphi_1 \to (\varphi_2 \to (\cdots \to \varphi_n))$. This is called *right-associativity*.

Proposition 1.33 (Laws of logic)**.**

1. Laws of identity: $\varphi \equiv \varphi$

2. Law of excluded middle: $\varphi \vee \neg\varphi$
3. Law of non-contradiction: $\neg(\varphi \wedge \neg\varphi)$
4. Principle of explosion: $(\varphi \wedge \neg\varphi) \to \psi$
5. Disjunctive introduction: $\varphi \to (\varphi \vee \psi)$
   Disjunctive syllogism: $\neg\varphi \to ((\varphi \vee \psi) \to \psi)$
6. Conjunctive simplification: $(\varphi \wedge \psi) \to \varphi$
   Conjunctive introduction: $\varphi \to (\psi \to (\varphi \wedge \psi))$
7. Reductio ad absurdum: $[(\varphi \to \psi) \wedge (\varphi \to \neg\psi)] \to \neg\varphi$
8. Modus ponens: $[(\varphi \to \psi) \wedge \varphi] \to \psi$
9. Modus tollens: $[(\varphi \to \psi) \wedge \neg\psi] \to \neg\varphi$
10. Hypothetical syllogism: $[(\varphi \to \psi) \wedge (\psi \to \xi)] \to (\varphi \to \xi)$
11. Disjunctive syllogism: $[(\varphi \vee \psi) \wedge \neg\varphi] \to \psi$
12. Resolution: $[(\varphi \vee \psi) \wedge (\neg\varphi \vee \xi)] \to (\psi \vee \xi)$
13. Exportation: $(\varphi \wedge \psi) \to \xi \equiv \varphi \to (\psi \to \xi)$
14. Proof by case: $[(\varphi \vee \psi) \wedge (\varphi \to \xi) \wedge (\psi \to \xi)] \to \xi$
15. Proof by contradiction: $[\neg\varphi \to (\psi \wedge \neg\psi)] \to \varphi$
16. Gentzen's cut-elimination:

$$[(\varphi \to (\psi \wedge \xi)) \wedge ((\psi \wedge \chi) \to \gamma)] \to [(\varphi \wedge \chi) \to (\xi \wedge \gamma)]$$

1.2 Predicate Logic

List of preliminary references:

1. Herbert B. Enderton. *A Mathematical Introduction to Logic*, 2nd edition (5 January 2001). Academic Press.

1.2.1 Language

The symbols Σ of predicate logic consists of

1. The countably infinite collection of *variables* Σ_V: for examples, $x_1, x_2, \ldots$
2. The collection of letters Σ_A, with each of them is associated with a counting number $n \geq 1$ called *arity*, and each counting number is associated with countably infinitely many letters: say $\varphi_1, \varphi_2, \ldots$.
3. The collection of parentheses Σ_O
4. The collection of logical symbols $\Sigma_L = \{\neg, \to, \forall\}$

The language L of predicate logic, consisting of well-formed formulas, is recursively constructed as follows

1. If $\alpha \in \Sigma_A$ with arity n, and $x_1, x_2, \ldots, x_n$ are variables (possibly some of them are the same), then $\alpha(x_1, x_2, \ldots, x_n)$ is a wff. Such wff is called *atomic*.
2. If φ is a wff, and x is a variable, then $\forall x(\varphi)$ is a wff.
3. If φ is a wff, then $\neg(\varphi)$ is a wff.
4. If φ, ψ are wffs, then $(\varphi) \to (\psi)$ is a wff.

Formulas that are not atomic (i.e. not of the form $\alpha(x_1, x_2, \ldots, x_n)$ where $\alpha \in \Sigma_A$ and $x_i \in \Sigma_V$) are called *compound*.
Note: as usual, we can omit some parentheses (e.g. multiple $\neg$ and $\forall$ stacked consecutively) to reduce the accumulation of them.

1.2.2 Free Variables

Along with each wff φ, there are an additional collection associated with φ, $FV(\varphi)$, which contains *free variables* in φ. It is defined inductively as follows

1. If $\varphi := \alpha(x_1, x_2, \ldots, x_n)$ (where $\alpha \in \Sigma_A$), then $FV(\varphi) = \{x_1, x_2, \ldots, x_n\}$.
2. If $\varphi := \forall x(\lambda)$, then $FV(\varphi) = FV(\lambda) \setminus \{x\}$ (i.e. the free variables of φ are those of λ except x).

3. If $\varphi := \lambda \to \kappa$, then $FV(\varphi) = FV(\lambda) \cup FV(\kappa)$ (i.e. the free variables of φ are exactly the ones of either λ or κ).

4. If $\varphi := \neg\lambda$, then $FV(\varphi) = FV(\lambda)$.

From there, the *arity* of a general expression φ is defined as the cardinality of $FV(\varphi)$. In other words, the arity of φ is exactly the number of free variables inside φ.
Note: this concept of arity for each wff is very similar yet different to the arity of each symbol inside Σ_A (the symbol will become predicate/wff once we apply the correct number of variables to it!). They are the same, however, when considering atomic formulas.

Example 1.34. Suppose $\alpha, \beta, \lambda \in \sigma$ (we don't designate the arity and let the context sorts them out)

1. $\varphi := \alpha(x, x, z) \wedge \forall w(\beta(x, y, w, w))$. The free variables are x, y, z, so the arity of φ is 3. Once the free variables are identified, we usually write φ as $\varphi(x, y, z)$ for a clear view of what are the (free) variables that determines the truth value of φ (the order of appearance may vary, but one needs to include all free variables).

$$\varphi(x, y, z) := \alpha(x, x, z) \wedge \forall w(\beta(x, y, w, w))$$

2. $\varphi := \forall x(\alpha(x)) \wedge \forall x(\beta(x, y)) \wedge \forall z \forall y(\lambda(y, z, x))$. Even though we use the same variable for different context/scope, following the inductive rules above, it should be clear that $FV(\varphi) = \{x, y\}$ and $BV(\varphi) = \varnothing$.

$$\varphi(x, y) := \forall x(\alpha(x)) \wedge \forall x(\beta(x, y)) \wedge \forall z \forall y(\lambda(y, z, x))$$

3. $\varphi := \forall y \forall z \forall x \forall w(\alpha(x, z, t, w, s, y))$. Here, $FV(\varphi) = \{t, s\}$ and $BV(\varphi) = \{x, y, z, w\}$

$$\varphi(t, s) := \forall y \forall z \forall x \forall w(\alpha(x, z, t, w, s, y))$$

If a wff has arity of 0 (i.e. no free variable), we call it a *sentence*.

1.2.3 Substitution

The *substitution* of a variable y of φ (which may not even occur in the formula) for t, written $\varphi[y := t]$, is defined inductively as follows

1. If $\varphi := \alpha(x_1, \ldots, x_i, y, x_{i+1}, \ldots, x_n)$ is an atomic predicate, then $\varphi[y := t] := \alpha(x_1, \ldots, x_i, t, x_{i+1}, \ldots, x_n)$. Otherwise, if $\varphi := \alpha(x_1, \ldots, x_n)$ such that y is distinct from x_i, then $\varphi[y := t] := \varphi$ simply.

2. If $\varphi := \forall x(\lambda)$, then

   (a) $\varphi[y := t] := \varphi$, given $y = x$.

   (b) $\varphi[y := t] := \forall x(\lambda[y := t])$, given $y \neq x$.

3. If $\varphi := \lambda \to \kappa$, then $\varphi[y := t] := \lambda[y := t] \to \kappa[y := t]$.
4. If $\varphi := \neg\lambda$, then $\varphi[y := t] := \neg(\lambda[y := t])$.

Remark 1.35.

1. $\varphi[y := t]$ is always a well-formed formula.
2. $\varphi[y := y] = \varphi$, i.e. the substitution of y for the same variable will not affect the formula.
3. $\varphi[y := t] = \varphi$, given y is not free in φ. In other words, substitution of a non-free variable will not affect the formula.
4. After substitution, the free variables should be
   (a) If $y \notin FV(\varphi)$, then $FV(\varphi[y := t]) = FV(\varphi)$. This follows from the previous point.
   (b) If $\varphi := \alpha(x_1, \ldots, x_i, y, x_{i+1}, \ldots, x_n)$, then
   $$FV(\varphi[y := t]) = \{x_1, \ldots, x_i, t, x_{i+1}, \ldots, x_n\} = (FV(\varphi) \setminus \{y\}) \cup \{t\}$$
   (c) If $\varphi := \forall x(\lambda)$, $FV(\varphi[y := t]) = FV(\lambda[y := t]) \setminus \{x\}$
   (d) If $\varphi := \lambda \to \kappa$, then $FV(\varphi[y := t]) = FV(\lambda[y := t]) \cup FV(\kappa[y := t])$.
   (e) If $\varphi := \neg\lambda$, then $FV(\varphi[y := t]) = FV(\lambda[y := t])$.

There can be a downside to arbitrary substitution, however. Take $\varphi(y) := \neg\forall t(y = t)$, for example. The sentence $\forall y(\varphi(y))$ (i.e. for all y, there exists t such that $y \neq t$) is valid when it is interpreted in a model with at least 2 elements. However, a particular instance of it, $\varphi(t) = \neg\forall t(t = t)$, is not valid in any structure since it says that there exists t such that $t \neq t$. This is due to the substituted variable (which is t) is captured immediately by quantifier, so in a way, t is not *substitutable* for y in φ.

In general, given a wff φ and variables y, t, we say that t is *substitutable* for y in φ if, inductively,

1. If φ is atomic, then t is always substitutable for y (even when y is not free in φ).
2. If $\varphi := \forall x(\lambda)$, then t is substitutable for y if either
   (a) y does not occur free in φ (or $\forall x(\lambda)$).
   (b) $t \neq x$ and t is substitutable for y in λ.
3. If $\varphi := \neg\lambda$, then t is substitutable for y in φ if and only if it is substitutable in λ.
4. If $\varphi := \lambda \to \kappa$, then t is substitutable for y in φ if and only if it is substitutable in both λ and κ.

1.2.4 Hilbert system

The Hilbert system (for predicate logic) is a deductive system such that

1. Given wffs α, β, γ, the followings are axioms (for propositional part):

   (IP1) $\alpha \to (\beta \to \alpha)$

   (IP2) $[\alpha \to (\beta \to \gamma)] \to [(\alpha \to \beta) \to (\alpha \to \gamma)]$

   (PNC) $(\neg\alpha \to \neg\beta) \to [(\neg\alpha \to \beta) \to \alpha]$

2. Given variables x, t and wffs α, β, the followings are axioms (for quantified part):

   (UE) $\forall x(\alpha) \to \alpha[x := t]$, as long as t is substitutable for x in α.

   (UI) $\forall x(\alpha \to \beta) \to [\forall x(\alpha) \to \forall x(\beta)]$.

   (UG) $\alpha \to \forall x(\alpha)$, given x is not free in α.

3. (GR) $\forall x(\alpha)$ is an axiom whenever α is an axiom and x is a variable.

4. (MP) Modus ponens hold, i.e. $\alpha, \alpha \to \beta \vdash \beta$ given wffs α and β.

Since it contains axioms (and reference rule) of propositional logic, everything from previous section still holds here, and that includes the truth table method: if a formula is a tautology, then it is a theorem. The atomic formulas in predicate logic served the same role as the atomic formulas in propositional logic.

Lemma 1.36. Suppose $\varphi \equiv \psi \pmod{\Gamma}$, then we have the following

1. $\varphi[x := t] \equiv \psi[x := t] \pmod{\Gamma}$, where x is not free in Γ.
2. t is substitutable for x in φ if and only if it is substitutable for x in ψ, given x is not free in Γ.
3. $FV(\varphi) \setminus \bigcup_{\lambda\in\Gamma} FV(\lambda) = FV(\psi) \setminus \bigcup_{\lambda\in\Gamma} FV(\lambda)$.

Proof. Part 1: since $\Gamma \vdash \varphi \to \psi$, □

Theorem 1.37.

1. If $\Gamma \vdash \varphi$, and x is a variable that is not free in any formula belong to Γ, then $\Gamma \vdash \forall x(\varphi)$.
2. If $\Gamma \vdash \varphi$, and x is a variable that is not free in Γ, then there exists a variable t not free in φ such that $\Gamma \vdash \forall t(\varphi[x := t])$.

3. If $\Gamma, \varphi[x := t] \vdash \psi$, and t is not free in φ, ψ, Γ, then $\Gamma, \exists x(\varphi) \vdash \psi$.

4. If $\varphi \equiv \psi \pmod{\Gamma}$, and x is a variable that is not free in Γ, then $\forall x(\varphi) \equiv \forall x(\psi) \pmod{\Gamma}$.

5. If t is not free while substitutable for x in φ, then $\forall x(\varphi) \equiv \forall t(\varphi[x := t])$.

6. Let φ be a wff, x, t are variables, then there exists another formula φ' such that $\varphi \equiv \varphi'$ and t is substitutable for x in φ'.

Example 1.38.

$$\exists e \forall x(xe = x), \exists f \forall x(fx = x) \vdash \exists h \forall x(xh = hx = x)$$

1. By universal instantiation,

1.2.5 Vectorial Notation

If φ is a wff, and $x_1, \ldots, x_n$ are its free variable, then we can shorten $\varphi(x_1, \ldots, x_n)$ into $\varphi(\vec{x})$ using vectorial notation (where $\vec{x} := (x_1, \ldots, x_n)$).

Similarly, if we have a consecutive sequence of universal quantification, or existential quantification, we can abbreviate them too! Given $\vec{x} = (x_1, \ldots, x_n)$, then

$$\begin{aligned} \forall \vec{x}(\varphi) &:= \forall x_1 \ldots \forall x_n(\varphi) \\ \exists \vec{x}(\varphi) &:= \exists x_1 \ldots \exists x_n(\varphi) \\ &\equiv \neg \forall x_1 \ldots \forall x_n(\neg \varphi) \equiv \neg \forall \vec{x}(\neg \varphi) \end{aligned}$$

1.2.6 Some important metatheorems (draft)

Theorem 1.39.

1. Universal Generalization: if $\Gamma, \varphi \vdash \psi$, and x is free in Γ, then $\Gamma, \forall x(\varphi) \vdash \forall x(\psi)$.

2. Existential Generalization: if $\Gamma, \varphi \vdash \psi$, and x is free in Γ, then $\Gamma, \exists x(\varphi) \vdash \exists x(\psi)$.

3. Quantifier Elimination: if x is not free in φ, then $\forall x(\varphi) \vdash \varphi$ and $\exists x(\varphi) \vdash \varphi$.

4. Change of variable: if w is a variable different from x that is not free in Γ and φ, then

   (a) $\Gamma \vdash \forall w(\varphi[x := w])$ whenever $\Gamma \vdash \forall x(\varphi)$.

   (b) $\Gamma \vdash \exists w(\varphi[x := w])$ whenever $\Gamma \vdash \exists x(\varphi)$.

1.3 First-order logic

1.3.1 Translation of function symbols

1.3.2 The problem of empty domain

1.3.3 Functions

A predicate φ depending on 2 tuples $\mathbf{x}$ (of size n) and $\mathbf{y}$ (of size m) is said to be

1. ***Exact*** if
$$\forall \mathbf{x} \forall \mathbf{y} \forall \mathbf{z} \left[(\varphi(\mathbf{x}, \mathbf{y}) \wedge \varphi(\mathbf{x}, \mathbf{z})) \to \mathbf{y} = \mathbf{z} \right]$$
with tuples of appropriate size so that the expression makes sense.
2. ***Total*** if
$$\forall \mathbf{x} \exists \mathbf{y} (\varphi(\mathbf{x}, \mathbf{y}))$$

We say φ is

1. A ***function*** (in n-variable returning m-values) if it is exact and total.
$$\forall \mathbf{x} \exists ! \mathbf{y} (\varphi(\mathbf{x}, \mathbf{y}))$$
The tuple $\mathbf{x}$ is called the ***input***, while $\mathbf{y}$ is called the ***output***
2. A ***partial function*** if the totality condition is dropped.
3. A ***multivalued function*** if the exactness condition is dropped.

Given a function φ, we say that $\mathbf{x}$ is mapped to $\mathbf{y}$ if $\varphi(\mathbf{x}, \mathbf{y})$ holds. Taking some random letter, say f, and write $f(\mathbf{x})$ to denote the unique tuple $\mathbf{y}$ that is mapped to $\mathbf{x}$. In other words,
$$\mathbf{y} = f(\mathbf{x}) := \varphi(\mathbf{x}, \mathbf{y})$$
Using substitution, $\varphi(\mathbf{x}, f(\mathbf{x}))$ is always true.

With such notation, we can "substitute" functions into other predicate or even functions by the following interpretation

1. Let say $\varphi(\mathbf{y})$ is a predicate, $\psi(\mathbf{x}, \mathbf{y})$ is a function with associated letter g (i.e. $\psi(\mathbf{x}, g(\mathbf{x}))$ always hold). Then
$$\varphi(f(\mathbf{x})) := \exists \mathbf{y} \left(\varphi(\mathbf{y}) \wedge \psi(\mathbf{x}, \mathbf{y}) \right)$$

2. On the other hand, let $\varphi(\mathbf{x}, \mathbf{y})$ and $\psi(\mathbf{y}, \mathbf{z})$ be functions such that the size of output of φ is the same as that of the input of ψ (hence the matching letter). If f, g are the functional symbols associated with φ, ψ respectively, then the predicate

$$\mathbf{z} = g(f(\mathbf{x})) := \exists \mathbf{y} \left(\varphi(\mathbf{x}, \mathbf{y}) \wedge \psi(\mathbf{y}, \mathbf{z})\right)$$

is associated with $g \circ f$, i.e. $g \circ f(\mathbf{x})$ is the same as $g(f(\mathbf{x}))$.

Note: as exactness and totality conditions are propositions, one must add them as axioms in order to guarantee that a predicate is a function.

1.3.4 Constants

A special case of function is when the input is of size 0. The condition $\forall \mathbf{x} \exists! \mathbf{y}(\varphi(\mathbf{x}, \mathbf{y}))$ just drop to $\exists! \mathbf{y}(\varphi(\mathbf{y}))$. Such predicate is called ***constant***.

Taking another random letter, like $\mathbf{c}$ to associate φ, i.e. $\varphi(\mathbf{c})$ is true. Similarly, we can also "substitute" constants into predicates and functions as follows

1. Let say $\mathbf{c}$ is a constant associated with a predicate $\psi(\mathbf{x})$, and $\varphi(\mathbf{x})$ is another predicate with the same arity as ψ. Then

$$\psi(\mathbf{c}) := \exists \mathbf{x} \left(\psi(\mathbf{x}) \wedge \varphi(\mathbf{x})\right)$$

2. Again, $\mathbf{c}$ is a constant associated with a predicate $\psi(\mathbf{x})$, and $\varphi(\mathbf{x}, \mathbf{y})$ is a function with associated symbol f. Then

$$\mathbf{y} = f(\mathbf{c}) := \exists \mathbf{x} \left(\psi(\mathbf{x}) \wedge \varphi(\mathbf{x}, \mathbf{y})\right)$$

Chapter 2

Axiomatic Set Theory

2.1 Zermelo-Frankael set theory with Axiom of Choice (ZFC)

Proposition 2.1. $(\prod_{i\in I} A_i) \cap (\prod_{i\in I} B_i) = \prod_{i\in I}(A_i \cap B_i)$

Proposition 2.2 (Commutativity of Cartesian product)**.** Let $\{S_i\}_{i\in I}$ be a family of sets. If π is a permutation on the index set I, there is a natural bijection between $\prod_{i\in I} S_i$ and $\prod_{i\in I} S_{\pi(i)}$.

Proposition 2.3 (Associativity of Cartesian product)**.** Let $\{\{S_i\}_{i\in I_\alpha}\}_{\alpha\in A}$ be a family of families of sets. Show that there is natural bijection between $\prod_{\alpha\in A}\prod_{i\in I_\alpha} S_i$ and $\prod_{(i,\alpha)\in\coprod_{\alpha\in A} I_\alpha} S_{i,\alpha}$

2.2 Powered Cofinality

1. The λ-*summatory cofinality* of κ is defined as $\mathrm{scf}_\lambda \kappa = \sup\{\sum_{i\in I}\kappa_i : \kappa_i < \kappa, |I| < \lambda\}$
2. The λ-*powered cofinality* of κ is defined as $\mathrm{pcf}_\lambda \kappa = \sup\{\prod_{i\in I}\kappa_i : \kappa_i < \kappa, |I| < \lambda\}$

2.3 Lattice operations on collection

Let S be a set, C be a class (of sets), and $\mathcal{A}, \mathcal{B}$ be families of subsets in S.

1. $\mathcal{A}^\vee_C = \{\bigcup \mathcal{F} : \mathcal{F} \subseteq \mathcal{A}, \mathcal{F} \in C\}$
2. $\mathcal{A}^\wedge_C = \{\bigcap \mathcal{F} : \mathcal{F} \subseteq \mathcal{A}, \mathcal{F} \in C\}$
3. $C_{\mathcal{A}}(\mathcal{B}) = \{A \setminus B : A \in \mathcal{A}, B \in \mathcal{B}\}$.

If

1. $C = V$ is the von Neumann universe, we denote $\mathcal{A}^\vee := \mathcal{A}^\vee_V$ and $\mathcal{A}^\wedge := \mathcal{A}^\wedge_V$.
2. If $C = V_{<\kappa}$ is the class of sets with cardinality less than κ, we denote $\mathcal{A}^\vee_{<\kappa} := \mathcal{A}^\vee_C$ and $\mathcal{A}^\wedge_{<\kappa} := \mathcal{A}^\wedge_C$.

Proposition 2.4.

1. If C contains all singletons, then $\mathcal{A}$ is a sub-collection of $\mathcal{A}^\vee_C$ and $\mathcal{A}^\wedge_C$.
2. If C contains empty set, then $\varnothing \in \mathcal{A}^\vee_C$, and $S \in \mathcal{A}^\wedge_C$

If

1. $\mathcal{A}^\vee_C = \mathcal{A}$, we say $\mathcal{A}$ is closed under C-union.
2. $\mathcal{A}^\wedge_C = \mathcal{A}$, we say $\mathcal{A}$ is closed under C-intersection.
3. $S \setminus \mathcal{A} = \mathcal{A}$, we say $\mathcal{A}$ is closed under complementary

Example 2.5.

1. A topology $\mathcal{T}$ on a set X is a collection containing the empty set and X, and satisfy $\mathcal{T}^\vee = \mathcal{T}$, $\mathcal{T}^\wedge_{<\aleph_0} = \mathcal{T}$.

2.4 Transfinite Lattice

Let X be a set, $\mathcal{A}$ be a family of subsets of X. We say $\mathcal{A}$ is a (κ, λ)-lattice if

1. It is closed under $< \kappa$-union: $\bigcup_{i\in I} S_i \in \mathcal{A}$ whenever $S_i \in \mathcal{A}$ and $|I| < \kappa$.
2. It is closed under $< \lambda$-intersection: $\bigcup_{i\in I} S_i \in \mathcal{A}$ whenever $S_i \in \mathcal{A}$ and $|I| < \kappa$.

When $\kappa = \infty$, we allow arbitrary union, and say $\mathcal{A}$ is a λ-meet lattice. When $\lambda = \infty$, we allow arbitrary intersection, and say $\mathcal{A}$ is a κ-join lattice.

Example 2.6.

1. A topological space is $\aleph_0$-meet lattice.
2. A σ-algebra is an $\aleph_1$-lattice.

2.5 Order Theory

Theorem 2.7. The cofinality of a totally ordered set exists.

Theorem 2.8 (Monad of Ultrafilters)**.** For each $\mathfrak{F} \in \beta\beta X$, the collection $\mathcal{F} = \{A : [A] \in \mathfrak{F}\}$ is an ultrafilter.

2.6 Class Induction

Theorem 2.9. Given a unary class function α, show that

$$\forall F[F \neq \emptyset \to \exists B(F \subseteq B \wedge \forall x \in B(\alpha(x) \in B))]$$

Meaning, for each non-empty set F, there exists another set B containing it and closed under α operation.

Proof. Define the Godel beta formula

$$\beta(n, C) := \exists M \exists f[(f : (n{+}1) \times F \to M \text{ is a function}) \wedge \forall x \in F(f(0, x) = x) \wedge \forall m < n, \forall x \in F(f(m{+}1, x) = \alpha(f(m, x))$$

□

Corollary 2.10. If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are class functions, then

$$\forall F[F \neq \emptyset \to \exists B[F \subseteq B \wedge \forall \vec{x}_1 \in B(\alpha_1(\vec{x}_1)) \wedge \cdots \wedge \forall \vec{x}_n \in B(\alpha_1(\vec{x}_n))]]$$

Corollary 2.11. If $\{\alpha_x\}_{x \in X}$ is a family of class functions (i.e. a formula φ such that $\alpha_x(y) = z$ iff $\varphi(x, y, z)$ holds)

$$\forall F[F \neq \emptyset \to \exists B[F \subseteq B \wedge \forall x \in X, \forall y \in B, \forall z(\varphi(x, y, z) \leftrightarrow z \in B)]$$

Chapter 3

Category Theory I

[Reference: Any regular mono is extremal]

Chapter 4

General Topology I

Any things we say will be applicable under ZFC, but we will always put a note after each theorem to clarify which axiom we need (in addition to ZF).

4.1 Axiom

A *topology* $\mathcal{T}$ on a set X is a collection of subsets of X (i.e. $\mathcal{T} \subseteq \mathcal{P}(X)$) such that

1. It contains the empty set and X.
2. It is closed under arbitrary union: for any subcollection $\mathcal{S}$ of $\mathcal{T}$, we have $\bigcup \mathcal{S} \in \mathcal{T}$.
3. It is closed under finite intersection: for any $A, B \in \mathcal{T}$, we have $A \cap B \in \mathcal{T}$.

Such pair $(X, \mathcal{T})$ is then called a *topological space* (or space), which can be abbreviated to just X.

4.1.1 Open set - Closed set

Any set in $\mathcal{T}$ is called an *open* set, and any set that is the complement of an open set is called a *closed* set.

Remark 4.1. The door (being either open or closed) analogy will not work here, as there are sets that are both open and closed, called *clopen*. These include the empty set and the space X itself.

Lemma 4.2. If U is open and C is closed, then $U \setminus C$ is open, while $C \setminus U$ is closed.

Proof. These properties follow from $U \setminus C = U \cap (X \setminus C)$ and $C \setminus U = C \cap (X \setminus U)$ □

Theorem 4.3 (Topology in terms of closed sets)**.** From a topology $\mathcal{T}$ on X, we can defined $X \setminus \mathcal{T}$ as the collection of closed sets in X. It satisfies the following properties:

1. It contains the empty set and X.
2. It is closed under arbitrary *intersection*: for any subcollection $\mathcal{S}$ of $\mathcal{T}$, we have $\bigcap \mathcal{S} \in \mathcal{T}$.
3. It is closed under finite union: for any $A, B \in \mathcal{T}$, we have $A \cup B \in \mathcal{T}$.

Vice versa, Given any collection $\mathcal{L}$ of subsets of X satisfying the 3 above properties, then $\mathcal{T} = \{X \setminus C : C \in \mathcal{L}\}$ is a topology on X, and the closed sets (with respect to $\mathcal{T}$) coincide with those in $\mathcal{L}$, i.e. $\mathcal{L} = X \setminus \mathcal{T}$

Example 4.4. There are few distinguished example of topologies on a set X

1. The *discrete topology* $\mathcal{T} = \mathcal{P}(X)$, where every subset is both open and closed.
2. The *indiscrete topology* $\mathcal{T} = \{\varnothing, X\}$, where the only open sets are exactly those in 1st condition. It is straightforward to see that any topology must contain the indiscrete topology.
3. For an infinite cardinal κ, the *co-κ topology* $\mathcal{T} = \{U \subseteq X : U = \varnothing \text{ or } |X \setminus U| < \kappa\}$ where the closed sets are either X or has cardinality smaller than κ.
   (a) If $\kappa = \aleph_0$, we call it the *cofinite topology*. The closed sets in this space are either X or has finite number of elements.
   (b) If $\kappa = \aleph_1$, we call it the *cocountable topology*. The closed sets in this space are either X or has countable number of elements.

Proof.
Part 1: result from the fact that union and intersection of subsets of X are still subsets of X.
Part 2: any family of sets in $\mathcal{T}$ can only contain either X or $\varnothing$ (or both), so we only need to verify that $X \cup \varnothing = X$ and $X \cap \varnothing = \varnothing$ are open.
Part 3: since we can formulate topology in terms of closed sets by Theorem 4.3, we can verify those conditions for closed sets (which is considerably easier). Additionally, we can consider sets that are not X itself, since union of sets with X is X itself, and the intersection of sets with X is just the same intersection (both are closed, assuming we can show them so!).

Let $\{C_i\}_{i \in I}$ be an arbitrary family of closed set. Then $|C_i| < \kappa$ for each $i \in I$, so $|\bigcap_{i \in I} C_i| \leq |C_i| < \kappa$, i.e. $\bigcap_{i \in I} C_i$ is closed. On the other hand, if K and L are closed, then by cardinal arithmetic, we get $|K \cup L| \leq |K| + |L| < \kappa + \kappa = \kappa$, i.e. $K \cup L$ is closed. □

Proposition 4.5. If $\kappa > |X|$, then the co-κ topology on X is the same as the discrete topology. The vice versa also holds. Thus, in particular,

1. The cofinite topology on a finite set is discrete, while it is not on infinite set.
2. The cocountable topology on a countable set is discrete, while it is not on uncountable set, say ω_1 or $\mathbb{R}$.

Proof. Suppose $\kappa > |X|$, then any subset A is closed, as $|A| \leq |X| < \kappa$. Vice versa, suppose every subset of X is closed under co-κ topology. If X is finite, then $\kappa \geq \aleph_0 > |X|$. If X is infinite, then there exists a proper subset S of X such that $|X| = |S|$. This means that $|S| < \kappa$, and so $|X| < \kappa$. □

Proof Note: we need "infinite = Dedekind-infinite" (or equivalently, if X is infinite, then $|X| \geq \aleph_0$).

4.1.2 Comparison of topologies

We denote the collection of topologies on X as $\mathrm{Top}(X)$

Proposition 4.6. $\mathrm{Top}(X)$ forms a complete lattice under inclusion

1. The meet of topologies are just their intersection
$$\bigwedge \mathfrak{U} = \bigcap \mathfrak{U}$$
2. The join of topologies are the intersection of all topologies containing their union
$$\bigvee \mathfrak{U} = \bigcap_{\bigcup \mathfrak{U} \subseteq \mathcal{T}} \mathcal{T}$$
3. Additionally, given a collection $\mathcal{A}$ of subsets of X, there is a smallest topology containing $\mathcal{A}$, called the topology *generated* by $\mathcal{A}$.

Proof. Let $\mathfrak{U}$ be a collection of topologies on X.

1. Meet: since $\varnothing, X \in \mathcal{T}$ for any topology, we have $\varnothing, X \in \bigcap \mathfrak{U}$. On the other hand,
   (a) If $\{V_i\}_{i \in I}$ is a family of sets in $\bigcap \mathfrak{U}$, then $V_i \in \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{U}$. Thus, $\bigcup_{i \in I} V_i \in \mathcal{T}$ for all $\mathcal{T} \in \mathfrak{U}$, and so $\bigcup_{i \in I} V_i \in \bigcap \mathfrak{U}$.
   (b) If $P, Q \in \mathfrak{U}$, then $P, Q \in \mathcal{T}$ for any $\mathcal{T} \in \mathfrak{U}$. Hence, $P \cap Q \in \mathcal{T}$, and so $P \cap Q \in \bigcap \mathfrak{U}$.
2. Join: note that $\mathrm{Top}(X)$ is a Dedekind-complete meet-semilattice with a top element - the discrete topology. This automatically makes it into a complete lattice with $\bigvee \mathfrak{U} = \bigwedge \{\mathcal{T} : \mathcal{T} \supseteq \mathcal{S} \text{ for all } \mathcal{S} \in \mathfrak{U}\} = \bigcap_{\bigcup \mathfrak{U} \subseteq \mathcal{T}} \mathcal{T}$.

3. Given $\mathcal{A} \subseteq \mathcal{P}(X)$, the intersection $\bigcap_{\mathcal{A} \subseteq \mathcal{T}} \mathcal{T}$ is non-empty and forms a topology on X. This is, by construction, the smallest topology containing $\mathcal{A}$ (note the intersection).

□

Given 2 topologies $\mathcal{T}_1$ and $\mathcal{T}_2$ on X, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$, we say that $\mathcal{T}_1$ is *coarser* than $\mathcal{T}_2$, or $\mathcal{T}_2$ is *finer* than $\mathcal{T}_1$.

Lemma 4.7. $\mathcal{T}_1$ is coarser than $\mathcal{T}_2$ if and only if $X \setminus \mathcal{T}_1 \subseteq X \setminus \mathcal{T}_2$.

Proof. Suppose $\mathcal{T}_1$ is coarser than $\mathcal{T}_2$. Let C be a closed set with respect to $\mathcal{T}_1$, then $X \setminus C \in \mathcal{T}_1$, and so $X \setminus C \in \mathcal{T}_2$. In other words, C is closed with respect to $\mathcal{T}_2$.
*Vice versa, suppose any closed set under $\mathcal{T}_1$ is also closed under $\mathcal{T}_2$. Let U be an open set under $\mathcal{T}_1$, then $X \setminus U \in X \setminus \mathcal{T}_1 \subseteq X \setminus \mathcal{T}_2$. Therefore, U is open under $\mathcal{T}_2$. □

Proposition 4.8. If $\mathcal{T}_n$ denote the co-κ_n topology ($n = 1, 2$), and $\kappa_1 \leq \kappa_2$. Then $\mathcal{T}_1$ is coarser than $\mathcal{T}_2$.

Proof. By the previous Lemma 4.7, we can instead verify the coarseness on closed sets. If $C \neq X$ is closed under co-κ_1 topology, then $|C| < \kappa_1 \leq \kappa_2$. We conclude C is also closed under co-κ_2 topology. □

4.2 Neighborhood. Basis

Given a space X, a subset N of X is said to be a *neighborhood* around a point p in X (i.e. $p \in X$) if it contains an open set U (not necessarily distinct from N) such that $x \in U$.

Remark 4.9.

1. If N is a neighborhood around any point p of a certain subset S, then there exists an open set U such that $S \subseteq U \subseteq N$. The converse holds trivially.

   We say that N is a neighborhood around S (so this definition generalizes the one at the beginning).

2. A set is open if and only if it is a neighborhood of each point in it.

Proof.
Part 1: let V be the union of all open sets contained in N. In particular, V is an open neighborhood of each point p in S, and $V \subseteq N$. Since N is an open neighborhood of S, for each $p \in S$, there exists some open set U_p such that $p \in U_p \subseteq N$. By definition, $U_p \subseteq V$. Hence, $p \in V$ for all $p \in S$, and so $S \subseteq V$.

Part 2: suppose U is open, then it is trivially a neighborhood of each point in it. Vice versa, suppose U is a neighborhood of its own, then there exists some open V (by part 1) such that $U \subseteq V \subseteq U$. We conclude that $U = V$ is open. □

Since neighborhood is more general than *open* neighborhood, there might be cases where we need distinguish between them. However, for all practical purposes, we do not, as evidenced by the following.

Theorem 4.10 (Semantic Equivalence)**.** Let X be a toplogical space, $\varphi(N)$ is a predicate on subsets of X (possibly with parameter) such that

$$\exists U \subseteq N(\varphi(U)) \to \varphi(N)$$

holds, then

$$\forall N \in \mathcal{N}_p(\varphi(N)) \leftrightarrow \forall U \in \mathcal{N}_p, U \text{ open}(\varphi(U))$$

Note another way to state the above condition on φ is that $\varphi(U) \to \varphi(N)$ whenever $U \subseteq N$. In other words, if we interpret φ as a collection of subsets of X where φ holds, then such collection is an upper set!

Proof. The forward direction is just specialization. As for the backward direction: let N be a neighborhood of p. By definition, there exists an open neighborhood U of p such that $U \subseteq N$. From the hypothesis, we know that $\varphi(U)$ holds. From the condition on φ, we conclude $\varphi(N)$ must be true. □

Corollary 4.11 (Semantic Equivalence - Existential Quantifier)**.** If φ is a predicate on subsets of a topological space X such that

$$\varphi(N) \to \forall U \subseteq N(\varphi(U))$$

holds, then

$$\exists N \in \mathcal{N}_p(\varphi(N)) \leftrightarrow \exists U \in \mathcal{N}_p, U \text{ open}(\varphi(U))$$

Proof. Let $\psi(N) := \neg\varphi(N)$, then by contraposition, $\exists U \subseteq N(\psi(U)) \to \psi(N)$ holds. Thus,

$$\begin{aligned} \forall N \in \mathcal{N}_p(\psi(N)) &\leftrightarrow \forall U \in \mathcal{N}_p, U \text{ open}(\psi(U)), \text{ or} \\ \forall N \in \mathcal{N}_p(\neg\varphi(N)) &\leftrightarrow \forall U \in \mathcal{N}_p, U \text{ open}(\neg\varphi(U)), \text{ or} \\ \exists N \in \mathcal{N}_p(\varphi(N)) &\leftrightarrow \exists U \in \mathcal{N}_p, U \text{ open}(\varphi(U)) \end{aligned}$$

□

Example 4.12. The following are examples of predicate satisfying "$\exists U \subseteq N(\varphi(U)) \to \varphi(N)$"

1. $\varphi(N) := p \in N$.
2. $\varphi(N) := N \not\subseteq A$, where A is a subset of X.
3. $\varphi(N) := N \cap A \neq \varnothing$.
4. $\varphi(N) := f^{-1}(N)$ is a neighborhood of p, where $f : X \to Y$ is a function between the underlying sets of topological spaces, and $p \in X, N \subseteq Y$.

A collection $\mathcal{B}$ of open sets is called a *basis* for X if for each open set U and a point x in U, there exists some $B \in \mathcal{B}$ such that $x \in B \subseteq U$

Lemma 4.13. Given a basis $\mathcal{B}$ of a topological space X, then every open set is a union of sets in $\mathcal{B}$. The empty set is a special case when the union is empty.

Proof. Given U an open set, let $V = \bigcup_{\substack{B \in \mathcal{B} \\ B \subseteq U}} B$, which is also open and a subset of U (even if V is empty). Now since $\mathcal{B}$ is a basis, for each $x \in U$, there exists some $B \in \mathcal{B}$ such that $x \in B \subseteq U$. By construction, $B \subseteq V$, so $x \in V$. In other words, $U \subseteq V \subseteq U$. Hence, U is a union of sets in $\mathcal{B}$. □

Lemma 4.14 (Topology comparison based on basis)**.** Let $\mathcal{B}_n$ be the basis of the topology $\mathcal{T}_n$ on X $(n = 1, 2)$. The following are equivalent (TFAE):

1. $\mathcal{T}_2$ is finer than $\mathcal{T}_1$.
2. For each $x \in X$ and $B_1 \in \mathcal{B}_1$ containing x, there exists $B_2 \in \mathcal{B}_2$ such that $x \in B_2 \subseteq B_1$.

Proof. Assume $\mathcal{T}_2 \supseteq \mathcal{T}_1$. Let $x \in B_1 \in \mathcal{B}_1$, then $x \in B_1 \in \mathcal{T}_2$, so there must exist some $B_2 \in \mathcal{B}_2$ such that $x \in B_2 \subseteq B_1$.

Vice versa, assume the 2nd condition. From the previous Lemma, each $U_1 \in \mathcal{T}_1$ is the union $U_1 = \bigcup_{\substack{B \in \mathcal{B}_1 \\ B \subseteq U_1}} B$. Let $U_2 = \bigcup_{\substack{B \in \mathcal{B}_2 \\ B \subseteq U_1}} B$. By definition, $U_2 \in \mathcal{T}_2$ and U_2 is a subset of U_1. On the other hand, for each $x \in U_1$, there exists some $B_1^x \in \mathcal{B}_1$ and $B_2^x \in \mathcal{B}_2$ such that $x \in B_2^x \subseteq B_1^x \subseteq U_1$. Thus, $x \in U_2$, which should prove that $U_1 = U_2 \in \mathcal{T}_2$. We conclude that $\mathcal{T}_1 \subseteq \mathcal{T}_2$. □

A *neighborhood basis* $\mathcal{B}_p$ at p is a collection of neighborhoods of p such that for each neighborhood N at p, there exist some neighborhood B in $\mathcal{B}_p$ such that $B \subseteq N$.

4.2.1 Topology generated by a collection

Let $\mathcal{A}$ be a collection of subset of X.

1. Denote $\mathcal{A}^M = \mathcal{A} \cup \{\varnothing, X\}$.
2. Add finite intersections of sets in $\mathcal{A}$ to form $\mathcal{A}^{FI} = \{\bigcap_{i=1}^n U_i : U_i \in \mathcal{A}\}$.
3. Add arbitrary union of sets in $\mathcal{A}$ to form $\mathcal{A}^U = \{\bigcup \mathcal{F} : \mathcal{F} \subseteq \mathcal{A}\}$

A common property these operations share is that they always enlarge $\mathcal{A}$.

Remark 4.15. By definition, $\mathcal{T}$ is a topology if and only if $\mathcal{T}^M = \mathcal{T}^{FI} = \mathcal{T}^U = \mathcal{T}$.

Remark 4.16.

1. Extensionality: $\mathcal{A} \subseteq \mathcal{A}^M, \mathcal{A}^{FI}, \mathcal{A}^U$.
2. Monotonicity: if $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A}^M \subseteq \mathcal{B}^M$, $\mathcal{A}^{FI} \subseteq \mathcal{B}^{FI}$, and $\mathcal{A}^U \subseteq \mathcal{B}^U$.

Proof. Part 1: $\mathcal{A} \subseteq \mathcal{A}^M$ is straightforward, $\mathcal{A} \subseteq \mathcal{A}^{FI}$ because finite intersection includes singleton intersection, and $\mathcal{A} \subseteq \mathcal{A}^U$ because $\bigcup\{A\} = A$ for $A \in \mathcal{A}$.
Part 2: $\mathcal{A}^M \subseteq \mathcal{B}^M$ is straightforward, $\mathcal{A}^{FI} \subseteq \mathcal{B}^{FI}$ since if $U_i \in \mathcal{A}$, then $U_i \in \mathcal{B}$, and so $\bigcap_{i=1}^n U_i \in \mathcal{B}^{FI}$. Finally, $\mathcal{A}^U \subseteq \mathcal{B}^U$ because whenever $\mathcal{F} \subseteq \mathcal{A} \subseteq \mathcal{B}$, we get $\bigcup \mathcal{F} \in \mathcal{B}^U$. □

Lemma 4.17. A collection $\mathcal{B}$ of subsets of X is a *topological basis* if

1. It covers X: $\bigcup \mathcal{B} = X$.
2. For each $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists some $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

Then the topology generated by $\mathcal{B}$ is exactly $\mathcal{B}^U$.

Proof. It is straightforward from the axioms to see that $\mathcal{B}^U \subseteq \mathcal{T}$ for any topology containing $\mathcal{B}$. If $\mathcal{B}^U$ is a topology, then we're pretty much done since $\mathcal{B}^U$ would be the smallest topology containing $\mathcal{B}$.

1. The first condition (of a basis) shows that $X \in \mathcal{B}^U$. We also have $\varnothing \in \mathcal{B}^U$ by considering empty union.

2. It is closed under arbitrary union: since any element of $\mathcal{B}^U$ has the form of $\bigcup \mathcal{F}$, let $\{\mathcal{F}_i\}_{i\in I}$ be a collection of family of sets in $\mathcal{A}$. Then

$$\begin{aligned}\bigcup_{i\in I}\bigcup \mathcal{F}_i &= \bigcup_{i\in I}\{x \mid \exists E \in \mathcal{F}_i : x \in E\} \\ &= \{x \mid \exists i \in I, \exists E \in \mathcal{F}_i : x \in E\} \\ &= \bigcup\{E \mid \exists i \in I : E \in \mathcal{F}_i\} \\ &= \bigcup\bigcup_{i\in I} \mathcal{F}_i\end{aligned}$$

3. It is closed under finite intersection: first observe that for any $B_1, B_2 \in \mathcal{B}$, we have $B_1 \cap B_2$ a union of sets in $\mathcal{B}$, i.e. $B_1 \cap B_2 \in \mathcal{B}^U$. Indeed, denote $C = \bigcup_{\substack{B\in\mathcal{B} \\ B\subseteq B_1\cap B_2}}$, which should be in $\mathcal{B}^U$. Under second condition (of a basis), each $x \in B_1 \cap B_2$ there exists some $C_x \in \mathcal{B}$ such that $x \in C_x \subseteq B_1 \cap B_2$. By C's definition, we get $x \in C$, so $B_1 \cap B_2 = C \in \mathcal{B}^U$.

   Let $\bigcup_{i\in I} F_i, \bigcup_{j\in J} G_j \in \mathcal{B}^U$, then

$$\bigcup_{i\in I} F_i \cap \bigcup_{j\in J} G_j = \bigcup_{\substack{i\in I \\ j\in J}} F_i \cap G_j \in \mathcal{B}^U$$

   Since $F_i \cap G_j \in \mathcal{B}^U$ whenever $F_i, G_j \in \mathcal{B}^U$, and the previous part (i.e. arbitrary union of elements in $\mathcal{B}^U$ is still in $\mathcal{B}^U$).

□

Theorem 4.18. Let $\mathcal{A}$ be a subcollection of $\mathcal{P}(X)$, and $\mathcal{T}$ is the topology generated by $\mathcal{A}$. Then $\mathcal{T} = ((\mathcal{A}^M)^{FI})^U = ((\mathcal{A}^{FI})^M)^U = ((\mathcal{A}^{FI})^U)^M$

Proof. Adding X and $\varnothing$ in any order will not affect the process much since they are the greatest and least element in the (complete) lattice $\mathcal{P}(X)$, meaning we can always omit them when taking union and intersection. This let us to prove just 1 equality, say the first one.

Next, $\mathcal{B} = (\mathcal{A}^M)^{FI}$ is a basis! Indeed,

1. It covers X as $X \in \mathcal{A}^M \subset \mathcal{B}$
2. For each $B_1, B_2 \in \mathcal{B}$, which is represented by $B_1 = \bigcap_{i=1}^n C_i$ and $B_2 = \bigcap_{j=1}^m D_j$, then $B_1 \cap B_2$ is a finite intersection of sets (C_i and D_j) in $\mathcal{A}^M$, so it's still in $\mathcal{B}$. Hence, for any $x \in B_1 \cap B_2$, just pick $B = B_1 \cap B_2$ and we have $x \in B \subseteq B_1 \cap B_2$.

By the previous Lemma 4.17, the topology generated by $\mathcal{B}$ (which contains $\mathcal{A}$) coincides with $\mathcal{B}^U = ((\mathcal{A}^M)^{FI})^U$. Therefore, $\mathcal{T} \subseteq \mathcal{B}^U$ by minimality. On the other hand,

1. Any topology $\mathcal{T}'$ containing $\mathcal{A}$ must contain $\mathcal{A}^M$.
2. Any topology $\mathcal{T}'$ containing $\mathcal{A}^M$ must contain $\mathcal{B}$ (closed under finite intersection).
3. Finally, any topology $\mathcal{T}'$ containing $\mathcal{B}$ must contain $\mathcal{B}^U$ (closed under arbitrary union).

This shows the reverse inclusion $\mathcal{T} \supseteq \mathcal{B}^U$. □

Lemma 4.19. Let $\mathcal{A} \subseteq \mathcal{B}$ be collections of subsets of X, and $\mathcal{T}, \mathcal{S}$ respectively be the topologies generated by $\mathcal{A}, \mathcal{B}$. Then $\mathcal{T} \subseteq \mathcal{S}$.

Proof. Since $\mathcal{S}$ also contains $\mathcal{A}$, we get $\mathcal{T} \subseteq \mathcal{S}$ by minimality. □

Given a topological space X, the weight $w(X)$ of X is defined as the least cardinality of a basis of X.

Remark 4.20.

1. Since a basis of a space must be a subcollection of its topology, the weight is always less than or equal to the cardinality of the topology, which in turns is less than or equal to the cardinality of the power set of the space. In simple term, $w(X) \leq |\mathcal{T}| \leq 2^{|X|}$.
2. Since the class of cardinal numbers is well-ordered, there exists a basis with cardinality equal to the weight.

Lemma 4.21 (Cardinality of generation)**.** If $w(X)$ is infinite, then $w(X)$ is also the least cardinality of any collection of open subsets of X that generates X.

Proof. Denote κ to be the least cardinality of any collection of open subsets of X that generates X. Then $\kappa \leq w(X)$ by definition. On the other hand, for each $\mathcal{S}$ that generates X, $\mathcal{B} = (\mathcal{S}^M)^{FI}$ must be a basis. As $|\mathcal{B}| \geq w(X)$ is infinite,

$\mathcal{S}$ cannot be finite (as the first step adds at most 2 sets, and the second step adds finite intersection of these sets). Hence, $|\mathcal{S}| \leq |\mathcal{S}^M| \leq |\mathcal{S}| + 2 = |\mathcal{S}|$. As for $(\mathcal{S}^M)^{FI}$,

$$|\mathcal{S}| = |\mathcal{S}^M| \leq \left|(\mathcal{S}^M)^{FI}\right| \leq \left|\bigcup_{n=1}^{\infty} (\mathcal{S}^M)^n\right| = \sum_{n=1}^{\infty} |\mathcal{S}^M|^n$$
$$= \sum_{n=1}^{\infty} |\mathcal{S}| = \aleph_0 |\mathcal{S}| = \max(\aleph_0, |\mathcal{S}|) = |\mathcal{S}|$$

Hence, $|\mathcal{B}| = |\mathcal{S}|$. So for any $\mathcal{S}$ that generates X, there exists a basis $\mathcal{B}$ of X with the same cardinality. Therefore, $w(X) \leq \kappa$, and so $\kappa = w(X)$.
Proof Note: we need axiom of choice for arbitrary cardinalities to be comparable. □

4.2.2 Some topologies on the real numbers

Let $\mathbb{R}$ be the set of real numbers. Denote the following topologies on $\mathbb{R}$ as

1. $\mathcal{T}_0$ for the *standard topology*: it is generated by the basis consisting of open intervals (a, b).
2. $\mathcal{T}_0^+$ for the topology generated by the basis consisting of (a, ∞).
3. $\mathcal{T}_0^-$ for the topology generated by the basis consisting of $(-\infty, b)$.
4. $\mathcal{T}_\omega$ for the cofinite topology
5. $\mathcal{T}_{\omega_1}$ for the cocountable topology
6. $\mathcal{T}_l$ for the *lower limit topology*: it is generated by half-open intervals of the form $[a, b)$. The resulted line is called *Sorgenfrey line*.
7. $\mathcal{T}_u$ for the *upper limit topology*: it is generated by half-open intervals of the form $(a, b]$.
8. $\mathcal{T}_l^+$ for the topology generated by $[a, \infty)$.
9. $\mathcal{T}_u^-$ for the topology generated by $(-\infty, b]$.
10. $\mathcal{T}_K$ for the *K-topology* generated by (a, b) and $(a, b) \setminus K$ where $K = \{1/n : n \in \mathbb{N}\}$

Proposition 4.22. The following holds

1. $\mathcal{T}_0$ is (strictly) finer than $\mathcal{T}_0^+, \mathcal{T}_0^-, \mathcal{T}_\omega$.
   $\mathcal{T}_0$ is coarser than $\mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_K$.
   $\mathcal{T}_0$ is incomparable to $\mathcal{T}_{\omega_1}, \mathcal{T}_l^+, \mathcal{T}_u^-$.

2. $\mathcal{T}_0^+$ is incomparable to $\mathcal{T}_0^-, \mathcal{T}_\omega, \mathcal{T}_{\omega_1}, \mathcal{T}_u^-$.
   $\mathcal{T}_0^+$ is coarser than $\mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_l^+, \mathcal{T}_K$.

3. $\mathcal{T}_0^-$ is incomparable to $\mathcal{T}_\omega, \mathcal{T}_{\omega_1}, \mathcal{T}_l^+$.
   $\mathcal{T}_0^-$ is coarser than $\mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_u^-, \mathcal{T}_K$.

4. $\mathcal{T}_\omega$ is coarser than $\mathcal{T}_{\omega_1}, \mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_K$.
   $\mathcal{T}_\omega$ is incomparable to $\mathcal{T}_l^+, \mathcal{T}_u^-$.

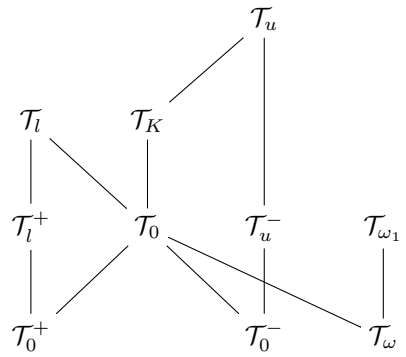
5. $\mathcal{T}_{\omega_1}$ is incomparable to $\mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_l^+, \mathcal{T}_u^-, \mathcal{T}_K$.

6. $\mathcal{T}_l$ is incomparable to $\mathcal{T}_u, \mathcal{T}_u^-, \mathcal{T}_K$.
   $\mathcal{T}_l$ is finer than $\mathcal{T}_l^+$.

7. $\mathcal{T}_u$ is incomparable to $\mathcal{T}_l^+$.
   $\mathcal{T}_u$ is finer than $\mathcal{T}_u^-, \mathcal{T}_K$.

8. $\mathcal{T}_l^+, \mathcal{T}_u^-, \mathcal{T}_K$ are mutually incomparable.

Here is the Hasse diagram



Proof.
1. $\mathcal{T}_0$ is finer than $\mathcal{T}_0^+$: use Lemma 4.14. For each $a \in \mathbb{R}$, $(a, \infty) = \bigcup_{b>a}(a, b)$. If $(-1, 1)$ is in $\mathcal{T}_0^+$, then there exists some (a, ∞) such that $0 \in (a, \infty) \subseteq (-1, 1)$, which impossible since for a big enough r (say $r > \max(1, a)$) then $r \in (a, \infty)$ yet $r \notin (-1, 1)$.
2. $\mathcal{T}_0$ is finer than $\mathcal{T}_0^+$: similar as (1)
3. $\mathcal{T}_0$ is finer than $\mathcal{T}_\omega$: use Lemma 4.7. Firstly, each point x is closed in $\mathcal{T}_0$ since $(-\infty, x) \cup (x, \infty) = \bigcup_{y<x}(y, x) \cup \bigcup_{y>x}(x, y)$ is open. Hence, if C is finite, then C is closed, since it is a finite union of closed singletons. On the other hand, $\mathbb{Z} = \mathbb{R} \setminus \bigcup_{n\in\mathbb{Z}}(n, n+1)$ is closed in $\mathcal{T}_0$ but not in $\mathcal{T}_\omega$.
4. $\mathcal{T}_0$ is coarser $\mathcal{T}_l$: $(a, b) = \bigcup_{\varepsilon>0}[a + \varepsilon, b)$. On the other hand, if $[0, 1)$ is open in $\mathcal{T}_0$, then there exists some (a, b) such that $0 \in (a, b) \subseteq [0, 1)$, meaning both $a < 0$ and $0 \leq a$, a contradiction.
5. $\mathcal{T}_0$ is coarser $\mathcal{T}_u$: similar as (4)
6. $\mathcal{T}_0$ is coarser $\mathcal{T}_K$: note (a, b) is already in $\mathcal{T}_K$. On the other hand, if $(-1, 1)\setminus K$

is open, then there exists an interval $(a, b) \subseteq (-1, 1) \setminus K$ around 0. Since $b > 0$, there exists some $n \in \mathbb{N}$ large enough so that $b > 1/n > 0$ (archimedean property). Thus $1/n \in (-1, 1) \setminus K$, a contradiction.
7. $\mathcal{T}_0$ is incomparable to $\mathcal{T}_{\omega_1}$: $[0, 1]$ is not closed in $\mathcal{T}_{\omega_1}$ since it is uncountable. On the other hand, $K = \{1/n : n \in \mathbb{N}\}$ is not closed in $\mathcal{T}_0$, since otherwise, there must be some (a, b) around 0 such that $(a, b) \subseteq \mathbb{R} \setminus K$.
8. $\mathcal{T}_0$ is incomparable to $\mathcal{T}_l^+$: $(-1, 1)$ is not open in $\mathcal{T}_l^+$ since otherwise, there exists some $[a, \infty)$ such that $0 \in [a, \infty) \subseteq (-1, 1)$ (impossible). On the other hand, $[0, \infty)$ is not open in $\mathcal{T}_0$ because any (a, b) containing 0 will contain negative numbers.
9. $\mathcal{T}_0$ is incomparable to $\mathcal{T}_u^-$: similar to (8).
10. $\mathcal{T}_0^+$ is incomparable to $\mathcal{T}_0^-$: similar to (8).
11. $\mathcal{T}_0^+$ is incomparable to $\mathcal{T}_\omega$: a singleton like $\{0\}$ cannot be closed in $\mathcal{T}_0^+$ since otherwise, there exists some (a, ∞) such that $-1 \in (a, \infty) \subseteq \mathbb{R} \setminus \{0\}$. This means that $a < -1$, so $0 \in (a, \infty)$, a contradiction. On the other hand, $(-\infty, 0]$ is not closed in $\mathcal{T}_\omega$ since it is uncountable.
12. $\mathcal{T}_0^+$ is incomparable to $\mathcal{T}_{\omega_1}$: same as (11).
13. $\mathcal{T}_0^+$ is incomparable to $\mathcal{T}_u^-$: similar to (8).
14. $\mathcal{T}_0^+$ is coarser than $\mathcal{T}_l$: combination of (1) and (4).
15. $\mathcal{T}_0^+$ is coarser than $\mathcal{T}_u$: combination of (1) and (5).
16. $\mathcal{T}_0^+$ is coarser than $\mathcal{T}_l^+$: $(a, \infty) = \bigcup_{\varepsilon > 0} [a + \varepsilon, \infty)$. On the other hand, $[0, \infty)$ cannot be open in $\mathcal{T}_0^+$ by an argument similar to (4).
17. $\mathcal{T}_0^+$ is coarser than $\mathcal{T}_K$: combination of (1) and (6).
18. $\mathcal{T}_0^-$ is incomparable to $\mathcal{T}_\omega$: similar to (11).
19. $\mathcal{T}_0^-$ is incomparable to $\mathcal{T}_{\omega_1}$: similar to (12).
20. $\mathcal{T}_0^-$ is incomparable to $\mathcal{T}_l^+$: similar to (8).
21. $\mathcal{T}_0^-$ is coarser than $\mathcal{T}_l$: combination of (2) and (4).
22. $\mathcal{T}_0^-$ is coarser than $\mathcal{T}_u$: combination of (2) and (5).
23. $\mathcal{T}_0^-$ is coarser than $\mathcal{T}_u$: similar to (16).
24. $\mathcal{T}_0^-$ is coarser than $\mathcal{T}_K$: combination of (2) and (6).
25. $\mathcal{T}_\omega$ is coarser than $\mathcal{T}_{\omega_1}$: result of Proposition 4.8. Additionally, $\mathbb{Z}$ is not closed in $\mathcal{T}_\omega$.
26. $\mathcal{T}_\omega$ is coarser than $\mathcal{T}_l$: combination of (3) and (4).
27. $\mathcal{T}_\omega$ is coarser than $\mathcal{T}_u$: combination of (3) and (5).
28. $\mathcal{T}_\omega$ is coarser than $\mathcal{T}_K$: combination of (3) and (6).
29. $\mathcal{T}_\omega$ is incomparable to $\mathcal{T}_l^+$ and $\mathcal{T}_u^-$: similar to (11).
30. $\mathcal{T}_{\omega_1}$ is incomparable to $\mathcal{T}_l, \mathcal{T}_u, \mathcal{T}_l^+, \mathcal{T}_u^-, \mathcal{T}_K$: similar to (7).
31. $\mathcal{T}_l$ is incomparable to $\mathcal{T}_u$: similar to (4).
32. $\mathcal{T}_l$ is incomparable to $\mathcal{T}_u^-$: similar to (8).
33. $\mathcal{T}_l$ is incomparable to $\mathcal{T}_K$: suppose $[0, 1)$ is open in $\mathcal{T}_K$, then either $0 \in (a, b) \subseteq [0, 1)$ or $0 \in (a, b) \setminus K$ for some $a, b \in \mathbb{R}$. In any case, we get $a < 0$ and $0 \leq a$, a contradiction. On the other hand, first note that $\mathbb{R} \setminus K =$

$(-\infty, 0] \cup \bigcup_{n=1}^{\infty} \left(\frac{1}{n+1}, \frac{1}{n}\right) \cup (1, \infty)$. So if K is closed in $\mathcal{T}_l$, then $0 \in [a, b) \subseteq \mathbb{R} \setminus K$ for some $a, b \in \mathbb{R}$. However, we have $0 < b$, so for large enough $n \in \mathbb{N}$, we get $0 < 1/n < b$, and so $1/n \in \mathbb{R} \setminus K$, a contradiction.

34. $\mathcal{T}_l$ is finer than $\mathcal{T}_l^+$: similar to (1).

35. $\mathcal{T}_u$ is incomparable to $\mathcal{T}_l^+$: similar to (8).

36. $\mathcal{T}_u$ is finer to $\mathcal{T}_u^-$: similar to (2).

37. $\mathcal{T}_u$ is finer to $\mathcal{T}_K$: note $(a, b) \in \mathcal{T}_u$ (see (4)), and $\mathbb{R} \setminus K = U_K \cup (-\infty, 0] = U_K \cup \bigcup_{a<0}(a, 0]$ (where $U_K \in \mathcal{T}_0 \subseteq \mathcal{T}_u$), we also get $(a, b) \setminus K \in \mathcal{T}_u$. In conclusion, $\mathcal{T}_u$ is finer than $\mathcal{T}_K$. Additionally, $(-1, 0]$ is not open in $\mathcal{T}_K$ since otherwise, either $0 \in (a, b) \subseteq (-1, 0]$, or $0 \in (a, b) \setminus K \subseteq (-1, 0]$. Either case gives us $b > 0$ and $b \leq 0$, a contradiction.

38. $\mathcal{T}_l^+$ is incomparable to $\mathcal{T}_u^-$: similar to (4) and (8).

39. $\mathcal{T}_l^+$ is incomparable to $\mathcal{T}_K$: similar to (33).

40. $\mathcal{T}_u^-$ is incomparable to $\mathcal{T}_K$: similar to (8). □

4.3 Subsets of topological spaces

Given a set A in a topological space X. We denote

1. $\text{int}(A)$ as the *interior* of A: this is the union of open sets contained in A

$$\text{int}(A) = \bigcup_{\substack{U \text{ open} \\ U \subseteq A}} U$$

2. $\text{cl}(A)$ as the *closure* of A: this is the intersection of closed sets *containing* A

$$\text{cl}(A) = \bigcap_{\substack{C \text{ closed} \\ C \supseteq A}} C$$

3. $\text{ext}(A)$ as the *exterior* of A: this is the union of open sets disjoint with A

$$\text{ext}(A) = \bigcup_{\substack{U \text{ open} \\ A \cap U = \varnothing}} U$$

4. ∂A as the *boundary* of A: this is the complement of the interior relative to the closure of A

$$\partial A = \text{cl}(A) \setminus \text{int}(A)$$

A point in $\text{int}(A)$ is called an interior point. A point in $\text{cl}(A)$ is called a closure point. A point in $\text{ext}(A)$ is called an exterior point. And a point in ∂A is called a boundary point.

Proposition 4.23 (Topological properties)**.**

1. Extremal properties:
   (a) $\text{int}(A)$ is the largest open set contained in A. Hence, if U is open, then $U \subseteq A$ if and only if $U \subseteq \text{int}(A)$.
   (b) $\text{cl}(A)$ is the smallest closed set containing A. Hence, if C is closed, then $A \subseteq C$ if and only if $\text{cl}(A) \subseteq C$.
   (c) $\text{ext}(A)$ is the largest open set disjoint with A. Hence, if U is open, then $A \cap U = \varnothing$ if and only if $U \subseteq \text{ext}(A)$
2. Partition properties:
   (a) $\text{int}(A)$ and ∂A forms a partition of $\text{cl}(A)$.
   (b) $\text{cl}(A)$ and $\text{ext}(A)$ forms a partition of the space X itself.
3. Characterizations of points

(a) p is an interior point of A if and only if A is a neighborhood of p, or equivalently, there exists some open set U such that $p \in U \subseteq A$.

(b) p is a closure point of A if and only if every neighborhood of p meets A at some point.

(c) p is an exterior point of A if and only if there exists some neighborhood of p that are disjoint from A.

(d) p is a boundary point of A if and only if every neighbohood of p meets both A and its complement at some point.

4. Open and closed properties:

(a) A is open if and only if $A = \text{int}(A)$

(b) A is closed if and only if $A = \text{cl}(A)$, or equivalently, $A = X \setminus \text{ext}(A)$

(c) ∂A is always closed. Additionally, A is clopen if and only if $\partial A = \varnothing$.

Proof.
Part 1: if V is any open set contained in A, then $V \subseteq \bigcup_{\substack{U \text{ open} \\ U \subseteq A}} U = \text{int}(A)$ by definition. Since $\text{int}(A)$ is union of open sets, it is itself open, and so it is the largest open set contained in A. The other properties follow similarly.
Part 2:

(a) Straight from definition, with the fact that $\text{int}(A) \subseteq A \subseteq \text{cl}(A)$ (part (1)).

(b) Note that $X \setminus \text{cl}(A)$ is open and disjoint from A (since $A \subseteq \text{cl}(A)$, so we get $X \setminus \text{cl}(A) \subseteq \text{ext}(A)$ by maximality. Equivalently, $X \subseteq \text{cl}(A) \cup \text{ext}(A)$. But these are subsets of X, so the equality must hold.

Let $p \in \text{ext}(A)$, then there exists some open U such that $p \in U \subseteq X \setminus A$ by definition. In other words $p \notin X \setminus U \supset A$. As $X \setminus U$ is closed, we get $X \setminus U \supseteq \text{cl}(A)$. Thus $p \notin \text{cl}(A)$.

Part 3:

(a) If p is an interior point of A, then take $U = \text{int}(A)$ and we get $p \in U \subseteq A$. Vice versa, if $p \in U \subseteq A$ where U is open, then by definition $U \subseteq \text{int}(A)$, and so $p \in \text{int}(A)$.

(b) If p is a closure point of A, and N is a neighborhood of p, let U be an open set such that $p \in U \subseteq N$. If $U \cap A = \varnothing$, then by definition $p \in \text{ext}(A)$, contradicting the partition property in part (2). Thus, we must have $U \cap A \neq \varnothing$, and so any neighborhood of p must intersect A. Vice versa, if any neighborhood N of p intersect A, then p cannot be in $\text{ext}(A)$ since it would imply an existence of a neighborhood (of p) that is disjoint from A. Therefore $p \in X \setminus \text{ext}(A) = \text{cl}(A)$.

(c) If p is an exterior point of A, take $U = \text{ext}(A)$ and we get a neighborhood of A that is disjoint from A. Vice versa, if there exists some neighborhood N of p that are disjoint from A, then there also exists an open set U such that $p \in U \subseteq N$. As U is a subset of N, it should be disjoint from A. Thus, $U \subseteq \text{ext}(A)$ and so $p \in \text{ext}(A)$.

(d) If p is a boundary point of A, then every neighborhood of p meets A (note $\partial A \subseteq \text{cl}(A)$ by definition). To show that each neighborhood of p meets also $X \setminus A$, suppose otherwise, that there exists some open neighborhood U of p (as we can always find such inside any neighborhood) satisfying $U \cap (X \setminus A) = \varnothing$. Equivalently, $U \subseteq A$, so by part (3a), we get $p \in \text{int}(A)$. But $\text{int}(A)$ is disjoint from ∂A, a contradiction to the assumption on p.

Vice versa, suppose every neighborhood of p meets both A and $X \setminus A$. This means, by part (3c), that $p \in \text{cl}(A)$. Now if $p \in \text{int}(A)$, then there exists some open neighborhood U of p such that $U \subseteq A$. However, this means U is disjoint with the complement of A, so there should be no intersection between them, a contradiction. Hence, $p \in \text{cl}(A) \setminus \text{int}(A) = \partial A$.

Part 4: if $A = \text{int}(A)$, then A is open (since $\text{int}(A)$ is always open, from part (1)). Vice versa, if A is open, then $A \subseteq \text{int}(A)$ by definition of interior. Yet $\text{int}(A) \subseteq A$ (again from part (1)), so we get $A = \text{int}(A)$. The 2nd property follow similarly. As for ∂A, from part (2), we know that $\partial A = X \setminus (\text{int}(A) \cup \text{ext}(A))$, so the boundary is always closed. □

Proposition 4.24 (Algebraic Properties)**.** Let $\mathcal{M}_X = \text{End}(\mathcal{P}(X))$ be the collection of all maps, or operators, on $\mathcal{P}(X)$.

1. With composition, and the identity map I, this forms a (noncommutative) monoid

2. Define the partial ordering $\varphi \leq \psi$ if and only if $\varphi(A) \subseteq \psi(A)$ for all $A \subseteq X$.

3. This ordering is right-compatible with the monoid structure: given $\varphi \leq \psi$, then $\varphi\lambda \leq \psi\lambda$ for any $\lambda \in \mathcal{M}_X$.

4. It is also a complete lattice: the least element $\perp$ is the map sending everything to $\varnothing$, and the greatest element $\top$ is the map sending everything to X.

$$\left(\bigvee_{j\in J} \varphi_j\right)(A) = \bigcup_{j\in J} \varphi_j(A)$$
$$\left(\bigwedge_{j\in J} \varphi_j\right)(A) = \bigcap_{j\in J} \varphi_j(A)$$

5. The join and meet left-distribute with composition:

$$\left(\bigvee_{j\in J}\varphi_j\right)\psi = \bigvee_{j\in J}(\varphi_j\psi)$$
$$\left(\bigwedge_{j\in J}\varphi_j\right)\psi = \bigwedge_{j\in J}(\varphi_j\psi)$$

6. Finally, the complement reverse join and meet

$$c\left(\bigvee_{j\in J}\varphi_j\right) = \bigwedge_{j\in J}(c\varphi_j)$$
$$c\left(\bigwedge_{j\in J}\varphi_j\right) = \bigvee_{j\in J}(c\varphi_j)$$

Denote the letters k, i, e, b, c denotes the the closure, interior, exterior, boundary, complement operators correspondingly (and $I, \bot, \top$ as above). Prove

1. Monotone properties: given $\varphi \le \psi$ and $A \subseteq B$
   (a) $i\varphi \le i\psi$, or $\operatorname{int}(A) \subseteq \operatorname{int}(B)$.
   (b) $k\varphi \le k\psi$, or $\operatorname{cl}(A) \subseteq \operatorname{cl}(B)$.
   (c) $e\varphi \ge e\psi$, or $\operatorname{ext}(A) \supseteq \operatorname{ext}(B)$.
   (d) $c\varphi \ge c\psi$, or $X \setminus A \supseteq X \setminus B$.
2. Lattice properties (interior):
   (a) $i\left(\bigwedge_{j\in J}\varphi_j\right) \le \bigwedge_{j\in J}(i\varphi_j)$, or $\operatorname{int}\left(\bigcap_{j\in J}A_j\right) \subseteq \bigcap_{j\in J}\operatorname{int}(A_j)$.
   (b) $i(\varphi\wedge\psi) = i\varphi \wedge i\psi$, or $\operatorname{int}(A\cap B) = \operatorname{int}(A)\cap\operatorname{int}(B)$.
   (c) $i\left(\bigvee_{j\in J}\varphi_j\right) \ge \bigvee_{j\in J}(i\varphi_j)$, or $\operatorname{int}\left(\bigcup_{j\in J}A_j\right) \supseteq \bigcup_{j\in J}\operatorname{int}(A_j)$.
3. Lattice properties (closure):
   (a) $k\left(\bigvee_{j\in J}\varphi_j\right) \supseteq \bigvee_{j\in J}(k\varphi_j)$, or $\operatorname{cl}\left(\bigcup_{j\in J}A_j\right) \supseteq \bigcup_{j\in J}\operatorname{cl}(A_j)$.
   (b) $k(\varphi\vee\psi) = k\varphi \vee k\psi$, or $\operatorname{cl}(A\cup B) = \operatorname{cl}(A)\cup\operatorname{cl}(B)$.
   (c) $k\left(\bigwedge_{j\in J}\varphi_j\right) \subseteq \bigwedge_{j\in J}(k\varphi_j)$, or $\operatorname{cl}\left(\bigcap_{j\in J}A_j\right) \subseteq \bigcap_{j\in J}\operatorname{cl}(A_j)$.
4. Lattice properties (exterior):
   (a) $e\left(\bigvee_{j\in J}\varphi_j\right) \le \bigwedge_{j\in J}(e\varphi_j)$, or $\operatorname{ext}\left(\bigcup_{j\in J}A_j\right) \subseteq \bigcap_{j\in J}\operatorname{ext}(A_j)$.

(b) $e(\varphi \vee \psi) = e\varphi \wedge e\psi$, or $\operatorname{ext}(A \cup B) = \operatorname{ext}(A) \cap \operatorname{ext}(B)$.

(c) $e\left(\bigwedge_{j\in J} \varphi_j\right) \geq \bigvee_{j\in J}(e\varphi_j)$, or $\operatorname{ext}\left(\bigcap_{j\in J} A_j\right) \supseteq \bigcup_{j\in J} \operatorname{ext}(A_j)$

5. Lattice properties (boundary):

(a) $b(\varphi \vee \psi) \leq b\varphi \vee b\psi$, or $\partial(A \cup B) \subseteq \partial A \cup \partial B$

(b) $b = k \wedge kc$, or $\partial A = \operatorname{cl}(A) \cap \operatorname{cl}(X \setminus A)$

6. Other Identities and Inequalities:

(a) $\top\varphi = \top$, $\bot\varphi = \bot$ for all $\varphi \in \operatorname{End}(\mathcal{P}(X))$.

(b) $I\varphi = \varphi I = \varphi$ for all $\varphi \in \operatorname{End}(\mathcal{P}(X))$.

(c) $\varphi\top = \top$ for $\varphi \in \{c, k, i, e\}$ and $b\top = \bot$.

(d) $\varphi\bot = \bot$ for $\varphi \in \{c, k, i, e, b\}$.

(e) $c\bot = \top$ and $c\top = \bot$.

(f) $c^2 = I$.

(g) $ck = e = ic$, or $X \setminus \operatorname{cl}(A) = \operatorname{ext}(A) = \operatorname{int}(X \setminus A)$.

(h) $cb = i \vee e = eb$, or $X \setminus \partial A = \operatorname{int}(A) \cup \operatorname{ext}(A) = \operatorname{ext}(\partial A)$

(i) $k^2 = k$, or $\operatorname{cl}(\operatorname{cl}(A)) = \operatorname{cl}(A)$.

(j) $kb = b$, or $\operatorname{cl}(\partial A) = \partial A$.

(k) $kc = ci$.

(l) $ik = e^2$, or $\operatorname{int}(\operatorname{cl}(A)) = \operatorname{ext}(\operatorname{ext}(A))$.

(m) $i^2 = i$, or $\operatorname{int}(\operatorname{int}(A)) = \operatorname{int}(A)$.

(n) $ie = e$, or $\operatorname{int}(\operatorname{ext}(A)) = \operatorname{ext}(A)$.

(o) $ibi = ibe = ibk = \bot$, or more generally, the interior of the boundary of either an open set or a closed set is always empty.

(p) $ek = e$, or $\operatorname{ext}(\operatorname{cl}(A)) = \operatorname{ext}(A)$.

(q) $bc = b$, or $\partial(X \setminus A) = \partial A$.

(r) $bi, bk, b^2 \leq b$, or $\partial(\operatorname{int}(A)), \partial(\operatorname{cl}(A)), \partial\partial A \subseteq \partial A$.
Additionally, $\partial\partial A = \partial A$ iff $\operatorname{int}(\partial A) = \varnothing$

(s) $b^3 = b^2$, or $\partial\partial\partial A = \partial\partial A$

(t) $kck = kckckck$

(u) $e^4 = e^2$

(v) $ik = (ik)^2$

Proof. Part 1:

(a) Since $\text{int}(A)$ is open, and $\text{int}(A) \subseteq A \subseteq B$, we have $\text{int}(A) \subseteq \text{int}(B)$ by minimality.

(b) Similar argument as the previous one

(c) Using part (1b), we get $\text{ext}(A) = X \setminus \text{cl}(A) \supseteq X \setminus \text{cl}(B) = \text{ext}(B)$.

(d) This is just set-theoretic result.

Part 2:

(a) Let $p \in \text{int}(\bigcap_{j\in J} A_j)$, then there exists some open neighborhood U of p such that $U \subseteq \bigcap_{j\in J} A_j$. In other words, $U \subseteq A_j$ for each $j \in J$. From there, we have $p \in \text{int}(A_j)$ for all $j \in J$.

(b) Vice versa, if $p \in \text{int}(A) \cap \text{int}(B)$, then there exists some open neighborhoods $U \subseteq A$ and $V \subseteq B$ of p. Together, we get $U \cap V \subseteq A \cap B$ is an open neighborhood of p. Hence, $p \in \text{int}(A \cap B)$.

(c) Since $A_k \in \subseteq \bigcup_{j\in J} A_j$, by monotone property, we have $\text{int}(A_k) \subseteq \text{int}\left(\bigcup_{j\in J} A_j\right)$. Hence, $\bigcup_{j\in J} \text{int}(A_j) \subseteq \text{int}\left(\bigcup_{j\in J} A_j\right)$.

Part 3:

(a) Let $p \in \bigcup_{j\in J} \text{cl}(A_j)$, then there exists some $k \in J$ such that $p \in \text{cl}(A_k)$. Let U be a neighborhood of p. From part (3), we know that $U \cap A_k \neq \varnothing$, and so is $U \cap \bigcup_{j\in J} A_j \neq \varnothing$. But U is arbitrary, so we should have $p \in \text{cl}\left(\bigcup_{j\in J} A_j\right)$.

(b)

$$\begin{aligned}
\text{cl}(A \cup B) &= X \setminus \text{int}(X \setminus (A \cup B)) \\
&= X \setminus \text{int}((X \setminus A) \cap (X \setminus B)) \\
&= X \setminus (\text{int}(X \setminus A) \cap \text{int}(X \setminus B)) \\
&= (X \setminus \text{int}(X \setminus A)) \cup (X \setminus \text{int}(X \setminus B)) \\
&= \text{cl}(A) \cup \text{cl}(B)
\end{aligned}$$

(c) Since $\bigcap_{j\in J} A_j \subseteq A_j$, we have $\text{cl}\left(\bigcap_{j\in J} A_j\right) \subseteq \text{cl}(A_j)$ by monotonicity. Thus, $\text{cl}\left(\bigcap_{j\in J} A_j\right) \subseteq \bigcap_{j\in J} \text{cl}(A_j)$.

Part 4:

(a) Since $\text{ext}(A)$ and $\text{cl}(A)$ forms a partition on X, from the inclusion $\bigcup_{j\in J}\text{cl}(A_j) \subseteq \text{cl}\left(\bigcup_{j\in J} A_j\right)$ (part (3)), we get

$$\bigcap_{j\in J}\text{ext}(A_j) = \bigcup_{j\in J} X\setminus\text{cl}(A_j) = X\setminus\bigcup_{j\in J}\text{cl}(A_j) \supseteq X\setminus\text{cl}\left(\bigcup_{j\in J}A_j\right) = \text{ext}\left(\bigcup_{j\in J}A_j\right)$$

(b) Similarly, $\text{ext}(A\cup B) = X\setminus\text{cl}(A\cup B) = X\setminus(\text{cl}(A)\cup\text{cl}(B)) = (X\setminus\text{cl}(A))\cap(X\setminus\text{cl}(B)) = \text{ext}(A)\cap\text{ext}(B)$.

(c) Since $\bigcap_{j\in J}A_j \subseteq A_j$, we have $\text{ext}\left(\bigcap_{j\in J}A_j\right) \supseteq \text{ext}(A_j)$ by antitonicity. Hence, $\text{ext}\left(\bigcap_{j\in J}A_j\right) \supseteq \bigcup_{j\in J}\text{ext}(A_j)$.

Part 5:

1. Taking the complement of both side of the inclusion, we can instead prove the following one

$$\text{ext}\,(A\cup B)\cup\text{int}\,(A\cup B) \supseteq (\text{ext}(A)\cup\text{int}(A))\cap(\text{ext}(B)\cup\text{int}(B))$$

   Let p be a point in the RHS, then $p\in(\text{ext}(A)\cup\text{int}(A))$ and $p\in(\text{ext}(B)\cup\text{int}(B))$. Consider the following 2 cases

   (a) If either $p\in\text{int}(A)$ or $p\in\text{int}(B)$, WLOG, say $p\in\text{int}(A)$. Let U be an open neighborhood of p contained in A. Then U is also contained in $A\cup B$. From part (1), we get $p\in\text{int}\,(A\cup B)$.

   (b) If $p\in\text{ext}(A)$ *and* $p\in\text{ext}(B)$, then $p\in\text{ext}(A)\cap\text{ext}(B) = \text{ext}(A\cup B)$.

2. Apply part (3), we get $\partial A\subseteq\text{cl}(A)$ and $\partial A\subseteq\text{cl}(X\setminus A)$, so $\partial A\subseteq\text{cl}(A)\cap\text{cl}(X\setminus A)$. Vice versa, if $p\in\text{cl}(A)\cap\text{cl}(X\setminus A)$, then any neighborhood of p must intersect both A and $X\setminus A$, hence $p\in\partial A$.

Part 6: the first 6 are trivial.

(g) $ck = e$ is clear from partition property. As for $e = ic$, note that $p\in\text{ext}(A)$ iff $p\in U\cap A = \varnothing$ for some open set U, and iff $p\in U\subseteq X\setminus A$, which is just $p\in\text{int}(X\setminus A)$.

(h) $cb = i\vee e$ is also from partition property. For $cb = eb$, note that $X\setminus\partial A$ is open, so $\text{ext}(\partial A)\supseteq X\setminus\partial A$. On the other hand, $\text{ext}(\partial A)\cap\partial A = \varnothing$ by construction, so we get $X\setminus\partial A = \text{ext}(\partial A)$.

(i) $k^2 = k$ follows from $A = \text{cl}(A)$ iff A is closed.

(j) Same for $kb = b$ as the boundary of a set is always closed.

(k) Left-multiply and right-multiply c on both side of $ck = ic$, we get $kc = Ikc = c(ck)c = c(ic)c = ciI = ci$.

(l) $ik = iIk = (ic)(ck) = e^2$.

(m) $i^2 = i$ follows from $A = \text{int}(A)$ iff A is open.

(n) Same as above.

(o) Firstly, suppose U is open, we need to show that $\text{int}(\partial U) = \varnothing$. Since U is open, $\text{int}(U) = U$, and $\partial U = \text{cl}(U) \setminus U = \text{cl}(U) \cap (X \setminus U)$. Hence,

$$\text{int}(\partial U) = \text{int}(\text{cl}(U)) \cap \text{int}(X \setminus U) = \text{int}(\text{cl}(U)) \cap (X \setminus \text{cl}(U))$$

Now, note that $A = \text{int}(\text{cl}(U))$ is a subset of $\text{cl}(U)$, while $B = X \setminus \text{cl}(U)$ is disjoint from $\text{cl}(U)$. Thus, $\text{int}(\partial U) = A \cap B = \varnothing$.

Secondly, suppose C is closed, then $\partial C = C \setminus \text{int}(C) = C \cap (X \setminus \text{int}(C))$ and so

$$\text{int}(\partial C) = \text{int}(C) \cap \text{int}(X \setminus \text{int}(C)) = \text{int}(C) \cap (X \setminus \text{cl}(\text{int}(C)))$$

As $\text{cl}(\text{int}(C))$ contains $\text{int}(C)$, $X \setminus \text{cl}(\text{int}(C))$ must be disjoint from $\text{int}(C)$ due to it being a subset of $X \setminus \text{int}(C)$. Therefore, $\text{int}(\partial C) = \varnothing$.

(p) $ek = (ck)k = ck^2 = ck = e$.

(q) Since $e = ic$ (part 6g), $bc = c(i \vee ic)c = cic \wedge cic^2 = ci \wedge cic = c(i \vee ic) = c(cb) = Ib = b$ (part 6h).

(r) Note $b = k \wedge kc = k \wedge ci$ (part 5b and 6k), so

  (a) $bk = (k \wedge kc)k = k^2 \wedge kck = k \wedge k(ck)$. As $k \geq I$, we get $ck \leq cI = c$ and so $bk = k \wedge k(ck) \leq k \wedge kc = b$.

  (b) $bi = (k \wedge ci)i = ki \wedge ci^2 = ki \wedge ci = ki \wedge kc$. As $i \leq I$, we get $ki \leq kI = k$, and so $bi = ki \wedge kc \leq k \wedge kc = b$.

  (c) $b^2 = (k \wedge kc)b = kb \wedge kcb = b \wedge kcb \leq b$

Suppose $ib = \bot$, then $b^2 = (k \wedge ci)b = kb \wedge cib = b \wedge c\bot = b \wedge \top = b$. Translated back to the language of sets, we have

$$\partial(\partial A) = \text{cl}(\partial A) \cap (X \setminus \text{int}(\partial A)) = \partial A \cap (X \setminus \varnothing) = \partial A \cap X = \partial A$$

Vice versa, suppose $b^2 = b$, then $b = b^2 = b \wedge c(ib)$, which only happens if $c(ib) \geq b$, or $ib \leq cb = eb$. However $i \wedge e = \bot$ (as the interior and exterior of a set are always disjoint), which shows that $ib = ib \wedge eb = (i \wedge e)b = \bot b = \bot$.

(s) Since $ibk = \bot$ (part 6o), we can use previous part 6r to show that $b^2(kb) = b(kb)$, which results in $b^3 = b^2$. Translated to the language of sets, since $\text{int}(\partial(\text{cl}(A))) = \varnothing$, substituting ∂A to A, we get

$$\partial\partial(\text{cl}(\partial A)) = \partial(\text{cl}(\partial A)) \text{ or}$$
$$\partial\partial\partial A = \partial\partial A$$

(t) With $k(ckc)kck = k(ic^2)kck = kikck$ and $i(kck) \leq kck$, we get $kckckck \leq k(kck) = kck$. On the other hand,

$$\begin{aligned} ik &\leq k \\ k(ik) &\leq k^2 = k \\ ck(ik) &\geq ck \\ kck(ckc)k = kck(ik) &\geq kck \end{aligned}$$

since k is monotonic and c is antitonic. Thus, $kckckck = kck$.

(u) Multiply c to the previous equation in part (t) on the left, we get $(ck)^2 = (ck)^4$, or $e^2 = e^4$ (from part g)

(v) A result of part (l) and (u).

□

4.3.1 Denseness

A set A is

1. *Dense* (in X) if $X = \mathrm{cl}(A)$.
2. *Nowhere dense* (in X) if $\mathrm{int}(\mathrm{cl}(A)) = \varnothing$

Proposition 4.25.

1. Properties of dense sets:
   (a) A set is dense if and only if it meets with every non-empty open set.
   (b) Supersets of dense set is still dense.
   (c) The collection of *open* dense sets forms a filter on X (may contain only X).
   (d) The only dense set in a discrete space is the space itself.
   (e) Any nonempty subset of an indiscrete space is dense.
2. Properties of nowhere dense sets:
   (a) TFAE:
      i. A is nowhere dense.
      ii. The complement of A contains an open dense subset.
      iii. The closure of A is the boundary of some open set (in particular, the closure of A is exactly the boundary of $X \setminus \mathrm{cl}(A)$).
      iv. A is a subset of the boundary of some open set.
      v. For each non-empty open set U, there exists a non-empty open subset V of U that is disjoint from A.
   (b) The collection of nowhere dense sets forms an ideal on X.

Proof.
Part 1:

(a) Suppose A is dense, and let U be a non-empty open set. Take any $p \in U$, then since $\mathrm{cl}(A) = X$, U must meets A at some point (possibly p). Vice versa, suppose A meets with every non-empty open set, we let $p \in X$ and U be any open neighborhood of p. By assumption, $U \cap A = \varnothing$, so p must lie in the closure of A. In other words, $X \subseteq \mathrm{cl}(A)$ (since p is arbitrary), and so $X = \mathrm{cl}(A)$.

(b) Suppose $\mathrm{cl}(A) = X$, and $A \subseteq B$. Then by monotonicity, $\mathrm{cl}(A) \subseteq \mathrm{cl}(B) \subseteq X$, so $\mathrm{cl}(B)$ must be the whole space X, i.e. B is dense.

(c) By part (1b), we only need to show intersection of 2 open dense sets is still dense (the openness is implied by the axioms of topology). Suppose A, B are open dense sets, and U an arbitrary *non-empty* open set. Then $A \cap U$ is also non-empty and open (by part 1a), and $(A \cap B) \cap U = B \cap (A \cap U)$ is also non-empty and open. Hence, by equivalent definition, $A \cap B$ must be dense.

(d) If A is dense in a discrete space, then $A = \mathrm{cl}(A) = X$.

(e) Since the only non-empty closed set in an indiscrete space X is X itself, $\mathrm{cl}(A) = X$ for any non-empty A.

Part 2:

(a) We prove $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$. Note that $(iii) \Rightarrow (iv)$ is straight forward.

Suppose A is nowhere dense, then

$$\mathrm{int}(\mathrm{cl}(A)) = \varnothing \leftrightarrow X = X \setminus \mathrm{int}(\mathrm{cl}(A)) = \mathrm{cl}(X \,\mathrm{cl}(A)) = \mathrm{cl}(\mathrm{int}(X \setminus A))$$

Hence, $\mathrm{int}(X \setminus A)$ is dense and open. Since $\mathrm{int}(X \setminus A) \subseteq X \setminus A$, the complement of A contains an open dense subset.

Suppose $V \subseteq X \setminus A$ is an open dense set, then $\mathrm{int}(X \setminus A) = X \setminus \mathrm{cl}(A)$, which must contains V by maximality, is also open dense. Yet

$$X = \mathrm{cl}(X \setminus \mathrm{cl}(A)) = X \setminus \mathrm{int}(\mathrm{cl}(A))$$

Hence, $\mathrm{int}(\mathrm{cl}(A)) = \varnothing$, which results in

$$\partial(X \setminus \mathrm{cl}(A)) = \partial(\mathrm{cl}(A)) = \mathrm{cl}(\mathrm{cl}(A)) \setminus \mathrm{int}(\mathrm{cl}(A)) = \mathrm{cl}(A)$$

Suppose $A \subseteq \partial U$ for some open set U, then $\mathrm{int}(\mathrm{cl}(A)) \subseteq \mathrm{int}(\mathrm{cl}(\partial U)) = \mathrm{int}(\partial U)$ by monotonicity. However, $\mathrm{int}(\partial U) = \varnothing$ by Proposition 4.24 (part 6o), so $\mathrm{int}(\mathrm{cl}(A)) = \varnothing$. From there, we get $\mathrm{cl}(\mathrm{int}(X \setminus A)) = X \setminus \mathrm{int}(\mathrm{cl}(A)) =$

X. Denote $W = \text{int}(X \setminus A)$, which is a dense open set. For each non-empty open set U, let $V = U \cap W$ be an open subset of U. Since W is dense, V is not empty (part 1a), so we have found a non-empty open subset of any non-empty open set that is disjoint from A (as $V \subseteq W \subseteq X \setminus A$).

Suppose for each non-empty open set U, there exists a non-empty open subset V of U that is disjoint from A. Take $W = \text{int}(X \setminus A) = X \setminus \text{cl}(A)$, let U be any non-empty open subset of X. By the hypothesis, let V be a non-empty open subset of U that is disjoint from A (so $V \subseteq X \setminus A$). By maximality, $V \subseteq W$, so $W \cap U$ is not empty. By part 1a, W is dense, i.e. $\text{cl}(\text{int}(X \setminus A)) = X \setminus \text{int}(\text{cl}(A)) = X$, so $\text{int}(\text{cl}(A)) = \varnothing$, i.e. A is nowhere dense.

(b) We first show if $A \subseteq B$, and B is nowhere dense, then A is nowhere dense. Indeed, $\text{int}(\text{cl}(A)) \subseteq \text{int}(\text{cl}(B)) = \varnothing$ by monotonicity.

Next, we show that union of 2 nowhere dense sets A and B is again nowhere dense. Let U be any non-empty open set, then from part (2a), there exists a non-empty open subset V of U disjoint from A, and a non-empty open subset W of V disjoint B. Thus, W is disjoint from $A \cup B$ (as $W \subseteq V$, and $V \cap A = \varnothing$). By equivalent definition, $A \cup B$ must be nowhere dense set.

□

The *density* $d(X)$ of X is the least cardinality of a dense set in X.

Remark 4.26.

1. Since X is a dense set itself, the density of the space is $\leq |X|$.
2. Since the class of cardinal numbers is well-ordered, there must exist a dense subset of X with cardinality equal to the density.

Proposition 4.27. If $\mathcal{B}$ is a basis on X, then there exists a dense subset S of X such that $|S| \leq |\mathcal{B}|$. As a result, $d(X) \leq w(X)$.

Proof. Assuming axiom of choice, pick $x_B \in B$ for each $B \in \mathcal{B}$, and collect them into a set S. The map from $\mathcal{B}$ to S by sending B to x_B is then a surjection, so we must have $|\mathcal{B}| \leq |S|$. We now show that S is dense. For any $p \in X$ and N a neighborhood of p, there exists some $B \in \mathcal{B}$ such that $p \in B \subseteq N$. Since $x_B \in B$, $S \cap N$ is non-empty. By definition, $p \in \text{cl}(S)$, and so $\text{cl}(S) = X$. □

4.3.2 Regularity

An set U is *regularly open* if $U = \text{int}(\text{cl}(U))$. Likewise, a set C is *regularly closed* if $C = \text{cl}(\text{int}(C))$. As defined, a regularly open set is open, while a regularly closed set is closed.

Proposition 4.28.

1. Both X and $\varnothing$ are regular open and regular closed sets.
2. A set is regularly open iff its complement is regularly closed.
3. An open set U is regular iff $\partial \operatorname{cl}(U) = \partial U$.
4. A closed set C is regular iff $\partial \operatorname{int}(C) = \partial C$.

Proof. Part 1:

$$\begin{aligned} \operatorname{cl}(\operatorname{int}(X)) &= \operatorname{cl}(X) = X \\ \operatorname{cl}(\operatorname{int}(\varnothing)) &= \operatorname{cl}(\varnothing) = \varnothing \\ \operatorname{int}(\operatorname{cl}(X)) &= \operatorname{cl}(X) = X \\ \operatorname{int}(\operatorname{cl}(\varnothing)) &= \operatorname{cl}(\varnothing) = \varnothing \end{aligned}$$

Part 2: if U is regularly open, then $\operatorname{cl}(\operatorname{int}(X \setminus U)) = \operatorname{cl}(X \setminus \operatorname{cl}(U)) = X \setminus \operatorname{int}(\operatorname{cl}(U)) = X \setminus U$. Hence $X \setminus U$ is regularly closed. The reverse direction is similar.
Part 3: if U is regularly open, then $\partial \operatorname{cl}(U) = \operatorname{cl}(\operatorname{cl}(U)) \setminus \operatorname{int}(\operatorname{cl}(U)) = \operatorname{cl}(U) \setminus U = \partial U$. Vice versa, if $\partial U = \partial \operatorname{cl}(U)$, then

$$\operatorname{int}(U) = \operatorname{cl}(U) \setminus \partial U = \operatorname{cl}(\operatorname{cl}(U)) \setminus \partial \operatorname{cl}(U) = \operatorname{int}(\operatorname{cl}(U))$$

Since U is open we get $U = \operatorname{int}(U) = \operatorname{int}(\operatorname{cl}(U))$. Hence U is regular
Part 4: a combination of part (2) and (3). Indeed, suppose C is regularly closed. Then $X \setminus C$ is regularly open, and thus, $\partial C = \partial(X \setminus C) = \partial \operatorname{cl}(X \setminus C) = \partial(X \setminus \operatorname{int}(C)) = \partial \operatorname{int}(C)$. Vice versa, if $\partial C = \partial \operatorname{int}(C)$, then

$$C = \operatorname{cl}(C) = \partial C \cup \operatorname{int}(C) = \partial \operatorname{int}(C) \cup \operatorname{int}(\operatorname{int}(C)) = \operatorname{cl}(\operatorname{int}(C))$$

which shows that C is regular. □

A space X is called a *partition space* if there exists a basis on X such that it is also a partition on X.

Theorem 4.29. The followings are equivalent

1. X is a partition space.
2. Every closed set is regular.
3. Every open set is regular.
4. Every closed set is open.
5. Every open set is closed.

Proof. We show that (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (1).
*Suppose X is a partition space: let $\mathcal{B}$ be a basis in X such that it is also a partition on X. Given C a closed set, then $U = X \setminus C$ is open and so $U = \bigcup_{\substack{B \subseteq U \\ B \in \mathcal{B}}} B$. Yet $\mathcal{B}$ is a partition, so $C = X \setminus U = \bigcup_{\substack{B \not\subseteq U \\ B \in \mathcal{B}}} B = \bigcup_{\substack{B \subseteq C \\ B \in \mathcal{B}}} B$. By definition, C is open, and so $\mathrm{cl}(\mathrm{int}(C)) = \mathrm{cl}(C) = C$.

*Suppose every closed set is regular: since an open set is regular if and only if its complement, which is closed, is also regular, we conclude that every open set is regular.

*Suppose every open set is regular: let A be a *non-empty* subset of X. Denote $R = \mathrm{int}(X \setminus A)$, which is closed. Since $\mathrm{cl}(R \cup A) = \mathrm{cl}(R) \cup \mathrm{cl}(A) \supseteq R \cup \mathrm{cl}(A) = (X \setminus \mathrm{cl}(A)) \cup \mathrm{cl}(A) = X$, we know that $R \cup A$ is dense in X.

If $A \subseteq R$, then by minimality, $\mathrm{cl}(A) \subseteq R$. Hence, $X = \mathrm{cl}(R \cup A) = \mathrm{cl}(R)$. By regularity hypothesis, $X = \mathrm{int}(\mathrm{cl}(R)) = R \subseteq X \setminus A$, which forces A to be empty, a contradiction. Hence, $A \not\subseteq \mathrm{cl}(R)$, or equivalently, $A \cap (X \setminus \mathrm{cl}(R)) \neq \varnothing$. In particular, by setting $A = \{p\}$ to be a singleton, we get $p \in X \setminus \mathrm{cl}(\mathrm{int}(X \setminus \{p\}))$

Furthermore, $\mathrm{cl}(X \setminus \mathrm{cl}(R)) = X \setminus \mathrm{int}(\mathrm{cl}(R)) = X \setminus R = \mathrm{cl}(A)$. Let $A = \{p\}$, we get $\mathrm{cl}(\mathrm{int}(X \setminus \{p\}))) = \mathrm{cl}(\{p\})$

So given K a closed set, and any $p \in K$, we know that $\mathrm{cl}(\{p\}) \subseteq K$ and thus,

$$p \in X \setminus \mathrm{cl}(\mathrm{int}(X \setminus \{p\})) \subseteq \mathrm{cl}[X \setminus \mathrm{cl}(\mathrm{int}(X \setminus \{p\}))] = \mathrm{cl}(\{p\}) \subseteq K$$

In short, we have found an open neighborhood $U = X \setminus \mathrm{cl}(\mathrm{int}(X \setminus \{p\}))$ of p such that $U \subseteq K$. As p varies in K, we know that K is open.

*Suppose every closed set is open: by complementary, we know that every open set is closed.

*Suppose every open set is closed: let E_p be the intersection of all open set containing $p \in X$. By the hypothesis, E_p is closed, so $\mathrm{cl}(\{p\}) \subseteq E_p$ by minimality. Vice versa, if C is a closed containing p, then C is also open, as $X \setminus C$ will be both open and closed set. Hence, $E_p \subseteq C$ for all closed sets containing p. By definition, $E_p \subseteq \bigcap_{\substack{C \text{ closed} \\ p \in C}} C = \mathrm{cl}(\{p\})$.

Next, if $x \in E_p = \mathrm{cl}(\{p\})$, then $E_x = \mathrm{cl}(\{x\}) \subseteq E_p$ by minimality. On the other hand, for any open neighborhood V of x, we must have $V \cap \{p\} \neq \varnothing$, or $p \in V$ simply. Thus, p is in the E_x, the intersection of open neighborhoods of x. We then argue similarly that $E_p \subseteq E_x$, and so $E_p = E_x$ whenever $x \in E_p$.

It is then true that $\{E_p : p \in X\}$ forms a basis and a partition on X. Indeed

1. $\bigcup_{p \in X} E_p = \bigcup_{p \in X} \mathrm{cl}(\{p\}) \supseteq \bigcup_{p \in X} \{p\} = X$.

2. If $E_p \cap E_q$ is nonempty, let $x \in E_p \cap E_q$. By the previous observation, we get $E_p = E_x = E_q$.

3. Since E_p is closed, $X \setminus E_p$ is open. By the hypothesis, $X \setminus E_p$ is closed, and so E_p is open.

4. Finally, if U is open, and $p \in U$, then U is closed (by the hypothesis), and so $E_p = \mathrm{cl}(\{p\}) \subseteq \mathrm{cl}(U) = U$

□

4.3.3 Locally finite collection

A collection $\mathcal{A}$ of subsets of X is *locally finite* if there exists a neighborhood around each point that intersects finitely many sets of $\mathcal{A}$. This includes any finite family of subsets of X

Proposition 4.30. Suppose $\{A_j\}_{j\in J}$ is a locally finite collection, show that

1. $\mathrm{cl}\left(\bigcup_{j\in J} A_j\right) = \bigcup_{j\in J} \mathrm{cl}(A_j)$

2. $\mathrm{int}\left(\bigcap_{j\in J} A_j\right) = \bigcap_{j\in J} \mathrm{int}(A_j)$

3. $\mathrm{ext}\left(\bigcup_{j\in J} A_j\right) = \bigcap_{j\in J} \mathrm{ext}(A_j)$

Proof. Part 1: by Proposition 4.23, we just need to show that $\mathrm{cl}\left(\bigcup_{j\in J} A_j\right) \subseteq \bigcup_{j\in J} \mathrm{cl}(A_j)$. Let $p \notin \bigcup_{j\in J} \mathrm{cl}(A_j)$, then $p \notin \mathrm{cl}(A_j)$ for all $j \in J$. By local finiteness, there exists an open neighborhood V of p such that $Z = \{j \in J : A_j \cap U \neq \varnothing\}$ is finite. For each $j \in Z$, let U_j be some open neighborhood of p that is disjoint from A_j. Denote $W = V \cap \bigcup_{j\in Z} U_j$, which is open since it is an intersection of $|Z| + 1$ open sets.

We verify that W is disjoint with A_j for *all* $j \in J$. Indeed, consider 2 cases

1. $j \in Z$: since $W \subseteq U_j$, W is disjoint from A_j by construction of U_j.

2. $j \notin Z$: by definition, $V \cap A_j = \varnothing$. With $W \subseteq V$, we conclude that W is disjoint from A_j.

Thus, we have found, for each $p \notin \bigcup_{j\in J} \mathrm{cl}(A_j)$, an open neighborhood around it that is disjoint with $\bigcup_{j\in J} \mathrm{cl}(A_j)$. This means $\bigcup_{j\in J} \mathrm{cl}(A_j)$ is closed, and so, by minimality of closure, we get $\mathrm{cl}\left(\bigcup_{j\in J} A_j\right) \subseteq \bigcup_{j\in J} \mathrm{cl}(A_j)$ (note both side contains $\bigcup_{j\in J} A_j$).

Part 2: follows from part 1 directly

$$\begin{aligned}\operatorname{int}\left(\bigcap_{j\in J} A_j\right) &= X \setminus \operatorname{cl}\left(X \setminus \bigcap_{j\in J} A_j\right) = X \setminus \operatorname{cl}\left(\bigcup_{j\in J} X \setminus A_j\right) \\ &= X \setminus \bigcup_{j\in J} \operatorname{cl}(X \setminus A_j) = \bigcap_{j\in J} X \setminus \operatorname{cl}(X \setminus A_j) \\ &= \bigcap_{j\in J} \operatorname{int}(A_j)\end{aligned}$$

Part 3: similar to part 2

$$\begin{aligned}\operatorname{ext}\left(\bigcup_{j\in J} A_j\right) &= X \setminus \operatorname{cl}\left(\bigcup_{j\in J} A_j\right) = X \setminus \bigcup_{j\in J} \operatorname{cl}(A_j) \\ &= \bigcap_{j\in J} X \setminus \operatorname{cl}(A_j) = \bigcap_{j\in J} \operatorname{ext}(A_j)\end{aligned}$$

□

Proof Note: in part 1, we only pick a finite number of U_j so there's no need for axiom of choice.

4.3.4 Limit points. Isolated points

Let A be a subset of a topological space X. A point p in X is said to be

1. A *limit* point of A if each neighborhood U of p contains at least some point of A other than p itself. In other words, $A \cap (U \setminus \{p\}) \neq \varnothing$.
2. An *isolated* point of A if $p \in A$, and there exists a neighborhood U of p such that p is the only common point between U and A. In other words, $A \cap (U \setminus \{p\}) = \varnothing$.

The set of limit points is denoted with $\operatorname{lp}(A)$ (or A'), called the *derived set* of A. The set if isolated points is denoted with $\operatorname{ip}(A)$ (which is a subset of A).

Proposition 4.31.

1. $\operatorname{lp}(A)$ and $\operatorname{ip}(A)$ forms a partition on $\operatorname{cl}(A)$.
2. $\operatorname{lp}(A)$ and A covers $\operatorname{cl}(A)$.
3. A is closed iff $\operatorname{lp}(A) \subseteq A$.
4. If A is closed, then $\operatorname{lp}(A)$ is closed.
5. TFAE:

(a) p is a limit point of A.

(b) p lies in the closure of $A \setminus \{p\}$.

(c) p is a limit point of $A \setminus \{p\}$.

6. If $\{p\}$ is open, then p is an isolated point of any set containing it. We also have a partial converse, if p is an isolated point of the space, then $\{p\}$ is open.

Proof. Part 1: if p is both a limit point and an isolated point, then $p \in A$ and there exists some neighborhood U of p such that

1. U contains some point in A other than p.

2. The only point in common with U and A is p.

A contradiction! So $\text{lp}(A)$ and $\text{ip}(A)$ are disjoint. Now for each $p \in \text{cl}(A)$ since any neighborhood U of p must meet A at some point (i.e. $|U \cap A| \geq 1$), there are 2 cases

1. Any neighborhood U contains at least some point in A other than p. By definition, p is a limit point.

2. Some neighborhood U contains no point in A other than p. Since $U \cap A \neq \varnothing$, p has to be in the intersection, and so $U \cap A = \{p\}$. By definition, p is an isolated point.

Part 2: let $p \in \text{cl}(A)$, if $p \notin A$ then since each neighborhood U must meet A at some point, such point cannot be p. By definition, p has to be a limit point of A.
Part 3: by previous part (2), we have $\text{cl}(A) = A \cup \text{lp}(A) \subseteq A$. Yet $A \subseteq \text{cl}(A)$ always hold, so we obtain $A = \text{cl}(A)$, i.e. A is closed. Vice versa, if A is closed, then $A = \text{cl}(A) = A \cup \text{lp}(A)$, which holds only if $\text{lp}(A) \subseteq A$.
Part 4: since A is closed, we have $\text{lp}(A) \subseteq A$. We will show that $\text{lp}(\text{lp}(A)) \subseteq \text{lp}(A)$ (which in turns imply that $\text{lp}(A)$ is closed by part 2). Let $p \in \text{lp}(\text{lp}(A))$, and U be a neighborhood of p. Then U meets $\text{lp}(A)$ at some point other than p. Yet $\text{lp}(A) \subseteq A$, so we have U meets A at some point other than p. Hence, $p \in \text{lp}(A)$.
Part 5: we will prove in the direction $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$

1. Suppose p is a limit point of A, then $A \cap (U \setminus \{p\}) \neq \varnothing$ whenever U is a neighborhood of p. Note that $A \cap (U \setminus \{p\}) = A \cap [U \cap (X \setminus \{p\})] = [A \cap (X \setminus \{p\})] \cap U = (A \setminus \{p\}) \cap U$, we conclude that p must lie in $\text{cl}(A \setminus \{p\})$.

2. Suppose p lies in the closure of $A \setminus \{p\}$, let U be an open neighborhood of p. Then $U \cap (A \setminus \{p\}) \neq \varnothing$, note that any point in such intersection must be different from p due to $A \setminus \{p\}$. Hence, p must be a limit point of $A \setminus \{p\}$.

3. Suppose p is a limit point $A \setminus \{p\}$, then each neighborhood of p must contain some point, other than p, in $A \setminus \{p\} \subseteq A$. This means p is also a limit point of A.

Part 6: just take $U = \{p\}$ to be the open neighborhood of p. Vice versa, if p is an isolated point of X, then there exists an open neighborhood U of p such that $U \cap X = \{p\}$. But $U \subseteq X$, so $U = \{p\}$, i.e. $\{p\}$ is open. □

Proposition 4.32.

1. Properties of derived set:
   (a) $A \subseteq B$ implies $\mathrm{lp}(A) \subseteq \mathrm{lp}(B)$
   (b) $\mathrm{lp}(\mathrm{lp}(A)) \subseteq \mathrm{lp}(A) \cup A = \mathrm{cl}(A)$
   (c) $\bigcup_{i\in I} \mathrm{lp}(A_i) \subseteq \mathrm{lp}\left(\bigcup_{i\in I} A_i\right)$
   (d) $\mathrm{lp}(\mathrm{cl}(A)) = \mathrm{cl}(\mathrm{lp}(A))$
2. Properties of set of isolated points:
   (a) $\mathrm{ip}(A) \subseteq A$
   (b) $\mathrm{ip}(\mathrm{ip}(A)) = \mathrm{ip}(A)$
   (c) $\mathrm{ip}\left(\bigcup_{i\in I} A_i\right) \subseteq \bigcup_{i\in I} \mathrm{ip}(A_i)$

Proof. Part 1:

(a) Let $p \in \mathrm{lp}(A)$, and U be an open neighborhood of p. There exists some $q \in U \cap A$ that is distinct from p. But $A \subseteq B$, so $q \in U \cap B$. Thus, $p \in \mathrm{lp}(B)$.

(b) Let $p \in \mathrm{lp}(\mathrm{lp}(A))$ yet $p \notin A$, and U be an open neighborhood of p. Then there exists some $q \in U \cap \mathrm{lp}(A)$ such that $q \neq p$. As U is also a neighborhood of $q \in \mathrm{lp}(A)$, there exists another $r \in U \cap A$ such that $r \neq q$. Since $r \in A$ yet $p \notin A$, we must have $r \neq p$. Thus, there is always some point in $U \cap A$ other than (possibly) p. But U is arbitrary, so we get $p \in \mathrm{lp}(A)$. On the other hand, the equality $\mathrm{lp}(A) \cup A = \mathrm{cl}(A)$ comes from previous proposition.

(c) From part (1a), we already have $\mathrm{lp}(A_i) \subseteq \mathrm{lp}\left(\bigcup_{i\in I} A_i\right)$, or $\bigcup_{i\in I} \mathrm{lp}(A) \subseteq \mathrm{lp}\left(\bigcup_{i\in I} A_i\right)$.

(d) For the reverse inclusion, suppose $p \notin \mathrm{lp}(A \cup B)$, we will show that $p \notin \mathrm{lp}(A) \cup \mathrm{lp}(B)$. By the hypothesis, there is some neighborhood U, V of p such that $U \cap A$ and $V \cap B$ is either empty or contains only p. Taking $W = U \cap V$ to be neighborhood of p, then we get $W \cap (A \cup B)$ is either empty or contains only p (since $W \subseteq U, V$). In other words, $p \notin \mathrm{lp}(A \cup B)$.

(e) Applying part (1c), we get $\mathrm{lp}(\mathrm{cl}(A)) = \mathrm{lp}(\mathrm{lp}(A) \cup A) = \mathrm{lp}(\mathrm{lp}(A)) \cup \mathrm{lp}(A) = \mathrm{cl}(\mathrm{lp}(A))$.

Part 2:

(b) Note that $\mathrm{ip}(A) \subseteq A$ (by definition) for any set A, so $\mathrm{ip}(\mathrm{ip}(A)) \subseteq \mathrm{ip}(A)$. Vice versa, let $p \in \mathrm{ip}(A)$, then there exists some neighborhood N of p such that $N \cap A = \varnothing$. Since A contains $\mathrm{ip}(A)$, it also holds that $N \cap \mathrm{ip}(A) = \varnothing$. By definition, $p \in \mathrm{ip}(\mathrm{ip}(A))$.

(c) If $p \in \mathrm{ip}\left(\bigcup_{i\in I} A_i\right)$, then there exists a neighborhood N such that $N \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I} N \cap A_i = \{p\}$. In particular, there must be some $k \in I$ such that $N \cap A_k = \{p\}$ (otherwise, the union will be empty). By definition, $p \in \mathrm{ip}(A_k)$, hence $p \in \bigcup_{i\in I} \mathrm{ip}(A_i)$.

□

A set is *discrete* if every point in it is an isolated point (i.e. $\mathrm{ip}(A) = A$). On the opposite side, a set is *dense-in-itself* if no point in the set is an isolated point (i.e. $\mathrm{ip}(A) = \varnothing$). A *perfect* set is then a dense-in-itself, closed set.

Remark 4.33. A discrete set is not necessarily closed. Take $K = \{1/n : n \in \mathbb{N}\}$ in $\mathbb{R}$ as an example.

Proposition 4.34.

1. Subsets of a discrete set are discrete themselves.
2. Union of dense-in-itself sets is dense-in-itself.
3. Open subsets of a dense-in-itself space are dense-in-itself themselves.
4. A set is perfect iff every point in it is a limit point. In other words, A is perfect if and only if $\mathrm{lp}(A) = A$.
5. A set is dense-in-itself iff its closure is perfect. In particular, the closure of a dense-in-itself set is still dense-in-itself.

Proof.
Part 1: let A be a discrete set in X, and $B \subseteq A$. Then for each $p \in B$, there must exists some neighborhood N of p such that $N \cap A = \{p\}$. Since $p \in B \subseteq A$, we also get $N \cap B = \{p\}$. Therefore, p is an isolated point of B.
Part 2: let $\{A_i\}_{i\in I}$ be a family of dense-in-itself sets. Let $p \in \bigcup_{i\in I} A_i$, then $p \in A_k$ for some $k \in I$. In other words, for all neighborhood N of p, N meets A_k at another point other than p ($N \cap A_k$ cannot be empty since $p \in A_k$). Since $A_k \subseteq \bigcup_{i\in I} A_i$, N must also meet at $\bigcup_{i\in I} A_i$ another point other than p. Hence, p cannot be an isolated point of the union of A_i.
Part 3: suppose X is dense-in-itself, and U is open. If $p \in U$ is an isolated point, then there exists open neighborhood W of p such that $U \cap W = \{p\}$. Since finite intersection of open sets is open, $\{p\}$ is open in X. Therefore, $V = \{p\}$ is an open neighborhod of p such that $V \cap X = \{p\}$, i.e. p is an isolated point of X, contradicting the assumption.

Part 4: suppose A is perfect, then $A = \mathrm{cl}(A) = \mathrm{lp}(A) \cup \mathrm{ip}(A)$. But it is dense-in-itself, so it must be the case that $\mathrm{ip}(A) = \varnothing$, or $A = \mathrm{lp}(A)$. Vice versa, if $A = \mathrm{lp}(A)$, then A must be closed in particular, and so $\mathrm{lp}(A) = A = \mathrm{cl}(A) = \mathrm{lp}(A) \cup \mathrm{ip}(A)$. Yet $\mathrm{ip}(A) \cap \mathrm{lp}(A) = \varnothing$, so we get $\mathrm{ip}(A) = \varnothing$, or equivalently, A is dense-in-itself. Combined with the fact A is closed, we conclude A is perfect.
Part 5: if A is dense-in-itself, then $\mathrm{cl}(A) = \mathrm{lp}(A) \cup \mathrm{ip}(A) = \mathrm{lp}(A)$. On the other hand, $\mathrm{lp}(\mathrm{cl}(A)) = \mathrm{cl}(\mathrm{lp}(A))$ (for any subset A of X), so $\mathrm{lp}(\mathrm{cl}(A)) = \mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$. By part (4), $\mathrm{cl}(A)$ must be perfect. Vice versa, if $\mathrm{cl}(A) = \mathrm{lp}(A) = \mathrm{lp}(A) \cup \mathrm{ip}(A)$, then $\mathrm{ip}(A)$ must be empty (since $\mathrm{lp}(A)$ and $\mathrm{ip}(A)$ are disjoint). By definition, A is dense-in-itself. □

4.4 Convergence

Let $\mathcal{B}$ be a (non-empty) filter base on a topological space X. We say that $\mathcal{B}$ *converges* to a point x if for each neighborhood N of x, there exists some $B \in \mathcal{B}$ such that $B \subseteq N$. We denote $\mathcal{B} \to x$, or $x \in \lim \mathcal{B}$.
Note: if $\lim \mathcal{B}$ is a singleton $\{x\}$, then we can write $\lim \mathcal{B} = x$ instead of the convoluted set notation $x \in \lim \mathcal{B}$.

Remark 4.35. By definition of refinement, $\mathcal{B}$ converges to x if and only if it refines the neighborhood filter $\mathcal{N}_x$ of x. Since the refinement relation is reflexive, this also means that $\mathcal{N}_x \to x$ for each $x \in X$.

Remark 4.36. If $\mathcal{B}$ is a trivial filter base, i.e. $\varnothing \in \mathcal{B}$ (which can happen, for example, if $\mathcal{B}$ contains 2 sets that are disjoint), then it converges to any point in X. Thus, from now on, any filter base we mention is non-trivial (and thus also exclude, for example, the empty topological space).

Proposition 4.37.

1. Refinement: if $\mathcal{B} \to x$, and $\mathcal{C}$ refines $\mathcal{B}$, then $\mathcal{C} \to x$ also. In particular, $\lim \mathcal{C} \supseteq \lim \mathcal{B}$.
2. Generated Filter: If $\mathcal{B}$ is a filter base, and $\mathcal{F}$ is the filter (which is itself a filter base) generated by $\mathcal{B}$, then $\mathcal{B} \to x$ if and only if $\mathcal{F} \to x$.
3. Meet: if $\{\mathcal{B}_i\}_{i \in I}$ is a family of filter bases converging to some x, then there exists some other filter base $\mathcal{C}$, also converging to x, such that $\mathcal{B}_i$ refines $\mathcal{C}$.
4. Subsequentiality: let $x \in X$. If $\mathcal{B}$ is a filter base such that, whenever $\mathcal{C}$ refines $\mathcal{B}$, there exists another $\mathcal{D}$ refining $\mathcal{C}$ that converges to x, then $\mathcal{B}$ converges to x.

Proof.
Part 1: by assumption, for each neighborhood N of x, there exists $B \in \mathcal{B}$ and $C \in \mathcal{C}$ such that $B \subseteq N$ and $C \subseteq B$. Hence, there always exists $C \in \mathcal{C}$ such that $C \subseteq N$. As N varies, $\mathcal{C} \to x$.
Part 2: note that $\mathcal{B} \subseteq \mathcal{F}$ (and so $\mathcal{F}$ refines $\mathcal{B}$), $\mathcal{B} \to x$ implies $\mathcal{F} \to x$ by part (1). Vice versa, given $\mathcal{B}$ refines $\mathcal{F}$ as, by definition, there exists some $B \in \mathcal{B}$ such that $B \subseteq F$ whenever $F \in \mathcal{F}$, then $\mathcal{F} \to x$ implies $\mathcal{B} \to x$.
Part 3: let $\mathcal{B}_i \to x$ for each $i \in I$. We define $\mathcal{C} = \{C \subseteq X \mid \forall i \in I, \exists B \in \mathcal{B}_i : B_i \subseteq C\}$. This a filter base (and in fact, a filter) since

1. Because $\mathcal{B}_i$ does not contain empty set, each set C in $\mathcal{C}$ is non-empty since they need to contain some set in $\mathcal{B}_i$.
2. If $C_1, C_2 \in \mathcal{C}$, there exists, for each $i \in I$, $B_1^i, B_2^i \in \mathcal{B}$ such that $B_n^i \subseteq C_n$ $(n = 1, 2)$. Since $\mathcal{B}_i$ is a filter base, there exists another $B_i \in \mathcal{B}_i$ such that $B_i \subseteq B_1^i \cap B_2^i$. Thus, $B_i \subseteq B_1^i \cap B_2^i \subseteq C_1 \cap C_2$ for every $i \in I$, meaning $C_1 \cap C_2 \in \mathcal{C}$.

$\mathcal{C}$ is also coarser than $\mathcal{B}_i$ because every C is defined to contain some set in $\mathcal{B}_i$ for each $i \in I$. Finally, we also have $\mathcal{C} \to x$. Indeed, for each neighborhood N of x, there exists some $B_i \in \mathcal{B}_i$ such that $B_i \subseteq N$. By definition, $N \in \mathcal{C}$ and so $\mathcal{C}$ contains the neighborhood filter at x.
Part 4: suppose $\mathcal{B}$ does not converge to x. Equivalently, there exists some neighborhood N_0 of x such that $B \not\subseteq N_0$ for all $B \in \mathcal{B}$. In particular, $\mathcal{B}_0 = \{B \setminus N_0 : B \in \mathcal{B}\}$ does not contain empty set. Furthermore, $\mathcal{B}_0$ is a filter base that refines $\mathcal{B}$

1. If $B_1 \setminus N_0, B_2 \setminus N_0 \in \mathcal{B}_0$, there exists some $B \in \mathcal{B}$ such that $B \subseteq B_1 \cap B_2$. Therefore, $B \setminus N_0 \subseteq (B_1 \cap B_2) \setminus N_0 = (B_1 \setminus N_0) \cap (B_2 \setminus N_0)$.
2. For each $B \in \mathcal{B}$, we have $B \setminus N_0 \in \mathcal{B}_0$ and $B \setminus N_0 \subseteq B$.

Thus, by the hypothesis, there exists some $\mathcal{D}$ refining $\mathcal{B}_0$ that converges to x. In particular, pick any $B \in \mathcal{B}$, and there exists some $D_0, D_b \in \mathcal{D}$ such that $D_0 \subseteq N_0$ and $D_b \setminus B \setminus N_0$. There also some $D \in \mathcal{D}$ such that $D \subseteq D_0 \cap D_b \subseteq N_0 \cap (B \setminus N_0) = \varnothing$, contradicting the non-triviality of a filter base. □

Theorem 4.38 (Convergence as categorical limit)**.** Let X be a topological space (with neighborhood filter $\mathcal{N}_x$) and $\mathfrak{F}X$ be the set of filters on X. Denote $\mathfrak{F}_{\mathcal{F},x} = \{\mathcal{G} \in \mathfrak{F}X : \mathcal{G} \supseteq \mathcal{F} \cup \mathcal{N}_x\}$. Note that both $\mathfrak{F}X$, $\mathfrak{F}_{\mathcal{F},x}$, and any filter on X can be considered as categories (as they are partially ordered under inclusion).

Next, define $E_x : \mathfrak{F}_{\mathcal{F},x} \to \mathfrak{F}X$ to be the inclusion functor. Then $\mathcal{F} \to x$ if and only if $\mathcal{F}$ is the limit of E_x

Proof. Suppose $\mathcal{F}$ is a limit of E_x, then for any cone $\mathcal{N}$ to E_x, i.e. $\mathcal{N}$ is a filter such that $\mathcal{N} \subseteq \mathcal{G}$ for any $\mathcal{G} \in \mathfrak{F}_{\mathcal{F},x}$, one must have $\mathcal{N} \subseteq \mathcal{F}$. In other words, $\mathcal{N} \subseteq \bigcap_{\mathcal{F} \cup \mathcal{N}_x \subseteq \mathcal{G}} \mathcal{G} = \mathcal{W}$ implies $\mathcal{N} \subseteq \mathcal{F}$. By construction, $\mathcal{W}$ is the filter generated by $\mathcal{F}$ and neighborhoods of x. Taking $\mathcal{N} = \mathcal{W}$, we get $\mathcal{W} \subseteq \mathcal{F}$. Hence $\mathcal{N}_x \subseteq \mathcal{F}$ in particular. By monotonicity, $\mathcal{F} \to x$.

Vice versa, suppose $\mathcal{F} \to x$, then $\mathcal{F} \supseteq \mathcal{N}_x$ (Proposition 4.37), and so $\mathfrak{F}_{\mathcal{F},x} = \{\mathcal{G} \in \mathfrak{F}X : \mathcal{G} \supseteq \mathcal{F}\}$. Let $\mathcal{N}$ be a cone to E_x, i.e. $\mathcal{N}$ is a filter such that $\mathcal{N} \subseteq \mathcal{G}$ for all $\mathcal{G} \in \mathfrak{F}_{\mathcal{F},x}$ (or $\mathcal{G} \supseteq \mathcal{F}$). This means that $\mathcal{N} \subseteq \bigcap_{\mathcal{G} \supseteq \mathcal{F}} \mathcal{F} = \mathcal{F}$. By definition, $\mathcal{F}$ is the limit of $E_x : \mathfrak{F}_{\mathcal{F},x} \to \mathfrak{F}X$ □

4.4.1 Eventuality. Stationary

Given a filter base $\mathcal{B}$ and a subset S of X, we say that

1. S is $\mathcal{B}$-eventual, if there exists some $B \in \mathcal{B}$ such that $x \in S$ for all $x \in B$.
2. S is $\mathcal{B}$-frequent, or $\mathcal{B}$-stationary, if for each $B \in \mathcal{B}$, there exists some $x \in B$ such that $x \in S$.

Lemma 4.39.

1. "$\mathcal{B}$-eventual" implies "$\mathcal{B}$-frequent".
2. S is $\mathcal{B}$-eventual iff its complement is $\mathcal{B}$-frequent.
3. Monotonicity (filter): let $\mathcal{C}$ refines $\mathcal{B}$ (both are filter bases), then
   (a) S is $\mathcal{B}$-eventual implies it is $\mathcal{C}$-eventual.
   (b) S is $\mathcal{C}$-frequent implies it is $\mathcal{B}$-frequent.
4. Monotonicity (subset): let $S \subseteq T$ be subsets of X
   (a) S is $\mathcal{B}$-eventual implies T is $\mathcal{B}$-eventual.
   (b) S is $\mathcal{B}$-frequent implies T is $\mathcal{B}$-frequent.
5. Lattice properties (subset):
   (a) S and T are $\mathcal{B}$-eventual iff $S \cap T$ is $\mathcal{B}$-eventual.
   (b) S or T is $\mathcal{B}$-frequent iff $S \cup T$ is $\mathcal{B}$-frequent.
6. Lattice properties (filter): let $\{\mathcal{F}_\lambda\}_{\lambda \in L}$ be a family of filters on X.
   (a) S is $\mathcal{F}_\lambda$-eventual for every $\lambda \in L$ iff S is $\bigcap_{\lambda \in L} \mathcal{F}_\lambda$-eventual.
   (b) S is $\mathcal{F}_\lambda$-frequent for some $\lambda \in L$ iff S is $\bigcap_{\lambda \in L} \mathcal{F}_\lambda$-frequent.
7. Subsequentiality:
   (a) If $B \subseteq S$ for all $B \in \mathcal{B}$, then S is both $\mathcal{B}$-frequent and $\mathcal{B}$-eventual.
   (b) If S is $\mathcal{B}$-frequent, then there exists another filter base $\mathcal{C}$ refining $\mathcal{B}$ such that $C \subseteq S$ for all $C \in \mathcal{C}$.

Proof. Part 1: suppose S is $\mathcal{B}$-eventual, then there exists some $B_0 \in \mathcal{B}$ such that $B_0 \subseteq S$. Now for each $B \in \mathcal{B}$, there exists another $B' \in \mathcal{B}$ such that $B' \subseteq B \cap B_0$. Yet $B_0 \subseteq S$, so we know that $B' \subseteq B \cap S$, i.e. $B \cap S$ is not empty for all $B \in \mathcal{B}$ (again, as $\mathcal{B}$ is non-trivial).
Part 2: we have the following logical equivalence

$$\begin{aligned} &S \text{ is } \mathcal{B}\text{-eventual} \\ &\quad \Leftrightarrow \exists B \in \mathcal{B}, \forall x \in B : x \in S \\ &\quad \Leftrightarrow \exists B \in \mathcal{B}, \forall x \in B[\neg(x \notin S)] \\ &\quad \Leftrightarrow \neg[\forall B \in \mathcal{B}, \exists x \in B : x \in X \setminus S] \\ &\quad \Leftrightarrow \neg[X \setminus S \text{ is } \mathcal{B}\text{-frequent}] \end{aligned}$$

Part 3:

(a) Suppose S is $\mathcal{B}$-eventual: there exists $B_0 \in \mathcal{B}$ such that $B_0 \subseteq S$. Since $\mathcal{C}$ refines $\mathcal{B}$, there exists some $C_0 \in \mathcal{C}$ such that $C_0 \subseteq B_0$. Hence, $C_0 \subseteq S$ for some $C_0 \in \mathcal{C}$.

(b) Applying previous part (3a) and (2), we have:

$$\begin{aligned} S \text{ is } \mathcal{C}\text{-frequent} &\Leftrightarrow \neg[X \setminus S \text{ is } \mathcal{C}\text{-eventual}] \\ &\Rightarrow \neg[X \setminus S \text{ is } \mathcal{B}\text{-eventual}] \\ &\Leftrightarrow S \text{ is } \mathcal{B}\text{-frequent} \end{aligned}$$

Part 4:

(a) Suppose S is $\mathcal{B}$-eventual: there exists $B_0 \in \mathcal{B}$ such that $B_0 \subseteq S$. Since $S \subseteq T$, we also have $B_0 \subseteq T$.

(b) Apply previous part (4a) and (2), we have (note $X \setminus T \subseteq X \setminus S$):

$$\begin{aligned} S \text{ is } \mathcal{B}\text{-frequent} &\Leftrightarrow \neg[X \setminus S \text{ is } \mathcal{B}\text{-eventual}] \\ &\Rightarrow \neg[X \setminus T \text{ is } \mathcal{B}\text{-eventual}] \\ &\Leftrightarrow T \text{ is } \mathcal{B}\text{-frequent} \end{aligned}$$

Part 5:

(a) Suppose S and T are $\mathcal{B}$-eventual. This means there exists $B_s, B_t \in \mathcal{B}$ such that $B_s \subseteq S$ and $B_t \subseteq T$. Since $\mathcal{B}$ is a filter base, there exists $B_0 \in \mathcal{B}$ such that $B_0 \subseteq B_s \cap B_t \subseteq S \cap T$.

(b) Apply the previous part (5a) and (2), we get

$$\begin{aligned} &(S \text{ is } \mathcal{B}\text{-frequent}) \vee (T \text{ is } \mathcal{B}\text{-frequent}) \\ &\quad \Leftrightarrow \neg(X \setminus S \text{ is } \mathcal{B}\text{-eventual}) \vee \neg(X \setminus T \text{ is } \mathcal{B}\text{-eventual}) \\ &\quad \Leftrightarrow \neg[(X \setminus S \text{ is } \mathcal{B}\text{-eventual}) \wedge (X \setminus T \text{ is } \mathcal{B}\text{-eventual})] \\ &\quad \Leftrightarrow \neg[(X \setminus S) \cap (X \setminus T) \text{ is } \mathcal{B}\text{-eventual}] \\ &\quad \Leftrightarrow \neg[X \setminus (S \cup T) \text{ is } \mathcal{B}\text{-eventual}] \\ &\quad \Leftrightarrow S \cup T \text{ is } \mathcal{B}\text{-frequent} \end{aligned}$$

Part 6:

(a) Suppose S is $\mathcal{F}_\lambda$-eventual for all $\lambda \in L$. This means, for each $\lambda \in L$, there exists some $F \in \mathcal{F}_\lambda$ such that $F \subseteq S$. However, $\mathcal{F}_\lambda$ is a filter so $S \in \mathcal{F}_\lambda$, and so $S \in \bigcap_{\lambda \in L} \mathcal{F}_\lambda$. In particular, S is $\bigcap_{\lambda \in L} \mathcal{F}_\lambda$-eventual.

(b) Apply part (6a) and (2), we get

$$\begin{aligned} &\exists \lambda \in L : S \text{ is } \mathcal{F}_\lambda\text{-frequent} \\ &\quad \Leftrightarrow \exists \lambda \in L : \neg(X \setminus S \text{ is } \mathcal{F}_\lambda\text{-eventual}) \\ &\quad \Leftrightarrow \neg[\forall \lambda \in L : X \setminus S \text{ is } \mathcal{F}_\lambda\text{-eventual}] \\ &\quad \Leftrightarrow \neg \left[X \setminus S \text{ is } \bigcap_{\lambda \in L} \mathcal{F}_\lambda\text{-eventual} \right] \\ &\quad \leftrightarrow S \text{ is } \bigcap_{\lambda \in L} \mathcal{F}_\lambda\text{-frequent} \end{aligned}$$

Part 7:

(a) This follows from part (1) and (3) by first setting $\mathcal{I}_0 = \{S\}$ (assuming S is not empty) to be a filter base. Then S is $\mathcal{I}_0$-eventual, since the only option is S. From there, we conlcude S is also $\mathcal{I}_0$-frequent.

(b) Suppose S is $\mathcal{B}$-frequent. Define $\mathcal{B}_s = \{B \cap S : B \in \mathcal{B}\}$, then $\mathcal{B}_s$ is a filter base.

   (a) By assumption, $B \cap S$ is nonempty for all $B \in \mathcal{B}$.

   (b) Given $B_n \cap S \in \mathcal{B}_s$ $(n = 1, 2)$, then there exists some $B \in \mathcal{B}$ such that $B \subseteq B_1 \cap B_2$. We then have $B \cap S \subseteq (B_1 \cap B_2) \cap S = (B_1 \cap S) \cap (B_2 \cap S)$.

   We also have $B \cap S \subseteq S$ for every $B \cap S \in \mathcal{B}_s$, and $B \cap S \subseteq B$ for every $B \in \mathcal{B}$, so $\mathcal{B}_s$ refines $\mathcal{B}$ in particular.

□

Theorem 4.40 (Topology in terms of convergence)**.**

1. $x \in \mathrm{cl}(S)$ if and only if there exists a filter base $\mathcal{B}$ converging to x such that S is $\mathcal{B}$-frequent.

2. $x \in \mathrm{int}(S)$ if and only if S is $\mathcal{B}$-eventual whenever $\mathcal{B}$ converges to x.

Proof.
Part 1: suppose $x \in \mathrm{cl}(S)$, then the neighborhood filter $\mathcal{N}_x$ converges to x. Since $N \cap S$ is non-empty for all $N \in \mathcal{N}_x$ by equivalent definition of closure, S must be $\mathcal{B}$-frequent. Vice versa, suppose there exists a filter base $\mathcal{B}$ converging to x such that S is $\mathcal{B}$-frequent. Then for each neighborhood N of x, there exists some $B \in \mathcal{B}$ such that $B \subseteq N$. But $B \cap S$ is not empty, so $N \cap S$ is also not empty. By definition, $x \in \mathrm{cl}(S)$.
Part 2: we make use of the previous part (1) and the previous Lemma 4.39.

$$\begin{aligned} x \in \mathrm{int}(S) &\Leftrightarrow x \notin X \setminus \mathrm{int}(S) = \mathrm{cl}(X \setminus S) \\ &\Leftrightarrow \neg[\exists \mathcal{B} : \mathcal{B} \to x \wedge (X \setminus S \text{ is } \mathcal{B}\text{-frequent})] \\ &\Leftrightarrow \neg[\exists \mathcal{B} \text{ filter base} : (\mathcal{B} \to x) \wedge (X \setminus S \text{ is } \mathcal{B}\text{-frequent})] \\ &\Leftrightarrow \forall \mathcal{B} \text{ filter base} : \neg(\mathcal{B} \to x) \vee \neg(X \setminus S \text{ is } \mathcal{B}\text{-frequent}) \\ &\Leftrightarrow \forall \mathcal{B} \text{ filter base} : (\mathcal{B} \to x) \to (S \text{ is } \mathcal{B}\text{-eventual}) \end{aligned}$$

□

4.4.2 Cluster points

A point x in X is a *cluster* point of a filter base $\mathcal{B}$ if every neighborhood of x is $\mathcal{B}$-frequent. Equivalently, x is a cluster point of $\mathcal{B}$ iff $B \cap N$ is not empty for all $B \in \mathcal{B}$ and neighborhood N of x.

Remark 4.41. If $\{x\} \in \mathcal{B}$, then x is a cluster point.

Proof. Note that for any $B \in \mathcal{B}$, $B \cap \{x\}$ has to be not empty, so $\{x\}$ must be a subset of B. In other words, $x \in B$ for all $B \in \mathcal{B}$. Therefore, since any neighborhood of a point must contain that point, $B \cap N$ is not empty for all $B \in \mathcal{B}$ and neighborhood N of x. □

Proposition 4.42. TFAE

1. x is a cluster point of $\mathcal{B}$.
2. x is in the closure of B for each $B \in \mathcal{B}$.
3. x is a limit of some filter base $\mathcal{C}$ that refines $\mathcal{B}$.

Proof. We will show (1) ⇒ (2) ⇒ (3) ⇒ (1).

*Suppose x is a cluster points of $(x_i)_{i\in I}$: let $B \in \mathcal{B}$. Then $B \cap N$ is not empty for all neighborhood N of x. By definition, x is in the closure of B.

*Suppose x is in the closure of B for all $B \in \mathcal{B}$: this means that $B \cap N$ is not empty for all neighborhood N of x. Define $\mathcal{C} = \{B \cap N : B \in \mathcal{B} \wedge N \text{ is a neighborhood of } x\}$

1. $\varnothing \notin \mathcal{C}$ by the hypothesis.
2. If $B_1 \cap N_1, B_2 \cap N_2 \in \mathcal{C}$, then there exists some $B \in \mathcal{B}$ such that $B \subseteq B_1 \cap B_2$. Since $N = N_1 \cap N_2$ is another neighborhood of x, we have $B \cap N \in \mathcal{C}$ such that $B \cap N \subseteq (B_1 \cap B_2) \cap (N_1 \cap N_2) = (B_1 \cap N_1) \cap (B_2 \cap N_2)$.
3. $B \cap N \subseteq B, N$ for all $B \in \mathcal{B}$ and neighborhood N of x.

We conclude that $\mathcal{C}$ is a filter base that refines $\mathcal{B}$ and converges to x.

*Finally, suppose x is a limit of some filter base $\mathcal{C}$ that refines $\mathcal{B}$: for each $B \in \mathcal{B}$, there exists some $C_b \in \mathcal{C}$ such that $C_b \subseteq B$. Additionally, for each neighborhood N of x, there also exists some $C_n \in \mathcal{C}$ such that $C_n \subseteq N$. Yet $\mathcal{C}$ is a filter base, so there exists some *non-empty* $C \in \mathcal{C}$ such that $C \subseteq C_b \cap C_n \subseteq B \cap N$. But B and N are arbitrary (in their respective domains), so x must be a cluster point of $\mathcal{B}$ by definition. □

4.4.3 Filtered Sequence

[Reference: Where has this common generalization of nets and filters been written down? (Stack Exchange)]
A *$\mathcal{B}$-filtered sequence* $(x_i)_{i\in I}$ in X (indexed by I) is just a function $x : I \to X$ along with a filter base $\mathcal{B}$ on I.

Remark 4.43. For each $\mathcal{B}$-filtered sequence $(x_i)_{i\in I}$ in X, we can then form the *induced* filter base on X to be the image of $\mathcal{B}$ under x.

Vice versa, given a filter base $\mathcal{B}$ on X, we can form the *canonical* filtered sequence to be the pair $(\mathrm{id}_X, \mathcal{B})$ where id_X is the identity map on X.

With these transformations, we can declare a filtered sequence satisfies a (topological) property P (e.g. convergence, cluster, eventuality, stationary) if its canonical filter base also satifies P.

Remark 4.44. Let $(x_i)_{i\in I}$ be a $\mathcal{B}$-filtered sequence on X

1. x converges to p along $\mathcal{B}$ if and only if $x(\mathcal{B})$ refines the neighborhood filter at p.
2. p is a cluster point of $(x_i)_{i\in I}$ with respect to $\mathcal{B}$ if and only if for all $B \in \mathcal{B}$ and neighborhood N of p, there exists some $i \in B$ such that $x_i \in N$.
3. $(x_i)_{i\in I}$ is $\mathcal{B}$-eventually in $S \subseteq X$ if and only if there exists some $B_0 \in \mathcal{B}$ such that $x_i \in S$ for all $i \in B$.
4. Vice versa, $(x_i)_{i\in I}$ is $\mathcal{B}$-frequently in S if and only if for each $B \in \mathcal{B}$, there exists some $i \in B$ such that $x_i \in S$.

Given a $\mathcal{B}$-filtered sequence $(x_i)_{i\in I}$, there are 2 types of $\mathcal{C}$-filtered subsequences $(y_j)_{j\in J}$ to discuss

1. $(y_j)_{j\in J}$ is a *KW-subsequence* (stands for Kelley and Willard) if there exists a map $\varphi : J \to I$, called the re-indexing map, such that
   (a) $y_j = x_{\varphi(j)}$ for all $j \in J$.
   (b) $\varphi(\mathcal{C})$ is finer than $\mathcal{B}$.
2. $(y_j)_{j\in J}$ is an *AA-subsequence* (stands for Aarnes and Andenæs) if $y(\mathcal{C})$ refines $x(\mathcal{B})$ (as both are filter bases on X).

Remark 4.45.

1. The notion of AA-subsequences is derived directly from their corresponding (canonical) filter base on X.
2. Every KW-subsequence is AA-subsequence
3. Both notions of subsequences are (pre-)orderings, in that they satisfies reflexivity and transitivity.

Proof.
Part 2: if $\varphi : J \to I$ is the re-indexing map from $\mathcal{C}$-filtered sequence $(y_j)_{j\in J}$ to $\mathcal{B}$-filtered sequence $(x_i)_{i\in I}$, then $y = x \circ \varphi$, and $\varphi(\mathcal{C})$ refines $\mathcal{B}$. Since image under a function preserves refinement relation, we also have $y(\mathcal{C}) = x(\varphi(\mathcal{C}))$ refines $x(\mathcal{B})$, i.e. $(y_j)_{j\in J}$ is AA-subsequence of $(x_i)_{i\in I}$
Part 3: each filtered sequence is a KW-subsequence of itself by letting the identity map on the index set to serve as the reindexing map. It is also a AA-subsequence of itself since the refinement relation is reflexive.

The transitivity of AA-subsequence follows from the transitivity of refinement. As for KW-subsequence, let $(y_j)_{j\in J}$ be a $\mathcal{C}$-filtered subsequence of $(x_i)_{i\in I}$ (with $\varphi : J \to I$ as the re-indexing map), and $(z_k)_{k\in K}$ be a $\mathcal{D}$-filtered subsequence of $(y_j)_{j\in J}$ (with $\psi : K \to J$ as the re-indexing map). Consider $\lambda = \varphi \circ \psi$, which is a function from $K \to I$

1. $z = y \circ \psi = (x \circ \varphi) \circ \psi = x \circ (\varphi \circ \psi) = x \circ \lambda$.
2. Since $\psi(\mathcal{D})$ refines $\mathcal{C}$, $\varphi(\psi(\mathcal{D}))$ refines $\varphi(\mathcal{C})$. However $\varphi(\mathcal{C})$ refines $\mathcal{B}$, so $\lambda(\mathcal{D})$ must refine $\mathcal{B}$.

Hence, $(z_k)_{k\in K}$ is a KW-subsequence of $(x_i)_{i\in I}$, with $\lambda : K \to I$ serves as the re-indexing map. □

Proposition 4.46 (Equivalence of AA-subsequence)**.** Let $(x_i)_{i\in I}$ be a $\mathcal{B}$-filtered sequence and $(y_j)_{j\in J}$ be a $\mathcal{C}$-filtered sequence. TFAE

1. $(y, \mathcal{C})$ is an AA-subsequence of $(x, \mathcal{B})$.
2. For any $S \subseteq X$, $(y_j)_{j\in J}$ is $\mathcal{C}$-frequently in S implies $(x_i)_{i\in I}$ is $\mathcal{B}$-frequently in S.
3. For any $S \subseteq X$, $(x_i)_{i\in I}$ is $\mathcal{B}$-eventually in S implies $(y_j)_{j\in J}$ is $\mathcal{C}$-eventually in S.
4. If $L \subseteq I$ is $\mathcal{B}$-eventual (i.e. L contains some $B \in \mathcal{B}$), then $y^{-1}(x(L))$ is $\mathcal{C}$-eventual.

Proof. We show (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (1).

*Suppose $(y, \mathcal{C})$ is an AA-subsequence of $(x, \mathcal{B})$: then $y(\mathcal{C})$ refines $x(\mathcal{B})$. Note that $(x_i)_{i\in I}$ is $\mathcal{B}$-frequently in S iff S is $x(\mathcal{B})$-eventual. Hence, by Lemma 4.39, we have $(y_j)_{j\in J}$ is $\mathcal{C}$-frequently in S (or S is $y(\mathcal{C})$-frequent), then $(x_i)_{i\in I}$ is $\mathcal{B}$-frequently in S (or S is $x(\mathcal{B})$-frequent).

*Suppose $(x_i)_{i\in I}$ is $\mathcal{B}$-frequently in S whenever $(y_j)_{j\in J}$ is $\mathcal{C}$-frequently in S: again, from Lemma 4.39, we know that $(x_i)_{i\in I}$ is $\mathcal{B}$-eventually in S iff $(x_i)_{i\in I}$ is not $\mathcal{B}$-frequently in $X \setminus S$, which in turn implies $(y_j)_{j\in J}$ is not $\mathcal{C}$-frequently in $X \setminus S$, and so $(y_j)_{j\in J}$ is $\mathcal{C}$-eventually in S.

*Suppose $(y_j)_{j\in J}$ is $\mathcal{C}$-eventually in S whenever $(x_i)_{i\in I}$ is $\mathcal{B}$-eventually in S: Let $L \subseteq I$ be such that $B \subseteq L$ for some $B \in \mathcal{B}$. then $(x_i)_{i\in I}$ is $\mathcal{B}$-eventually in $x(L)$. Therefore, $(y_j)_{j\in J}$ is $\mathcal{C}$-eventually in $x(L)$ also. Equivalently, there exists some $C \in \mathcal{C}$ such that $y(C) \subseteq x(L)$, or $C \subseteq y^{-1}(x(L))$, or $y^{-1}(x(L))$ is $\mathcal{C}$-eventual.

*Suppose $y^{-1}(x(L))$ is $\mathcal{C}$-eventual whenever L is $\mathcal{B}$-eventual: note that $B \in \mathcal{B}$ is $\mathcal{B}$-eventual, so $y^{-1}(x(B))$ is $\mathcal{C}$-eventual, i.e. $C \subseteq y^{-1}(x(B))$ or $y(C) \subseteq x(B)$ for some $C \in \mathcal{C}$. By definition, $y(\mathcal{C})$ refines $x(\mathcal{B})$. □

While a KW-subsequence is more natural in functional and intuitive sense, the AA-subsequence is more natural in categorical sense. However, with respect to the filter bases they map to, they are "equivalent". With respect to the pre-ordering above, we say the 2 filtered sequences is *KW-equivalence* (resp. *AA-equivalence*) if each is a KW-subsequence (resp. AA-subsequence) of the other one.

Lemma 4.47. Given $\mathcal{B}$-filtered sequence $(x_i)_{i\in I}$ and $\mathcal{C}$-filtered sequence $(y_j)_{j\in J}$, TFAE

1. For all $B \in \mathcal{B}$ and $C \in \mathcal{C}$, there exists some $p \in X$ such that $x_i = y_j = p$ for some $i \in B, j \in C$.
2. $\mathcal{M} = \{S \subseteq X : S \supseteq x(B) \cap y(C) \text{ for some } B \in \mathcal{B}, C \in \mathcal{C}\}$ is a proper filter.
3. $x(\mathcal{B})$ and $y(\mathcal{C})$ have a common filter base refining both.
4. x and y have a common KW-subsequence. In fact, such subsequence can be chosen to be a maximal AA-subsequence.

Proof. We prove the following direction (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (1).

*Suppose there exists $i \in B, j \in C$ such that $x_i = y_j$ whenever $B \in \mathcal{B}, C \in \mathcal{C}$: equivalently, $x(B) \cap y(C) \neq \varnothing$. Hence, every set in $\mathcal{M}$ is non-empty. $\mathcal{M}$ is then a proper filter since

1. If $S \in \mathcal{M}$, then $x(B) \cap y(C) \subseteq S$ for some $B \in \mathcal{B}, C \in \mathcal{C}$. If $S \subseteq T$, then we also have $x(B) \cap y(C) \subseteq T$, and so $T \in \mathcal{M}$.
2. On the other hand, if $S_n \in \mathcal{M}$ $(n = 1, 2)$, then $x(B_n) \cap y(C_n) \subseteq S_n$ for some $B_n \in \mathcal{B}, C_n \in \mathcal{C}$. Since $\mathcal{B}, \mathcal{C}$ a filter base, there also exists some $B \in \mathcal{B}, C \in \mathcal{C}$ such that $B \subseteq B_1 \cap B_2, C \subseteq C_1 \cap C_2$. Let $S = x(B) \cap y(C) \in \mathcal{M}$, which then satisfies $S \subseteq x(B_1 \cap B_2) \cap y(C_1 \cap C_2) \subseteq x(B_1) \cap x(B_2) \cap y(C_1) \cap y(C_2) \subseteq S_1 \cap S_2$.

*Suppose $\mathcal{M}$ is a proper filter: in particular $x(B) \cap y(C)$ is not empty since it is in $\mathcal{M}$. Additionally $x(B) \cap y(C) \subseteq x(B)$ and $x(B) \cap y(C) \subseteq y(C)$, so $\mathcal{M}$ refines both $x(\mathcal{B})$ and $y(\mathcal{C})$.

*Suppose $\mathcal{F}$ is a filter base on X that refines $x(\mathcal{B})$ and $y(\mathcal{C})$: let $K = \{(i, j) \in I \times J : x_i = y_j\}$ and $\mathcal{D} = \{(B \times C) \cap K : B \in \mathcal{B}, C \in \mathcal{C}\}$.

1. By the hypothesis, for all $B \in \mathcal{B}$, $C \in \mathcal{C}$, there exists $F_b, F_c, F \in \mathcal{F}$ such that $F_b \subseteq x(B)$, $F_c \subseteq y(C)$, and $F \subseteq F_b \cap F_c \subseteq x(B) \cap y(C)$. In particular, K is not empty, since we can take some $(i, j) \in I \times J$ such that $p = x_i = y_j \in F$. This also shows that $\mathcal{D}$ is not empty, and contains no empty set.

2. $\mathcal{D}$ is a filter base on K: given $D_n = (B_n \times C_n) \cap K$ in $\mathcal{D}$ $(n = 1, 2)$, let $B \in \mathcal{B}$ and $C \in \mathcal{C}$ such that $B \subseteq B_1 \cap B_2$ and $C \subseteq C_1 \cap C_2$. We then have $D \in \mathcal{D}$ such that $D = (B \times C) \cap K \subseteq [(B_1 \cap B_2) \times (C_1 \cap C_2)] \cap K = [(B_1 \times C_1) \cap (B_2 \times C_2)] \cap K = D_1 \cap D_2$.

Let $z : K \to X$ where (i, j) is sent to $x_i = y_j \in X$, and consider the projection $\pi_I : K \to I$ and $\pi_J : K \to J$

1. $z_{ij} = x_i = x_{\pi_I(i,j)}$ and $z_{ij} = y_j = y_{\pi_J(i,j)}$.
2. We also have $\pi_I((B \times C) \cap K) \subseteq \pi_I(B \times C) = B$ and similarly, $\pi_J((B \times C) \cap K) \subseteq C$ for all $B \in \mathcal{B}, C \in \mathcal{C}$. Thus, π_I and π_J are the reindexing map from z to x and y respectively.

We conclude that z is a common KW-subsequence of x and y. Furthermore, z is a maximal AA-subsequence among common subsequence of x and y. Indeed, if $(w_l)_{l \in L}$ is a $\mathcal{E}$-filtered AA-subsequence of both x and y, then $w(\mathcal{E})$ refines both $x(\mathcal{B})$ and $y(\mathcal{C})$. In particular, for each $B \in \mathcal{B}$ and $C \in \mathcal{C}$, there exists some $E_b, E_c, E \in \mathcal{E}$ such that $w(E_b) \subseteq x(B)$, $w(E_c) \subseteq y(C)$, and $E \subseteq E_b \cap E_c$. Together, we get $w(E) \subseteq w(E_b) \cap w(E_c) \subseteq x(B) \cap y(C)$ for some $E \in \mathcal{E}$. Yet $z((B \times C) \cap K) = \{p \in X : x_i = y_j = p \text{ for some } i \in B, j \in C\} = x(B) \cap y(C)$, so as B, C varies, $w(\mathcal{E})$ refines $z(\mathcal{D})$ by definition.

*Finally, suppose $(z_k)_{k \in K}$ is a common $\mathcal{D}$-filtered KW-subsequence of x and y: then z is also AA-subsequence of x and y, and so $z(\mathcal{D})$ refines both $x(\mathcal{B})$ and $y(\mathcal{C})$. Again, by similar argument as above, we can show that $x(B) \cap y(C)$ is not empty (by containing some $z(D)$) for all $B \in \mathcal{B}, C \in \mathcal{C}$. □

Theorem 4.48. Any AA-subsequence of $(x_i)_{i \in I}$ is AA-equivalent to a KW-subsequence of $(x_i)_{i \in I}$.

Proof. If $(y_j)_{j \in J}$ is an AA-subsequence of $(x_i)_{i \in I}$, then $y(\mathcal{C})$ refines $x(\mathcal{B})$. In particular, given $B \in \mathcal{B}$ and $C \in \mathcal{C}$, there exists $C_b, C'\mathcal{C}$ such that $y(C_b) \subseteq x(B)$ and $C' \subseteq C_b \cap C$. Thus, $y(C') \subseteq y(C_b) \cap y(C) \subseteq y(C) \cap x(B)$, and so $y(C) \cap x(B)$ is not empty. By the previous lemma, there exists a common KW-subsequence z of both x and y that is also a maximal AA-subsequence. In other words, y must be an AA-subsequence of z. Since a KW-subsequence is also an AA-subsequence, we conclude z is AA-equivalent to y. □

Theorem 4.49 (Transformational Adjunction I)**.** Let $(x_i)_{i \in I}$ be a $\mathcal{B}$-filtered sequence in X, and $\mathcal{F}$ be a filter base on X. Then

$\mathcal{F}$ refines the induced filter base of $(x, \mathcal{B})$ if and only if the canonical sequence of $\mathcal{F}$ is an AA-subsequence to $(x, \mathcal{B})$.

Additionally, the induced filter base of the canonical sequence of $\mathcal{F}$ is just $\mathcal{F}$!

Proof. Suppose the canonical sequence of $\mathcal{F}$ is a KW-subsequence (or more generally AA-subsequence) of $(x, \mathcal{B})$, then $\mathcal{F} = \mathrm{id}_X(\mathcal{F})$ refines $x(\mathcal{B})$. Vice versa, if $\mathcal{F}$ refines $x(\mathcal{B})$, then $\mathrm{id}_X(\mathcal{F})$ refines $x(\mathcal{B})$, and so $(\mathrm{id}_X, \mathcal{F})$ is an AA-subsequence of $(x, \mathcal{B})$.

Since the canonical filtered sequence of $\mathcal{F}$ is $(\mathrm{id}_X, \mathcal{F})$, its induced filter base is $\mathrm{id}_X(\mathcal{F}) = \mathcal{F}$. □

4.4.4 Net on topological space

A *net* on X is a map $x : A \to X$ from a directed set A (i.e. A is a pre-ordered set where there is an upper bound to every 2 elements in A).

Remark 4.50. If $A = \mathbb{N}$, then a net turns into a sequence (of points in X)!

Remark 4.51. Given a net $(x_\alpha)_{\alpha \in A}$, the *eventuality filter base* on X (with respect to x) is defined as the image of $\mathcal{E}_A$ (under x) where

$$S \in \mathcal{E}_A \text{ if and only if } S = \uparrow \eta = \{\alpha \in A : \alpha \geq \eta\} \text{ for some } \eta \in A.$$

Vice versa, given a filter base $\mathcal{B}$ on X, let

1. $\langle X, \mathcal{B} \rangle$ be the subcollection of $X \times \mathcal{B}$ containing exactly (p, B) such that $p \in B$.
2. Ordering on $\langle X, \mathcal{B} \rangle$: $(p_1, B_1) \leq (p_2, B_2)$ iff $B_1 \supseteq B_2$.

Then the projection $\pi_{\mathcal{B}} : \langle X, \mathcal{B} \rangle \to X, (p, B) \mapsto p$ forms the canonical net of $\mathcal{B}$.

With these transformation, we again declare that a net has property P if its eventuality filter base on X has property P

Remark 4.52.

1. A net $(x_\alpha)_{\alpha \in A}$ converges to p if for each neighborhood N of p, there is some $\eta \in A$ such that $x_\alpha \in N$ for all $\alpha \geq \eta$.
2. p is a cluster point of $(x_\alpha)_{\alpha \in A}$ if for each neighborhood N of p and $\eta \in A$, there exists some $\alpha \geq \eta$ such that $x_\alpha \in N$.
3. $(x_\alpha)_{\alpha \in A}$ is eventually in S if there exists some $\gamma \in A$ such that $x_\alpha \in S$ for all $\alpha \geq \gamma$.
4. $(x_\alpha)_{\alpha \in A}$ is frequently in S if for each $\alpha \in A$, there exists some $\lambda \geq \alpha$ such that $x_\lambda \in S$.

Remark 4.53. Given a net $(x_\alpha)_{\alpha \in A}$ in X, if $\gamma \in A$ is a maximal element, then x_γ is a limit of the net.

Proof. Indeed, note that $\{\alpha \in A : \alpha \geq \gamma\} = \{\gamma\}$ by definition, so for any neighborhood N of x_γ, we have $x_\alpha \in N$ for all $\alpha \in \gamma$. □

Note: if A contains a maximal element, then such element is also the greatest element of A.

[Reference: `http://thales.doa.fmph.uniba.sk/sleziak/texty/rozne/topo/subnets2.pdf`] Given a net $x : A \to X$, there are 3 types of subnet in general

1. A *Willard subnet* is a map $y : B \to X$ such that there exists a cofinal monotone map $\varphi : B \to A$, and $y = x \circ \varphi$. Note that as φ is monotone, the cofinality condition is equivalent to
$$\forall \alpha \in A, \exists \beta \in B : \varphi(\beta) \geq \alpha$$
2. A *Kelley subnet* is like Willard subnet, but remove monotone condition on φ.
3. Finally, an *AA-subnet* (initials of the 2 authors Aarnes and Andenæs) is a map $y : B \to X$ such that the eventuality filer base $\mathcal{E}[y]$ refines $\mathcal{E}[x]$.

Remark 4.54. Straight from definition, we have the following inclusion: Willard subnets $\subseteq$ Kelley subnets $\subseteq$ AA-subnets.

Proof. The first inclusion is obvious, so we verify the second one. Let $y : B \to X$ be a Kelley subnet of $x : A \to X$, with $\varphi : B \to A$ be a cofinal map so that $y = x \circ \varphi$. Given $\eta \in A$, then there exists some $\gamma \in B$ such that $\varphi(\beta) \geq \eta$ for all $\beta \geq \gamma$. In other words,
$$\{y_\beta : \beta \geq \gamma\} = \{x_{\varphi(\beta)} : \beta \geq \gamma\} \subseteq \{x_\alpha : \alpha \geq \eta\}$$
This shows that $y(\mathcal{E}_B)$ refines $x(\mathcal{E}_A)$. □

Proposition 4.55 (Ordering of subnets)**.** The AA-subnet relation between nets is a pre-order on them. Same holds for other types of subnet.

Proof. For AA-subnet: by definition, $\mathcal{E}[x]$ refines $\mathcal{E}[y]$, which in turns refines $\mathcal{E}[z]$. Because refinement relation is transitive, we know that $\mathcal{E}[x]$ refines $\mathcal{E}[z]$. Additionally also refines itself, so every net is an AA-subnet of itself.

For Kelley subnet: let $\varphi : A \to B$ and $\psi : B \to C$ be monotone and cofinal maps such that $x = y \circ \varphi$ and $y = z \circ \psi$. Thus, $x = z \circ (\psi \circ \varphi)$. We will verify that $\psi \circ \varphi$ is cofinal. For each $\gamma_0 \in C$, there exists $\beta_0 \in B$ such that $\psi(\beta) \geq \gamma_0$ for all $\beta \geq \beta_0$. With such $\beta_0 \in B$, there then exists some $\alpha_0 \in A$ such that $\varphi(\alpha) \geq \beta_0$ whenever $\alpha \geq \alpha_0$. Therefore, $\psi(\varphi(\alpha)) \geq \gamma$ for all $\alpha \geq \alpha_0$.
On the other hand, given $x : A \to X$ a net, then $\mathrm{id}_A : A \to A$ is cofinal (just pick $\beta_0 = \alpha_0$ for any $\alpha_0 \in A$ to get $\mathrm{id}_A(\beta) = \beta \geq \beta_0 = \alpha_0$ for all $\beta \geq \beta_0$), and $x = x \circ \mathrm{id}_A$.

For Willard subnet: just copy the cofinal proof. The monotone one come from the fact that pre-ordered sets with monotone maps form a category. Explicitly,

1. Whenever $\alpha_1 \leq \alpha_2$, we get $\varphi(\alpha_1) \leq \varphi(\alpha_2)$ and $\psi(\varphi(\alpha_1)) \leq \psi(\varphi(\alpha_2))$.
2. The identity map on any pre-ordered set is always monotone: $\mathrm{id}_A(\alpha_1) = \alpha_1 \leq \alpha_2 = \mathrm{id}_A(\alpha_2)$ whenever $\alpha_1 \leq \alpha_2$

□

Proposition 4.56 (Equivalent definition of AA-subnet)**.** Let $x : A \to X$, $y : B \to X$ be nets. TFAE:

1. y is an AA-subnet of x.
2. A frequent subset with respect to y is also frequent with respect to x.
3. An eventual subset of with respect to x is also eventual with respect to y
4. If $D \subseteq A$ is eventual, i.e. $D \supseteq \{\alpha \in A : \alpha \geq \eta\}$ for some $\eta \in A$, then $y^{-1}(x(D))$ is eventual in B.

Proof. We show (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (1)

*Suppose y is an AA-subnet of x: let T be a frequent subset of X with respect to y and η be an element of A. Since $\mathcal{E}[y]$ refines $\mathcal{E}[x]$, there exists some $\gamma \in B$ such that $\{y_\beta : \beta \geq \gamma\} = \{x_{\varphi(\beta)} : \beta \geq \gamma\} \subseteq \{x_\alpha : \alpha \geq \eta\}$. By assumption, there is some $\beta_T \geq \gamma$ such that $y_{\beta_T} = x_{\varphi(\beta_T)} \in T$. In other words, there is some $\alpha_T = \varphi(\beta_T) \geq \eta$ such that $x_{\alpha_T} \in T$. By definition, T is frequent with respect to x.

*Suppose any frequent subset with respect to y is also frequent with respect to x: let K be a x-eventual subset of X. Apply Lemma 4.39, $X \setminus K$ must not be x-frequent. By the hypothesis, $X \setminus K$ is also not y-frequent. Equivalently, K is y-eventual.

*Suppose any eventual subset of with respect to x is also eventual with respect to y: let $D \subseteq A$ be eventual. Then $x(D)$ is x-eventual subset of X. Indeed, since D contains $\{\alpha \in A : \alpha \geq \eta\}$ for some $\eta \in A$, $x(D)$ contains $\{x_\alpha : \alpha \geq \eta\}$. By the hypothesis, $x(D)$ is also y-eventual, so there must exists some $\gamma \in B$ such that $y_\beta \in x(D)$ for all $\beta \geq \gamma$, or simply, $x(D)$ contains $\{y_\beta : \beta \geq \gamma\}$. Taking the inverse image (under y) on both side we get $y^{-1}(x(D))$ contains $y^{-1}\{y_\beta : \beta \geq \gamma\} \supseteq \{\beta \in B : \beta \geq \gamma\}$, i.e. $y^{-1}(x(D))$ is eventual in B.

*Finally, suppose $y^{-1}(x(D))$ is eventual in B whenever D is eventual in A: since $D = \{\alpha \in A : \alpha \geq \eta\}$ is automatically eventual in A, $y^{-1}(x(D)) = \{\beta \in B : y_\beta = x_\alpha \text{ for some } \alpha \geq \eta\}$ must contains $\{\beta \in B : \beta \geq \gamma\}$ for some $\gamma \in B$. In other words, $\{y_\beta : \beta \geq \gamma\} \subseteq x(D) = \{x_\alpha : \alpha \geq \eta\}$. As η varies in A, this shows that $\mathcal{E}[y]$ refines $\mathcal{E}[x]$, i.e. y is an AA-subnet of x. □

2 nets are called *AA-equivalent* if each of them is an AA-subnet of another one.

Lemma 4.57. Let $x : A \to X$, $y : B \to X$ be nets, and $\mathcal{E}[x], \mathcal{E}[y]$ be correspondingly the eventuality filter bases. Then TFAE

1. $F_x \cap F_y \neq \varnothing$ for every $F_x \in \mathcal{E}[x]$, $F_y \in \mathcal{E}[y]$.
2. $\mathcal{M} = \{S \subseteq X : S \supseteq F_x \cap F_y \text{ for some } F_x \in \mathcal{E}[x], F_y \in \mathcal{E}[y]\}$ is a proper filter.
3. $\mathcal{E}[x], \mathcal{E}[y]$ have a common filter base refining both.
4. x and y have a common AA-subnet.
5. x and y have a common Willard subnet. Furthermore, such subnet can be chosen to be a maximal AA-subnet (i.e. under AA-subnet relation).

Proof. We will show (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (1).

*Suppose $F_x \cap F_y \neq \varnothing$ for every $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$:

1. If $S \subseteq T$ with $S \in \mathcal{M}$, then $T \supseteq S \supseteq F_x \cap F_y$ for some $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$. By construction $T \in \mathcal{M}$.
2. If $S^k \supseteq F_x^k \cap F_y^k$ for $k = 1, 2$, where $F_x^k \in \mathcal{E}[x]$ and $F_y^k \in \mathcal{E}[y]$, then $S^1 \cap S^2 \supseteq (F_x^1 \cap F_y^1) \cap (F_x^2 \cap F_y^2) = (F_x^1 \cap F_x^2) \cap (F_y^1 \cap F_y^2)$. Since $\mathcal{E}[x], \mathcal{E}[y]$ are filter base, we know that $F_x \subseteq F_x^1 \cap F_x^2$ and $F_y \subseteq F_y^1 \cap F_y^2$ for some $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$. In summary, $S^1 \cap S^2 \supseteq (F_x^1 \cap F_x^2) \cap (F_y^1 \cap F_y^2) \supseteq F_x \cap F_y$, and so $S^1 \cap S^2 \in \mathcal{M}$.
3. If $\mathcal{M}$ is not proper, then $\varnothing \in \mathcal{M}$, which means $\varnothing \supseteq F_x \cap F_y$ for some $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$. In particular, $F_x \cap F_y = \varnothing$, contradicting the hypothesis.

*Suppose $\mathcal{M}$ is a proper: this means that $F_x \cap F_y$ can never be empty for all $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$. Therefore, $\mathcal{M}$ is a filter base that refines $\mathcal{E}[x]$ since $F_x \supseteq F_x \cap G$ for a fixed $G \in \mathcal{E}[y]$. Similarly, $\mathcal{M}$ refines $\mathcal{E}[y]$.

*Suppose $\mathcal{B}$ is a common filter base refining both $\mathcal{E}[x]$ and $\mathcal{E}[y]$: let $z : \mathcal{B} \to X$ be a derived net (which exists assuming choice). This leads to a filter base $\mathcal{E}[z]$ which should refines $\mathcal{B}$. Indeed, since $x_C \in C \subseteq B$ for any $C \subseteq B$ $(C, B \in \mathcal{B})$, we get $\{x_C : C \subseteq B\} \subseteq B$ for each $B \in \mathcal{B}$. Because refinement relation is transitive, $\mathcal{E}[z]$ refines both $\mathcal{E}[x]$ and $\mathcal{E}[y]$. By definition, z is an AA-subnet to both x and y.

*Suppose x and y has a common AA-subnet: $C = \{(\alpha, \beta) \in A \times B : x_\alpha = y_\beta\}$ ordered by $(\alpha_1, \beta_1) \leq (\alpha_2, \beta_2)$ iff $\alpha_1 \leq \alpha_2$ and $\beta_1 \leq \beta_2$. Construct $z : C \to X$ where $z(\alpha, \beta) = x_\alpha = y_\beta$. We verify the followings

1. C is a frequent subset of $A \times B$, and so it is directed: let $w : D \to X$ be a common AA-subnet of x and y. Given $\alpha_0 \in A$ and $\beta_0 \in B$, then there

exists $\omega_0^A, \omega_0^B, \omega_0 \in D$ such that

$$\{w_\delta : \delta \geq \omega_0^A\} \subseteq \{x_\alpha : \alpha \geq \alpha_0\}$$
$$\{w_\delta : \delta \geq \omega_0^B\} \subseteq \{y_\beta : \beta \geq \beta_0\}$$
$$\omega_0 \geq \omega_0^A, \omega_0^B$$

Therefore, $\{w_\delta : \delta \geq \omega_0\} \subseteq \{x_\alpha : \alpha \geq \alpha_0\} \cap \{y_\beta : \beta \geq \beta_0\}$. In particular, as α_0 and β_0 varies, C becomes a frequent subset of $A \times B$.

Hence, for all $(\alpha_k, \beta_k) \in C$ $(k = 1, 2)$, let α_0 and β_0 be respective upper bound of α_1, α_2 and β_1, β_2. Since C is frequent, there exists $(\alpha_f, \beta_f) \geq (\alpha_0, \beta_0) \geq (\alpha_k, \beta_k)$ $(k = 1, 2)$. We conclude that C is directed.

2. Let $\pi_A : C \to A$ and $\pi_B : C \to B$ be projection of coordinates in C. Then

   (a) Whenever $(\alpha_1, \beta_1) \leq (\alpha_2, \beta_2)$, we get $\pi_A(\alpha_1, \beta_1) = \alpha_1 \leq \alpha_2 = \pi_A(\alpha_2, \beta_2)$, i.e. π_A is monotone. Similarly, π_B is monotone.

   (b) Whenever $\alpha_0 \in A$, fix some $\beta_0 \in B$, and let $(\alpha_f, \beta_f) \in C$ such that $(\alpha_f, \beta_f) \geq (\alpha_0, \beta_0)$. Then $\pi_A(\alpha, \beta) = \alpha \geq \alpha_f \geq \alpha_0$ for all $(\alpha, \beta) \geq (\alpha_f, \beta_f)$. In other words, π_A (and π_B similarly) is cofinal.

3. Furthermore, $z(\alpha, \beta) = x_\alpha = y_\beta = x \circ \pi_A(\alpha, \beta) = y \circ \pi_B(\alpha, \beta)$ for all $(\alpha, \beta) \in C$. This concludes z to be a Willard subnet of both x and y.

4. Additionally, z is the maximal AA-subnet: let $w : D \to X$ to be any AA-subnet of x and y. Using the same argument in part (1), there exists $\omega_0 \in D$ for each $(\alpha_0, \beta_0) \in A \times B$ such that

$$\{w_\delta : \delta \geq \omega_0\} \subseteq \{x_\alpha : \alpha \geq \alpha_0\} \cap \{y_\beta : \beta \geq \beta_0\}$$

   If $\gamma_0 = (\alpha_0, \beta_0) \in C$, then $\{w_\delta : \delta \geq \omega_0\} \subseteq \{x_\alpha : \alpha \geq \alpha_0\} \cap \{y_\beta : \beta \geq \beta_0\} = \{z = x_\alpha = y_\beta : (\alpha, \beta) \geq (\alpha_0, \beta_0)\} = \{z_\gamma : \gamma \geq \gamma_0\}$. As $\gamma_0 \in C$ can be made arbitrary, w is an AA-subnet of z.

*Suppose $z : C \to X$ is a common Willard subnet to both x and y: as z is also AA-subnet of x and y, the eventuality filter base $\mathcal{E}[z]$ refines both $\mathcal{E}[x]$ and $\mathcal{E}[y]$. For each $F_x \in \mathcal{E}[x]$ and $F_y \in \mathcal{E}[y]$, let $G_x, G_y, H \in \mathcal{E}[z]$ be sets such that $G_x \subseteq F_x, G_y \subseteq F_y$, and $H \subseteq G_x \cap G_y$. By definition, H is a non-empty subset of $G_x \cap G_y \subseteq F_x \cap F_y$, so $F_x \cap F_y$ must be non-empty. □

Corollary 4.58. Every AA-subnet is AA-equivalent to a Willard subnet.

Proof. If $x : A \to X$ is net, and $y : B \to X$ is some AA-subnet of x, then by previous lemma, we can construct a Willard subnet to both x and y. In particular z is an AA-subnet of y. Additionally, z can be made into the maximal AA-subnet, indicating that y must be AA-subnet of z (note y is AA-subnet of itself). By definition, y is AA-equivalent to a Willard subnet z of x. □

Proposition 4.59 (Transformational Adjunction II)**.** Let $(x_\alpha)_{\alpha\in A}$ be a net and $\mathcal{B}$ be a filter base on X. Show that

$\mathcal{B}$ refines the eventuality filter base of $(x_\alpha)_{\alpha\in A}$ if and only if the canonical net of $\mathcal{B}$ is an AA-subnet of $(x_\alpha)_{\alpha\in A}$.

Additionally, the eventuality filter base of the canonical net of $\mathcal{B}$ is just $\mathcal{B}$!

Proof.
*Suppose $\mathcal{B}$ refines $x(\mathcal{E}_A)$: the canonical net of $\mathcal{B}$ is $\pi_\mathcal{B} : \langle X, \mathcal{B}\rangle \to X$. The eventuality filter base $\mathcal{F} = \pi_\mathcal{B}(\mathcal{E}_{\langle X,\mathcal{B}\rangle})$ of this net is: $S \in \mathcal{F}$ if and only if $S = \pi_\mathcal{B}\{(q, C) \in \langle X, \mathcal{B}\rangle : (q, C) \geq (p, B)\} = \{(q, C) \in \langle X, \mathcal{B}\rangle : C \subseteq B\} = B$ for some $p \in B \in \mathcal{B}$. In other words, $\mathcal{F} = \mathcal{B}$, so $\pi_\mathcal{B}$ is an AA-subnet of $(x_\alpha)_{\alpha\in A}$ (as $\mathcal{B}$ refines $x(\mathcal{E}_A)$ by assumption).

*Suppose $\pi_\mathcal{B} : \langle X, \mathcal{B}\rangle \to X$ is an AA-subnet of $(x_\alpha)_{\alpha\in A}$. Again, we know that $\mathcal{B}$ is just the eventuality filter base on X of $\pi_\mathcal{B}$. Therefore, $\mathcal{B}$ is finer than the eventuality filter base $x(\mathcal{E}_A)$ of x. □

Remark 4.60. Given a net $(p_\alpha)_{\alpha\in A}$, the canonical filtered sequence is the same p, but with $\mathcal{E}_A$ as the eventuality filter base on A.

Vice versa, given a $\mathcal{B}$-filtered sequence $(x_i)_{i\in I}$, then the canonical net of the sequence is the composition $\lambda : \langle I, \mathcal{B}\rangle \to X$ of the projection $\pi_\mathcal{B} : \langle I, \mathcal{B}\rangle \to I$ and $x : I \to X$ itself (recall that $\langle I, \mathcal{B}\rangle = \{(i, B) : i \in B \in \mathcal{B}\}$)

Proposition 4.61 (Transformational Bi-adjunction III)**.** Let $(x_i)_{i\in I}$ be a $\mathcal{B}$-filtered sequence and $(p_\alpha)_{\alpha\in A}$ be a net in X. Then

1. $(p_\alpha)_{\alpha\in A}$ is an AA-subnet of the canonical net of $(x_i)_{i\in I}$ if and only if the canonical sequence of $(p_\alpha)_{\alpha\in A}$ is an AA-subsequence of $(x_i)_{i\in I}$.

2. $(x_i)_{i\in I}$ is an AA-subsequence of the canonical sequence of $(p_\alpha)_{\alpha\in A}$ if and only if the canonical net of $(x_i)_{i\in I}$ is an AA-subnet of $(p_\alpha)_{\alpha\in A}$.

In the categorical sense, filtered sequence and net provides the same expressive power. They are still less expressive (but more general), though, than filter base as many filtered sequences and nets have the same induced filter base.

Proof.
*Suppose $(p_\alpha)_{\alpha\in A}$ is an AA-subnet of the the canonical net of $(x_i)_{i\in I}$: let $\lambda : \langle I, \mathcal{B}\rangle \to X$ be the composition of $\pi_\mathcal{B} : \langle I, \mathcal{B}\rangle \to I$ and $x : I \to X$. By assumption, $p(\mathcal{E}_A)$ refines $\lambda(\mathcal{E}_{\langle I,\mathcal{B}\rangle}) = x(\pi_\mathcal{B}(\mathcal{E}_{\langle I,\mathcal{B}\rangle}))$. Since we have proved $\pi_\mathcal{B}(\mathcal{E}_{\langle X,\mathcal{B}\rangle}) = \mathcal{B}$ (even when X is just a set) in Proposition 4.59), this shows that $p(\mathcal{E}_A)$ refines $x(\mathcal{B})$. Finally, note that the canonical sequence of $(p_\alpha)_{\alpha\in A}$ only adds $\mathcal{E}_A$ as a filter base on the index set A, we conclude the canonical

sequence of $(p_\alpha)_{\alpha\in A}$ is an AA-subsequence of $(x_i)_{i\in I}$.

Vice versa, suppose the canonical sequence of $(p_\alpha)_{\alpha\in A}$ is an AA-subsequence of $(x_i)_{i\in I}$: in other words, $p(\mathcal{E}_A)$ refines $x(\mathcal{B})$. By previous part, we know that the eventuality filter base of the canonical net $\lambda : \langle I, \mathcal{B}\rangle \to X$ is just $x(\mathcal{B})$ and so, by definition, $(p_\alpha)_{\alpha\in A}$ is an AA-subnet of the canonical net of $(x_i)_{i\in I}$.

*The reverse adjunction is proven similarly. □

Proposition 4.62 (Compositional relation between transformations)**.** Let $\mathcal{F}$ be a filter base, $(x_i)_{i\in I}$ be a $\mathcal{B}$-filtered sequence, and $(p_\alpha)_{\alpha\in A}$ be a net on X.

1. If the canonical sequence of $\mathcal{F}$ is $(x_i)_{i\in I}$, and the canonical net of $(x_i)_{i\in I}$ is $(p_\alpha)_{\alpha\in A}$, then the canonical net of $\mathcal{F}$ coincides with $(p_\alpha)_{\alpha\in A}$.
2. If the canonical net of $\mathcal{F}$ is $(p_\alpha)_{\alpha\in A}$, and the canonical sequence of $(p_\alpha)_{\alpha\in A}$ is $(x_i)_{i\in I}$, then canonical sequence of $\mathcal{F}$ admits $(x_i)_{i\in I}$ as a KW-subsequence.
3. If the induced filter base of $(x_i)_{i\in I}$ is $\mathcal{F}$, and the canonical net of $\mathcal{F}$ is $(p_\alpha)_{\alpha\in A}$, then the canonical net of $(x_i)_{i\in I}$ is a Willard subnet of $(p_\alpha)_{\alpha\in A}$.
4. If the canonical net of $(x_i)_{i\in I}$ is $(p_\alpha)_{\alpha\in A}$, and the eventuality filter base of $(p_\alpha)_{\alpha\in A}$ is $\mathcal{F}$, then the induced filter base of $(x_i)_{i\in I}$ coincides with $\mathcal{F}$.
5. If the eventuality filter base of $(p_\alpha)_{\alpha\in A}$ is $\mathcal{F}$, and the canonical sequence of $\mathcal{F}$ is $(x_i)_{i\in I}$, then the canonical sequence of $(p_\alpha)_{\alpha\in A}$ is a KW-subsequence of $(x_i)_{i\in I}$.
6. If the canonical sequence of $(p_\alpha)_{\alpha\in A}$ is $(x_i)_{i\in I}$, and the induced filter base of $(x_i)_{i\in I}$ is $\mathcal{F}$, then the eventuality filter base of $(p_\alpha)_{\alpha\in A}$ coincides with $\mathcal{F}$.

Proof.
Part 1: by the hypothesis, $x = \mathrm{id}_X, \mathcal{B} = \mathcal{F}$, and $A = \langle I, \mathcal{B}\rangle = \langle X, \mathcal{F}\rangle, p = \mathrm{id}_X \circ \pi_{\mathcal{F}} = \pi_{\mathcal{F}}$. On the other hand, the canonical net of $\mathcal{F}$ is $\pi_{\mathcal{F}} : \langle X, \mathcal{F}\rangle \to X$, which is the same as p.
Part 2: by the hypothesis, $p : A \to X$ is the same as $\pi_{\mathcal{F}} : \langle X, \mathcal{F}\rangle \to X$, and $(x, \mathcal{B})$ is the same as $(p, \mathcal{E}_A) = (\pi_{\mathcal{F}}, \mathcal{E}_{\langle X,\mathcal{F}\rangle})$. On the other hand, the canonical sequence of $\mathcal{F}$ is $(\mathrm{id}_X, \mathcal{F})$, so consider $\varphi = \pi_{\mathcal{F}}$, which should serves as the re-indexing map from $\pi_{\mathcal{F}}$ to id_X

1. $\pi_{\mathcal{F}} = \mathrm{id}_X \circ \varphi$.
2. Note that $\pi_{\mathcal{F}}(\mathcal{E}_{\langle X,\mathcal{F}\rangle}) = \mathcal{F}$, so $\varphi(\mathcal{E}_{\langle X,\mathcal{F}\rangle})$ refines $\mathcal{F}$ by reflexivity.

Part 3: by the hypothesis, $\mathcal{F} = x(\mathcal{B})$, and $A = \langle X, x(\mathcal{B})\rangle, p = \pi_{x(\mathcal{B})}$. On the other hand, the canonical net of $(x_i)_{i\in I}$ is $x \circ \pi_{\mathcal{B}} : \langle I, \mathcal{B}\rangle \to X$, so consider the map $\mu_x : \langle I, \mathcal{B}\rangle \to \langle X, x(\mathcal{B})\rangle$ where (i, B) is sent to $(x_i, x(B))$.

1. $p(\mu_x(i, B)) = \pi_{x(\mathcal{B})}(x_i, x(B)) = x_i = x(\pi_{\mathcal{B}}(i, B))$

2. μ_x is monotone: if $(i_1, B_1) \leq (i_2, B_2)$, then $B_1 \supseteq B_2$, $x(B_1) \supseteq x(B_2)$, and $(x_{i_1}, x(B_1) \leq (x_{i_2}, x(B_2))$, or equivalently, $\mu_x(i_1, B_1) \leq \mu_x(i_2, B_2)$.

3. μ_x is cofinal: given $(p, x(B))$ $(p \in x(B), B \in \mathcal{B})$, then pick some $i \in B$, we get $\mu_x(i, B) = (x_i, x(B))$. So for all $(j, C) \geq (i, B)$, we get $C \subseteq B$, $x(C) \subseteq x(B)$, and $\mu_x(j, C) = (x_j, x(C)) \geq (p, x(B))$

We conclude that the canonical net (which is $x \circ \pi_{\mathcal{B}}$) of $(x_i)_{i \in I}$ is a Willard subnet of $(p_\alpha)_{\alpha \in A}$ (or $\pi_{x(\mathcal{B})}$.
Part 4: by the hypothesis, $A = \langle I, \mathcal{B} \rangle, p = x \circ \pi_{\mathcal{B}}$, and $\mathcal{F} = p(\mathcal{E}_A) = x(\pi_{\mathcal{B}}(\mathcal{E}_{\langle I, \mathcal{B} \rangle})) = x(\mathcal{B})$. The induced filter base of $(x_i)_{i \in I}$ is just $x(\mathcal{B})$, so they coincide.
Part 5: by the hypothesis, $\mathcal{F} = p(\mathcal{E}_A)$ and $(x, \mathcal{B}) = (\mathrm{id}_X, \mathcal{F}) = (\mathrm{id}_X, p(\mathcal{E}_A))$. On the other hand, the canonical sequence of $(p_\alpha)_{\alpha \in A}$ is $(p, \mathcal{E}_A)$, so consider $p : A \to X$

1. $p = \mathrm{id}_X \circ p = x \circ p$

2. $p(\mathcal{E}_A) = \mathcal{F} = \mathcal{B}$, so $p(\mathcal{E}_A)$ refines $\mathcal{B}$ in particular.

Therefore, p serves as the re-indexing map from $(p, \mathcal{E}_A)$ to $(x, \mathcal{B})$.
Part 6: by the hypothesis, $(p, \mathcal{E}_A) = (x, \mathcal{B})$ and $\mathcal{F} = x(\mathcal{B}) = p(\mathcal{E}_A)$. On the other hand, the eventuality filter base of $(p_\alpha)_{\alpha \in A}$ is $p(\mathcal{E}_A)$ so they coincide. □

A *derived net* on a filter base $\mathcal{B}$ on X is a net $x : \mathcal{B} \to X$ (with $\mathcal{B}$ ordered under reverse inclusion) such that $x(B) \in B$ for all $B \in \mathcal{B}$.

Proposition 4.63. A filter base $\mathcal{B}$ converges to x if and only if every derived net on $\mathcal{B}$ converges to x.

Proof. Suppose $\mathcal{B} \to x$. Let $p : \mathcal{B} \to X$ to be any derived net on $\mathcal{B}$. Given a neighborhood U of x, there exists some $B \in \mathcal{B}$ such that $B \subseteq U$. Hence $p_F \in F \subseteq U$ for all $F \subseteq B, F \in \mathcal{B}$. By definition, the net $(p_B)_{B \in \mathcal{B}}$ converges to x.

Vice versa, suppose $\mathcal{B}$ does not converge to x. We will construct a derived net that also does not converge to x. By the hypothesis, there exists a neighborhood N of x such that $B \not\subseteq N$ for all $B \in \mathcal{B}$. Equivalently, $B \setminus N \neq \varnothing$ for all $B \in \mathcal{B}$. Hence, let $q : \mathcal{B} \to X$ be a derived net such that $q_B \notin N$ for all $B \in \mathcal{B}$. If $q_B \xrightarrow{B} x$, then given the neighborhood N of p, there exists some $C \in \mathcal{B}$ such that $q_B \in N$ for all $B \subseteq C$, a contradiction to the fact that $q_B \notin N$ for all $B \in \mathcal{B}$. Hence, q_B does not converge to x. □

Proof Note: the reverse direction essentially requires axiom of choice for the existence of a derived net.

4.5 Continuity

A *continuous* map $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ between topological spaces consists of a function $f : X \to Y$ between sets such that $f^{-1}(V)$ is open in X whenever V is open in Y.

Proposition 4.64. Let $f : X \to Y$ be a function between sets, $\mathcal{B}$ be a filter base on X, and S be a subset of X

1. If S is $\mathcal{B}$-eventual, then $f(S)$ is $f(\mathcal{B})$-eventual.
2. If S is $\mathcal{B}$-frequent, then $f(S)$ is $f(\mathcal{B})$-frequent.

Proof.
Part 1: by assumption, there exists $B \in \mathcal{B}$ such that $B \subseteq S$. Apply f on both side we get, $f(B) \subseteq f(S)$ (and $f(B) \in f(\mathcal{B})$).
Part 2: by assumption, for each $B \in \mathcal{B}$, $B \cap S \neq \varnothing$. Again, applying f on both side, we get $f(B \cap S) \neq \varnothing$. But $f(B \cap S) \subseteq f(B) \cap f(S)$, so we also have $f(B) \cap f(S) \neq \varnothing$. □

Theorem 4.65 (Equivalence of definitions)**.** Given a function $f : X \to Y$ between spaces, TFAE:

1. f is continuous.
2. [Closedness] The inverse image of a closed set (under f) is closed.
3. [Neighborhood] f is continuous at x for every $x \in X$. Here, f is *continuous at* x if for every neighborhood N of $f(x)$, $f^{-1}(N)$ is a neighborhood of x.
4. [Convergence] $f(\mathcal{B}) \to f(x)$ whenever $\mathcal{B} \to x$.
5. [Closure] $f(\mathrm{cl}(A)) \subseteq \mathrm{cl}(f(A))$ for all $A \subseteq X$.
6. [Interior] $f^{-1}(\mathrm{int}(B)) \subseteq \mathrm{int}(f^{-1}(B))$ for all $B \subseteq Y$.

Proof. We prove (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (5) ⇒ (6) ⇒ (1).

*Suppose f is continuous: given C a closed set in Y, then $f^{-1}(Y \setminus C) = X \setminus f^{-1}(C)$ is open in X, which shows that $f^{-1}(C)$ is closed in X.

*Suppose the inverse image of any closed set is closed: for every neighborhood N of $f(x)$, there exists some open neighborhood $V \subseteq N$ of $f(x)$ (and so $x \in f^{-1}(V) \subseteq f^{-1}(N)$). By assumption, $f^{-1}(Y \setminus V) = X \setminus f^{-1}(V)$ is a closed set. Thus, $f^{-1}(V)$ is open, and so $f^{-1}(N)$ is a neighborhood of x.

*Suppose f is continuous at x for all $x \in X$: given $\mathcal{B} \to x$, let N be a neighborhood of $f(x)$. Then $f^{-1}(N)$ is a neighborhood of x, so there exists some $B \in \mathcal{B}$ such that $B \subseteq f^{-1}(N)$. Equivalently, $f(B) \subseteq N$, so since N is arbitrary, $f(\mathcal{B}) \to f(x)$.

*Suppose $f(\mathcal{B}) \to f(x)$ whenever $\mathcal{B} \to x$: let $x \in \mathrm{cl}(A)$, we need to show that $f(x) \in \mathrm{cl}(f(A))$. By Theorem 4.40, there must be some filter base $\mathcal{B}$ converging to x such that A is $\mathcal{B}$-frequent. By the hypothesis and Proposition 4.64, $f(\mathcal{B})$ converges to $f(x)$ and $f(A)$ is $f(\mathcal{B})$-frequent. We conclude that $f(\mathrm{cl}(A)) \subseteq \mathrm{cl}(f(A))$.

*Suppose $f(\mathrm{cl}(A)) \subseteq \mathrm{cl}(f(A))$ for all $A \subseteq X$ (or $\mathrm{cl}(A) \subseteq f^{-1}(\mathrm{cl}(f(A)))$): substituting $A = f^{-1}(Y \setminus B)$, we get

$$\mathrm{cl}(f^{-1}(Y \setminus B)) = \mathrm{cl}(A) \subseteq f^{-1}(\mathrm{cl}(f(A))) = f^{-1}(\mathrm{cl}(f(f^{-1}(Y \setminus B)))$$

The LHS reduces to $\mathrm{cl}(f^{-1}(Y \setminus B)) = \mathrm{cl}(X \setminus f^{-1}(B)) = X \setminus \mathrm{int}(f^{-1}(B))$. For the RHS, since $f(f^{-1}(Y \setminus B)) \subseteq Y \setminus B$, we have $f^{-1}(\mathrm{cl}(f(f^{-1}(Y \setminus B))) \subseteq f^{-1}(\mathrm{cl}(Y \setminus B)) = f^{-1}(Y \setminus \mathrm{int}(B)) = X \setminus f^{-1}(\mathrm{int}(B))$. Combining the above results together, we obtain

$$X \setminus \mathrm{int}(f^{-1}(B)) \subseteq X \setminus f^{-1}(\mathrm{int}(B)) \text{ or}$$
$$\mathrm{int}(f^{-1}(B)) \supseteq f^{-1}(\mathrm{int}(B))$$

*Suppose $f^{-1}(\mathrm{int}(B)) \subseteq \mathrm{int}(f^{-1}(B))$ for all $B \subseteq Y$: let V be an open set of Y, and substitute $B = V$ into the hypothesis, we get $f^{-1}(V) = f^{-1}(\mathrm{int}(V)) \subseteq \mathrm{int}(f^{-1}(V))$. On the other hand, $\mathrm{int}(f^{-1}(V)) \subseteq f^{-1}(V)$, so we get $f^{-1}(V) = \mathrm{int}(f^{-1}(V))$, i.e. $f^{-1}(V)$ is open. □

Proposition 4.66 (Continuity in terms of generation)**.** Given $\mathcal{S}$ generates Y, then $f : X \to Y$ is continuous if and only if $f^{-1}(A)$ is open for every $A \in \mathcal{S}$.

Proof. The forward direction is straightforward, so we will prove the backward one. Since $\mathcal{S}$ generates Y, every open set U in Y is a union of (open sets) $\{V_i\}_{i \in I}$, with each V_i is a finite intersection of (open) sets in $\mathcal{S}$ (Theorem 4.18). The inverse image preserves both intersection and union, so each $f^{-1}(V_i)$ is a *finite* intersection of sets of the form $f^{-1}(A)$ ($A \in \mathcal{S}$), each is open by the hypothesis. Therefore, $f^{-1}(V_i)$ is open, and so $f^{-1}(U) = \bigcup_{i \in I} f^{-1}(V_i)$ is also open. □

Theorem 4.67 (Category of topological space)**.**

1. The topological spaces, together with continuous functiosn, form a category, denoted as **Top**.

2. The constant map $c_s : X \to Y, x \mapsto s$ for a fixed $s \in Y$ is continuous. In particular, any map from a singleton to another space is automatically continuous.

3. The intial object is the empty space $\varnothing$, and the termimal object is the singleton space $\{p\}$ (any such space are essentially the same categorically).

4. X is discrete if any function $f : X \to Y$ from X is continuous.
   Y is indiscrete if any function $f : X \to Y$ to Y is continuous.

5. If Y is discrete, then $f : X \to Y$ is continuous iff each fiber $f^{-1}(y)$ is open in X.
   If X is indiscrete, then $f : X \to Y$ is continuous iff every open sets in Y either contains the image $f(X)$ of X, or is disjoint from it.

Proof.
Part 1: the identity map id_X is continuous since $\mathrm{id}_X^{-1}(U) = U$ is open whenever U is open in X. Next, suppose $f : X \to Y$ and $g : Y \to Z$ are continuous. If W is open in Z, then $g^{-1}(W)$ is open in Y, and $f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W)$ is open in X. By definition $g \circ f$ is continuous.
Part 2: if $c_s(x) = s$ for all $x \in X$, then given an open set V in Y, either $s \in V$, which $c_s^{-1}(V) = X$, or $s \notin V$, and $c_s^{-1}(V) = \varnothing$. In either case, the inverse image of V is always open.
Part 3: there is only one map (the empty map) from $\varnothing$ to any space X, and it is trivially continuous as the inverse of any open subset of X is empty. Similarly, there is only one map (the constant map) from X to $\{p\}$, and we already show in part (2) that this map is continuous.
Part 4: if X is discrete, and $f : X \to Y$ is any map, then $f^{-1}(V)$ is always open regardless of whether V is open in Y. Vice versa, if any function from X is continuous, then consider $Y = \{0, 1\}$ with the discrete topology, with $f(S) = 0$ and $f(X \setminus S) = 1$. As f is continuous, $f^{-1}(0) = S$ is open (since Y is discrete). But S is an arbitrary subset of X, so X is discrete by definition.

If Y is indiscrete, and $f : X \to Y$ is any map, there are only 2 choices for open sets in Y: if $V = Y$, then $f^{-1}(V) = X$. If $V = \varnothing$, then $f^{-1}(V) = \varnothing$. We always have the inverse image of open sets in Y to be open, thus f is continuous. Vice versa, if any function to Y is continuous, then consider the identity map $\mathrm{id}_Y : Y \to Y$ (where the former is endowed with indiscrete topology). If U is open in Y, then U is open (through the inverse image under id_Y) with respect to the indiscrete topology on Y, and so either $U = Y$ or $U = Y$.
Part 5: if Y is discrete, then $\{y\}$ is open. Given $f : X \to Y$ a continuous map, the fiber $f^{-1}(y)$ must then be open in X. Vice versa, given every fiber is open in X, if V is open in Y, then $f^{-1}(V)$ must be open since it is a union of fibers $f^{-1}(y)$ for $y \in V$. By definition, f is continuous.

If X is indiscrete, and $f : X \to Y$ is continuous. For each open set V in Y, either $f^{-1}(V) = \varnothing$, which implies $f(X) \cap V = \varnothing$, or $f^{-1}(V) = X$, which implies $f(X) \subseteq V$. Vice versa, if for all open set V in Y, either $f(X) \cap V = \varnothing$ or $f(X) \subseteq V$, then either $f^{-1}(V) = \varnothing$ (since otherwise, there will be some $x \in X$ such that $f(x) \in V$), or $X \subseteq f^{-1}(V)$ (equivalently, $X = f^{-1}(V)$). As $\varnothing$ and X is open in X, we conclude f is continuous. □

Proposition 4.68.

1. $\mathrm{id}_X : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is continuous if and only if $\mathcal{T}_1 \supseteq \mathcal{T}_2$.

2. Covariance on the domain: if $f : (X, \mathcal{T}_1) \to Y$ is continuous, then $f : (X, \mathcal{T}_2) \to Y$ is continuous if $\mathcal{T}_1$ is coarser than $\mathcal{T}_2$.

3. Contravariance on the codomain: if $f : X \to (Y, \mathcal{S}_1)$ is continuous, then $f : X \to (Y, \mathcal{S}_2)$ is continuous if $\mathcal{S}_1$ is finer than $\mathcal{S}_2$.

Proof. Part 1: suppose $\mathrm{id}_X : (X, \mathcal{T}_1) \to (X, \mathcal{T}_2)$ is continuous. Whenever $U \in \mathcal{T}_2$, we get $\mathrm{id}_X^{-1}(U) = U \in \mathcal{T}_1$. On the other hand, if $\mathcal{T}_1 \supseteq \mathcal{T}_2$, then $U \in \mathcal{T}_2$ implies $U = \mathrm{id}_X^{-1}(U) \in \mathcal{T}_1$.
Part 2: if V is open in Y, then $f^{-1}(V) \in \mathcal{T}_1$. Since $\mathcal{T}_1 \subseteq \mathcal{T}_2$, $f^{-1}(V) \in \mathcal{T}_2$. But V is arbitrary, so $f : (X, \mathcal{T}_2) \to Y$ is continuous.
Part 3: if $V \in \mathcal{S}_2$, then $V \in \mathcal{S}_1$, thus $f^{-1}(V)$ is open in X. By definition, $f : X \to (Y, \mathcal{S}_2)$ is continuous. □

Proposition 4.69.

1. The closure of inverse image is a subset of the inverse image of closure.

2. Surjective image of a dense set is dense. In fact, if the image of the space is dense, then the image of any dense subset is also dense.

Proof.
Part 1: if $f : X \to Y$ is continuous and B is a subset of Y, then $f^{-1}(B) \subseteq f^{-1}(\mathrm{cl}(B))$ by monotonicity (of both closure and inverse image). By minimality, $\mathrm{cl}(f^{-1}(B)) \subseteq f^{-1}(\mathrm{cl}(B))$.
Part 2: suppose $f : X \to Y$ is a continuous function with dense image $\mathrm{cl}(f(X)) = Y$, and D is a dense set in X. By Theorem 4.65, $\mathrm{cl}(f(D)) \supseteq f(\mathrm{cl}(D)) = f(X)$. Taking the closure on both side, we get $\mathrm{cl}(f(D)) \supseteq \mathrm{cl}(f(X)) = Y$, or $\mathrm{cl}(f(D)) = Y$. □

Given X, Y topological spaces, we write $C(X, Y)$ to be the set of continuous functions $X \to Y$. If $X = Y$, we shorten to $C(X)$.

Let $S = \{0, 1\}$ with the topology to be $\{\varnothing, \{1\}, S\}$. Such space is called the *Sierpiński space*.

Theorem 4.70 (Representation Theorem)**.** The functor $\mathcal{T}$ that sends a space to its underlying topology, and a continuous map $f : X \to Y$ to $\mathcal{T}f : \mathcal{T}(Y) \to \mathcal{T}(X)$ (where $\mathcal{T}f(V) = f^{-1}(V)$) is representable, that is there is some space J such that $\mathcal{T}(-)$ is naturally isomorphic to $C(-, J)$

Proof. The space J we need is the Sierpiński space S. First, however, we shall verify that $\mathcal{T}$ is a contravariant functor. This is straight from construction and

1. $\mathcal{T}\,\mathrm{id}_X(U) = \mathrm{id}_X^{-1}(U) = U$ for any open set U, i.e. $\mathcal{T}$ maps identities to identities.

2. Given $f : X \to Y$ and $g : Y \to Z$ continuous, we get $\mathcal{T}f \circ \mathcal{T}g(W) = \mathcal{T}f(g^{-1}(W)) = f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W) = \mathcal{T}(g \circ f)(W)$ open in X whenver W is open in Z.

Next, we show that $\mathcal{T}(-) \cong C(-, S)$. Let $\eta_X : \mathcal{T}(X) \to C(X, S)$ be the map where $\eta_X(U) = 1_U$ (here, $1_U(x) = 1$ iff $x \in U$). Then η_X is

1. Well-defined: given U open in X, then 1_U is continuous. Indeed, as S only has 3 open sets, we just need to test whether $1_U^{-1}(1) = U$ is open, which is true.

2. Has an inverse: $\eta_X^{-1}(f) = f^{-1}(1)$. Indeed, the map is well-defined since $\{1\}$ is open in S. Furthermore, if $f = \eta_X(U)$, then $\eta_X^{-1}(f) = 1_U^{-1}(1) = U$, and if $U = \eta_X^{-1}(f)$, then we know that $f(U) = 1$ and $f(X \setminus U) = 0$, and so $f = 1_U = \eta_X(U)$.

3. Both $\eta : \mathcal{T} \to C(-, S)$ and its inverse λ are natural transformation: let $f : X \to Y$ be a continuous map.

   (a) $\eta_Y \circ \mathcal{T}f = C(f, S) \circ \eta_X$: indeed, for all $V \in \mathcal{T}(Y)$ and $x \in X$, we get

   $$(\eta_Y \circ \mathcal{T}f)(V)(x) = \eta_Y(f^{-1}(V))(x) = 1_{f^{-1}(V)}(x) = \begin{cases} 1 & \text{if } x \in f^{-1}(V) \\ 0 & \text{if } x \notin f^{-1}(V) \end{cases}$$

   $$(C(f, S) \circ \eta_X)(V) = (\eta_X(V) \circ f)(x) = 1_V(f(x)) = \begin{cases} 1 & \text{if } f(x) \in V \\ 0 & \text{if } f(x) \notin V \end{cases}$$

   Note that the last expression on each line is equal to each other.

   (b) $\lambda_X \circ C(f, S) = \mathcal{T}f \circ \lambda_Y$: indeed for all $g \in C(Y, S)$ and $x \in X$, we get

   $$(\lambda_X \circ C(f, S))(g) = \lambda_X(g \circ f) = (g \circ f)^{-1}(1)$$
   $$(\mathcal{T}f \circ \lambda_Y)(g) = \mathcal{T}f(g^{-1}(1)) = f^{-1}(g^{-1}(1))$$

□

An isomorphism in **Top** is called a *homeomorphism*. Any property that is preserve through homeomorphism is called a *topological invariant*. As for mono- and epi-morphisms, we will discuss later what they correspond to.

Remark 4.71. The followings are examples of topological invariants:

1. The cardinality of the space X, and the cardinality of the topology on it.

2. The weight $w(X)$ of a space X.

3. The density $d(X)$ of a space X.

Proof. Let $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ be a homeomorphism.
Part 1: As f is also a bijection between sets, we get $|X| = |Y|$. As for topology, note that if $g : Y \to X$ is the inverse map to f, then $\mathcal{T}f \circ \mathcal{T}g = \mathcal{T}(f \circ g) = \mathcal{T}\,\mathrm{id}_Y = \mathrm{id}_{\mathcal{T}(Y)}$ and $\mathcal{T}g \circ \mathcal{T}f = \mathrm{id}_{\mathcal{T}(X)}$ similarly. In other words, $\mathcal{T}f$ is a

bijection between topologies on X and Y, i.e. the cardinality of topologies are equal.
Part 2: we show that if $\mathcal{B}$ is a basis of X, then $f(\mathcal{B})$ is also a basis of Y. Since f^{-1} is also continuous, the inverse image of each $B \in \mathcal{B}$ under f^{-1} is open. We also have $f(X) = f\left(\bigcup \mathcal{B}\right) = \bigcup f(\mathcal{B}) = Y$ since f is a bijection. Finally, given any $B_1, B_2 \in \mathcal{B}$, then there exists some $B \in \mathcal{B}$ such that $B \subseteq B_1 \cap B_2$, hence $f(B) \subseteq f(B_1 \cap B_2) \subseteq f(B_1) \cap f(B_2)$.

From there, there is a one-to-one correspondence between bases of X and bases of Y, and $|\mathcal{B}| = |f(\mathcal{B})|$ for any basis $\mathcal{B}$ (or even arbitary collection of subsets of X). Hence, $w(X) = \inf\{|\mathcal{B}| : \mathcal{B} \text{ is a basis of } X\} = \inf\{|f(\mathcal{B})| : \mathcal{B} \text{ is a basis of } X\} = \inf\{|\mathcal{C}| : \mathcal{C} \text{ is a basis of } Y\} = w(Y)$.
Part 3: f is homeomorphism, so f is surjective in particular. Then any dense set in X is sent to another dense set in Y, so $d(X) \geq d(Y)$. Similarly, $d(Y) \geq d(X)$. □

4.5.1 Open map. Closed map

Let $f : X \to Y$ be a function between topological spaces. Then

1. f is open if the image of any open set is also open.
2. f is closed if the image of any closed set is also closed.

Proposition 4.72 (Equivalent definition)**.**

1. Open maps: TFAE
   (a) $f : X \to Y$ is open
   (b) [Basis] $f(B)$ is open whenever $B \in \mathcal{B}$, where $\mathcal{B}$ is a basis.
   (c) [Neighborhood] f sends each neighborhood of a point x to a neighborhood of $f(x)$.
   (d) $f(\text{int}(A)) \subseteq \text{int}(f(A))$ for all $A \subseteq X$.
   (e) Whenever C is closed in X, then $\{y \in Y : f^{-1}(y) \subseteq C\}$ is closed in Y.
2. Closed maps: $f : X \to Y$ is closed iff $\text{cl}(f(A)) \subseteq f(\text{cl}(A))$ for any subset A of X.

Proof.
Part 1: we will show $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a)$.
*Suppose $f : X \to Y$ is open: since a basis must contain only open set, we have $f(B)$ is open whenver $B \in \mathcal{B}$.

*Suppose, given a basis $\mathcal{B}$, that $f(B)$ is open whenever B is in $\mathcal{B}$: if N is a neighborhood of x, there must exists some U open and $B \in \mathcal{B}$ such that $p \in B \subseteq U \subseteq N$. We then have $f(B)$ is open and $f(p) \in f(B) \subseteq f(N)$. In

other words, $f(N)$ is a neighborhood of $f(p)$.

*Suppose f sends each neighborhood of x to a neighborhood of $f(x)$ for every $x \in X$: since $\text{int}(A)$ is a neighborhood of every point in it $f(\text{int}(A))$ is a neighborhood of every point in $f(\text{int}(A))$. By Remark 4.9, $f(\text{int}(A))$ is open. Since $\text{int}(A) \subseteq A$, we get $f(\text{int}(A)) \subseteq f(A)$. By maximality, $f(\text{int}(A)) \subseteq \text{int}(f(A))$.

*Suppose $f(\text{int}(A)) \subseteq \text{int}(f(A))$ for all $A \subseteq X$: let C be a closed set in X. Then $X \setminus C$ is open, so $f(X \setminus C) = f(\text{int}(X \setminus C)) \subseteq \text{int}(f(X \setminus C))$. On the other hand, $\text{int}(f(X \setminus C)) \subseteq f(X \setminus C)$, so we must have $f(X \setminus C) = \text{int}(f(X \setminus C))$, or simply, $f(X \setminus C)$ is open.

Consider $D = \{y \in Y : f^{-1}(y) \subseteq C\}$: if $y \notin D$, then $f^{-1}(y) \not\subseteq C$, or equivalently, $f^{-1}(y) \cap (X \setminus C) \neq \varnothing$. In other words, there exists $x \in X \setminus C$ such that $f(x) = y \in Y \setminus D$. Therefore, if we let y varies in $Y \setminus D$, $f(X \setminus C)$ must contain $Y \setminus D$, or equivalently, $D \supseteq Y \setminus f(X \setminus C)$. Vice versa, if $y \notin f(X \setminus C)$, then $y \neq f(x)$ for all $x \in X \setminus C$. Note that $y \neq f(x)$ is the same as saying $x \notin f^{-1}(y)$, so $X \setminus C$ is a subset of $X \setminus f^{-1}(y)$. Equivalently, $f^{-1}(y)$ is a subset of C, thus $y \in D$. We conclude that $D = Y \setminus f(X \setminus C)$ is closed, since we already show $f(X \setminus C)$ is open.

*Suppose $\{y \in Y : f^{-1}(y) \subseteq C\}$ is closed in Y whenever C is closed in X: note that we have prove that $\{y \in Y : f^{-1}(y) \subseteq C\}$ is the same as $Y \setminus f(X \setminus C)$. Given U open set in X, then $X \setminus U$ is closed, so $Y \setminus f(X \setminus (X \setminus U)) = Y \setminus f(U)$ must be closed, which in turn makes $f(U)$ open.

Part 2: if f is closed, then $f(\text{cl}(A))$ is a closed set containing $f(A)$ (since $\text{cl}(A)$ contains A). By minimality, $f(\text{cl}(A))$ contains $\text{cl}(f(A))$ also. Vice versa, suppose $f(\text{cl}(A)) \supseteq \text{cl}(f(A))$ for all $A \subseteq X$. Let C be a closed set in X, then $f(C) = f(\text{cl}(C)) \supseteq \text{cl}(f(C))$. We also have the reverse inclusion since the closure of a set always contain that set. Hence, $f(C) = \text{cl}(f(C))$, i.e. $f(C)$ is closed. □

Proposition 4.73.

1. Composition of open (resp. closed) maps is open (resp. closed).

2. A bijective map is open if and only if it is closed.

3. Every homeomorphism is both open and closed. In fact, a bijective continuous map is a homemorphism if and only if it is open, or equivalently, if and only if it is closed.

4. Inverse image of continuous and open maps: let $f : X \to Y$ be a function between topological spaces, and B be a subset of Y

(a) If f is continuous, then $\mathrm{cl}(f^{-1}(B)) \subseteq f^{-1}(\mathrm{cl}(B))$ and $f^{-1}(\mathrm{int}(B)) \subseteq \mathrm{int}(f^{-1}(B))$.

(b) If f is open, then $f^{-1}(\mathrm{cl}(B)) \subseteq \mathrm{cl}(f^{-1}(B))$ and $\mathrm{int}(f^{-1}(B)) \subseteq f^{-1}(\mathrm{int}(B))$.

5. Let $f : X \to Y$ be a continuous open map and $A \subseteq X, B \subseteq Y$

(a) $\mathrm{cl}(f^{-1}(B)) = f^{-1}(\mathrm{cl}(B))$

(b) $f^{-1}(\mathrm{int}(B)) = \mathrm{int}(f^{-1}(B))$

(c) $f^{-1}(\partial B) = \partial f^{-1}(B)$

(d) Suppose f is surjective, then B is regular open (resp. regular closed) in Y if and only if $f^{-1}(B)$ is regular open (resp. regular closed) in X

(e) If $\mathrm{cl}(A) = \mathrm{cl}(\mathrm{int}(A))$ then

$$\begin{aligned}\mathrm{cl}(f(A)) &= \mathrm{cl}(f(\mathrm{int}(A))) = \mathrm{cl}(f(\mathrm{int}(\mathrm{cl}(A))))\\ &= \mathrm{cl}(\mathrm{int}(f(A))) = \mathrm{cl}(f(\mathrm{cl}(A)))\end{aligned}$$

In particular,

i. If A is regular closed, then so is $\mathrm{cl}(f(A))$.

ii. If A is regular open, then so is $Y \setminus \mathrm{cl}(f(X \setminus A))$.

Proof. Part 1: suppose $f : X \to Y$ and $g : Y \to Z$ are open maps. Let U to be an open set of X, we need to show that $g \circ f(U)$ is also open in Z. Indeed, $f(U)$ is open in Y since f is open, and $g(f(U))$ is open in Z since g is open. Same argument for closed maps by replacing any occurrence of "open" with "closed".

Part 2: suppose $f : X \to Y$ is bijective. If f is open, then $f(X \setminus C) = Y \setminus f(C)$ (since f is both injective and surjective) is open whenever C is closed, and so $f(C)$ should be closed. Same argument for f closed (by swapping between "open" and "closed").

Part 3: suppose $f : X \to Y$ is a bijective continuous map, then by part (2), open and closed are equivalent, so we only need to prove one more equivalence between, say, open and homeomorphism. If f is homeomorphic, then the inverse is necessarily continuous, i.e. $f(U) = (f^{-1})^{-1}(U)$ is open in Y whenever U is open in X. Vice versa, if f is open, then the inverse is continuous since $(f^{-1})^{-1}(U) = f(U)$ is open whenever U is open. By definition, f is a homeomorphism.

Part 4:

(a) Since $f^{-1}(\mathrm{cl}(B))$ is a closed set containing $f^{-1}(B)$, we get $\mathrm{cl}(f^{-1}(B)) \subseteq f^{-1}(\mathrm{cl}(B))$ by minimality. Similarly, $f^{-1}(\mathrm{int}(B))$ is an open subset of $f^{-1}(B)$, so $f^{-1}(\mathrm{int}(B)) \subseteq \mathrm{int}(f^{-1}(B))$ by maximality.

(b) Since f is open, $f(X \setminus \mathrm{cl}(f^{-1}(B)))$ is open. Note that $\mathrm{cl}(f^{-1}(B)) \supseteq f^{-1}(B)$, so $X \setminus \mathrm{cl}(f^{-1}(B)) \subseteq X \setminus f^{-1}(B)$. Thus, $f(X \setminus \mathrm{cl}(f^{-1}(B))) \subseteq f(X \setminus f^{-1}(B)) = f(f^{-1}(Y \setminus B)) \subseteq Y \setminus B$. By minimality of interior, $f(X \setminus \mathrm{cl}(f^{-1}(B))) \subseteq \mathrm{int}(Y \setminus B) = Y \setminus \mathrm{cl}(B)$. Finally, applying the inverse image to both side, and note that $f^{-1}(f(A)) \supseteq A$, we get

$$\begin{aligned} X \setminus \mathrm{cl}(f^{-1}(B)) &\subseteq f^{-1}(f(X \setminus \mathrm{cl}(f^{-1}(B)))) \subseteq f^{-1}(Y \setminus \mathrm{cl}(B)) \\ &= X \setminus f^{-1}(\mathrm{cl}(B)) \end{aligned}$$

or equivalently, $f^{-1}(\mathrm{cl}(B)) \supseteq f^{-1}(\mathrm{cl}(B))$

Since $\mathrm{int}(f^{-1}(B)) \subseteq f^{-1}(B)$ and $f(f^{-1}(B)) \subseteq B$, we get $f(\mathrm{int}(f^{-1}(B))) \subseteq B$. Since f is open, $f(\mathrm{int}(f^{-1}(B)))$ is open, so $f(\mathrm{int}(f^{-1}(B))) \subseteq \mathrm{int}(B)$ by maximality. We then apply the inverse image to both side to get $f^{-1}(f(\mathrm{int}(f^{-1}(B)))) \subseteq f^{-1}(\mathrm{int}(B))$. As $f^{-1}(f(A))$ contains A, we conclude that $\mathrm{int}(f^{-1}(B)) \subseteq f^{-1}(\mathrm{int}(B))$.

Part 5:

(a) Follows from part (4).

(b) Follows from part (4).

(c) Applying part (5a) and the fact that $\partial B = \mathrm{cl}(B) \cap \mathrm{cl}(Y \setminus B)$, we get

$$\begin{aligned} f^{-1}(\partial B) &= f^{-1}(\mathrm{cl}(B) \cap \mathrm{cl}(Y \setminus B)) \\ &= \mathrm{cl}(f^{-1}(B)) \cap \mathrm{cl}(f^{-1}(Y \setminus B)) \\ &= \mathrm{cl}(f^{-1}(B)) \cap \mathrm{cl}(X \setminus f^{-1}(B)) = \partial f^{-1}(B) \end{aligned}$$

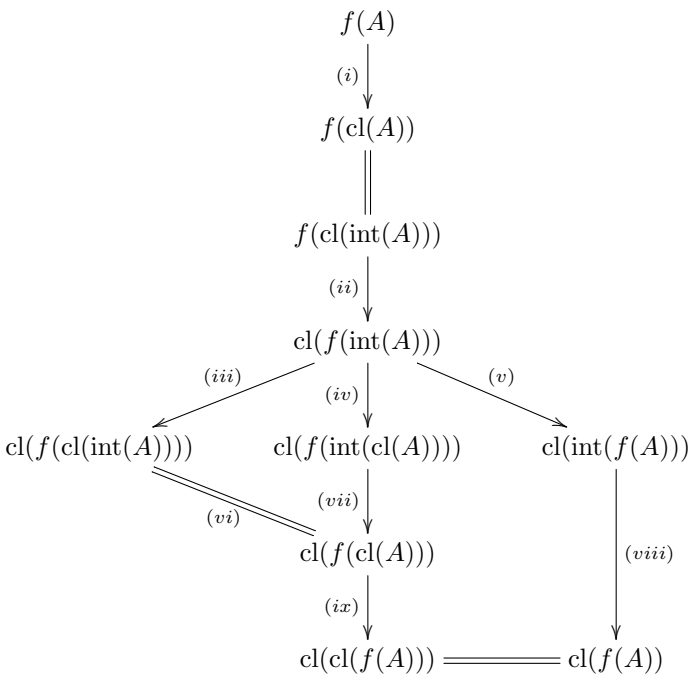
(d) Note that $f^{-1}(\mathrm{int}(\mathrm{cl}(B))) = \mathrm{int}(f^{-1}(\mathrm{cl}(B))) = \mathrm{int}(\mathrm{cl}(f^{-1}(B)))$. So if B is regular open, then $f^{-1}(B) = \mathrm{int}(\mathrm{cl}(f^{-1}(B)))$. Vice versa, if $f^{-1}(B)$ is regular open, we have $f^{-1}(\mathrm{int}(\mathrm{cl}(B))) = f^{-1}(B)$. Since f is surjective (i.e. $f(f^{-1}(T)) = T$ for any $T \subseteq Y$)

$$\mathrm{int}(\mathrm{cl}(B)) = f(f^{-1}(\mathrm{int}(\mathrm{cl}(B)))) = f(f^{-1}(B)) = B$$

Similar argument for regular closed case (or we can just use complementary to the previous result).

(e) Given $\mathrm{cl}(A) = \mathrm{cl}(\mathrm{int}(A))$, we observe the following inclusion using Hasse

diagram, with $S \to T$ represents the inclusion $S \subseteq T$



since

(i) Comes from apply the image (under f) to $A \subseteq \mathrm{cl}(A)$ (as it is monotonic)

(ii) Comes from substituting $S = \mathrm{int}(A)$ to $f(\mathrm{cl}(S)) \subseteq \mathrm{cl}(f(S))$ (a result of f being continuous).

(iii) Comes from applying image (under f) and then the closure operator to $\mathrm{cl}(\mathrm{int}(A)) \supseteq \mathrm{int}(A)$.

(iv) Comes from applying the interior operator, image, and the closure operator to $A \subseteq \mathrm{cl}(A)$.

(v) Comes from applying the closure operator to $f(\mathrm{int}(A)) \subseteq \mathrm{int}(f(A))$ (a result of f being open).

(vi) Comes from the hypothesis $\mathrm{cl}(\mathrm{int}(A)) = \mathrm{cl}(A)$

(vii) Comes from applying the image and closure operator to $\mathrm{int}(B) \subseteq B$ where $B = \mathrm{cl}(A)$.

(viii) Comes from applying the image and closure operator to $\mathrm{int}(B) \subseteq B$ where $B = f(A)$.

(ix) Comes from apply the closure operator to $f(\mathrm{cl}(A)) \subseteq \mathrm{cl}(f(A))$.

By minimality of closure operator on $f(A)$, any closed set in between

collapse to $\mathrm{cl}(f(A))$, i.e.

$$\begin{aligned}\mathrm{cl}(f(A)) &= \mathrm{cl}(f(\mathrm{int}(A))) = \mathrm{cl}(f(\mathrm{int}(\mathrm{cl}(A))))\\ &= \mathrm{cl}(\mathrm{int}(f(A))) = \mathrm{cl}(f(\mathrm{cl}(A)))\end{aligned}$$

In particular,

(i) If A is regular closed in X, we get $\mathrm{cl}(A) = A = \mathrm{cl}(\mathrm{int}(A))$, so

$$\mathrm{cl}(f(A)) = \mathrm{cl}(f(\mathrm{int}(A))) = \mathrm{cl}(\mathrm{int}(f(A)))$$

Since $\mathrm{int}(\mathrm{cl}(f(A))) \subseteq \mathrm{cl}(f(A))$, we get $\mathrm{cl}(\mathrm{int}(\mathrm{cl}(f(A)))) \subseteq \mathrm{cl}(\mathrm{cl}(f(A))) = \mathrm{cl}(f(A))$. For the reverse inclusion, since $\mathrm{cl}(f(A)) \supseteq f(A)$, we apply the closure and interior operators to get

$$\mathrm{cl}(\mathrm{int}(\mathrm{cl}(f(A)))) \supseteq \mathrm{cl}(\mathrm{int}(f(A))) = \mathrm{cl}(f(A))$$

We conclude $\mathrm{cl}(f(A))$ is regular closed.

(ii) if A is regular open, then $X \setminus A$ is regular closed, $f(X \setminus A)$ is regular closed, and $Y \setminus f(X \setminus A)$ is regular open.

□

4.5.2 Continuity at a point

Recall that a function $f : X \to Y$ between topological spaces is continuous at a point p in X if $f^{-1}(N)$ is a neighborhood of p whenever N is a neighborhood of $f(p)$.

Proposition 4.74 (Sequential continuity)**.** $f : X \to Y$ is continuous at p if and only if $f(\mathcal{B}) \to f(p)$ whenever $\mathcal{B} \to p$.

Proof. Suppose f is continuous at p, and $\mathcal{B}$ converges to p. For all neighborhood N of $f(p)$, we have $f^{-1}(N)$ is a neighborhood of p by definition. Thus, there exists some $B \in \mathcal{B}$ such that $B \subseteq f^{-1}(N)$. In other words, $f(B) \subseteq N$. By definition, $f(\mathcal{B}) \to f(p)$.

*Vice versa, suppose $f(\mathcal{B}) \to f(p)$ whenever $\mathcal{B} \to p$. In particular, $f(\mathcal{N}_p) \to f(p)$ where $\mathcal{N}_p$ is the neighborhood filter at p. Equivalently, for all M neighborhood of $f(p)$, there exists a neighborhood N of p such that $f(N) \subseteq M$, or $N \subseteq f^{-1}(M)$. Since $\mathcal{N}_p$ is a neighborhood filter, $f^{-1}(M)$ is also a neighborhood of p. As M varies, f is continuous at p. □

Proposition 4.75 (Pointed category)**.** If $f : X \to Y$ is continuous at p and $g : Y \to Z$ is continuous at $f(p)$, then $g \circ f$ is continuous at p.

Proof. If N is a neighborhood of $g(f(p))$, then $g^{-1}(N)$ is a neighborhood of $f(p)$, and so $f^{-1}(g^{-1}(N)) = (g \circ f)^{-1}(p)$ is a neighborhood of p. By definition, $g \circ f$ is continuous at p. □

More generally, given a function $f : X \to Y$ between topological spaces, and a filter base $\mathcal{B}$ that converges to p in X, we say q is a *$\mathcal{B}$-limit of f at p* if for all neighborhood N of q, there exists some $B \in \mathcal{B}$ such that $f(B) \subseteq N$. We denote $q \in \lim_{\mathcal{B} \to p} f(\mathcal{B})$

Remark 4.76. From Proposition 4.74, f is continuous at p if and only if $f(p) \in \lim_{\mathcal{B} \to p} f(\mathcal{B})$ for all $\mathcal{B} \to p$.

Example 4.77. Given $X = Y = \mathbb{R}$, there are several limits used in real analysis

1. When $\mathcal{B} = \{(a, p) : a < p\}$, the $\mathcal{B}$-limit is called *left-sided limit*. We denote such limit as $\lim_{x \to p^-} f(x)$.
2. When $\mathcal{B} = \{(p, b) : b > p\}$, the $\mathcal{B}$-limit is called *right-sided limit*. We denote such limit as $\lim_{x \to p^+} f(x)$.
3. When $\mathcal{B} = \{(a, b) \setminus \{p\} : a < p < b\}$, the $\mathcal{B}$-limit is called the *deleted limit* of f (at p). If the deleted limit coincide with the value of f at p (i.e. $f(p)$), then we recover the continuity of f at p.

Note:

1. The one-sided limits can be adopted to a more general situation when X is an ordered set, and Y is an arbitrary topological space (see the next section about the induced topology on ordered set).
2. As for the deleted limit of a function at a point, it works between topological spaces in general. However, when p is an isolated point of X, we have $Y = \lim_{\mathcal{B} \to p} f(\mathcal{B})$. So, in most case, we always assume p to be a limit point of X.

Proof. Note that when p is isolated point of X, this means that $\{p\}$ is itself open. Thus, $\mathcal{B} = \{N \setminus \{p\} : N \text{ is a neighborhood of } p\}$ contains empty set, and so for $f(\mathcal{B})$. Thus, $f(\mathcal{B})$ converges to each point in Y. □

4.6 Order topology

If $(L, \leq)$ is a linearly (or totally) ordered set, then the order topology on L is generated by the basis consisting of open intervals $(a, b) = \{x \in L : a < x < b\}$, and open rays $(a, \infty) = \{x \in L : x > a\}, (-\infty, b) = \{x \in L : x < b\}$.

Remark 4.78. In fact, the order topology can be generated by just open rays. The open intervals is just to make the collection into a basis.

Proof. Since $\bigcap_{i=1}^{n}(a_i, \infty) = (\max_{i=1}^{n} a_i, \infty) = (a, \infty)$, and $\bigcap_{j=1}^{m}(-\infty, b_j) = (-\infty, \min_{j=1}^{m} b_j) = (-\infty, b)$, we get

$$\bigcap_{i=1}^{n}(a_i, \infty) \cap \bigcap_{j=1}^{m}(-\infty, b_j) = (a, b)$$

Furthermore, given U_1, U_2 be either open intervals or rays, we consider the following cases

1. $U_1 = (a_1, b_1), U_2 = (a_2, \infty)$: then $U_1 \cap U_2 = (\max(a_1, a_2), b_1)$
2. $U_1 = (a_1, b_1), U_2 = (-\infty, b_2)$: then $U_1 \cap U_2 = (a_1, \min(b_1, b_2))$
3. $U_1 = (a_1, b_1), U_2 = (a_2, b_2)$: then $U_1 \cap U_2 = (\max(a_1, a_2), \min(b_1, b_2))$
4. $U_1 = (a_1, \infty), U_2 = (a_2, \infty)$: then $U_1 \cap U_2 = (\max(a_1, a_2), \infty)$
5. $U_1 = (a_1, \infty), U_2 = (-\infty, b_2)$: then $U_1 \cap U_2 = (a_1, b_2)$
6. $U_1 = (-\infty, b_1), U_2 = (-\infty, b_2)$: then $U_1 \cap U_2 = (-\infty \min(b_1, b_2))$

□

Remark 4.79. With the order topology, the closed intervals $[a, b]$ and closed rays $[a, \infty)$ and $(-\infty, b]$, as the name suggests, are closed.

Remark 4.80. The concept of eventuality and stationary for subsets of L is defined based upon the *eventuality filter base* on L.

Proposition 4.81. Suppose the order on X is dense

1. X is dense-in-itself, i.e. every point in X is a limit point.
2. The closure of the basic open sets are basic closed sets
   (a) $\mathrm{cl}(p, q) = [p, q]$.
   (b) $\mathrm{cl}(p, \infty) = [p, \infty)$.
   (c) $\mathrm{cl}(-\infty, q) = (-\infty, q]$.

   As a result, $\mathrm{cl}(p, q] = \mathrm{cl}[p, q) = [p, q]$.
3. The interior of basic closed sets are basic open sets

(a) $\text{int}[p,q] = (p,q)$, unless either $p = \min X$ or $q = \max X$.

(b) $\text{int}[p,\infty) = (p,\infty)$ unless $p = \min X$.

(c) $\text{int}(-\infty,q] = (-\infty,q)$ unless $q = \max X$.

Proof. Part 1: for each p and neighborhood N around p, we show that there is some point in N other than p. By definition of order topology, either

1. $p \in (a,b) \subseteq N$ for some $a,b \in X$: there exists some $q \in (a,p)$ (the ordering on X is dense), so N contains at least one point different from p.
2. $p \in (a,\infty) \subseteq N$ for some $a \in X$: there exists some $q \in (a,p)$, so N contains at least one point different from p.
3. $p \in (-\infty,b) \subseteq N$ for some $b \in X$: there exists some $q \in (p,b)$, so N contains at least one point different from p.

Part 2: since $[p,q] = X \setminus [(-\infty,p) \cup (q,\infty)]$ is closed, $\text{cl}(p,q) \subseteq [p,q]$. If N is a neighborhood of p, either

1. $p \in (a,b) \subseteq N$ for some $a,b \in X$: there exists some $r \in (p,b) \subseteq (p,q)$, so $N \cap (p,q)$ is not empty.
2. $p \in (-\infty,b) \subseteq N$ for some $b \in X$: there exists some $r \in (p,b) \subseteq (p,q)$, so $N \cap (p,q)$ is not empty.
3. $p \in (a,\infty) \subseteq N$ for some $a \in X$: then $(p,q) \subseteq (a,\infty)$, and so $N \cap (p,q)$ is not empty.

In any case $N \cap (p,q)$ is not empty, so $p \in \text{cl}(p,q)$. Similarly, $q \in \text{cl}(p,q)$. We conclude $\text{cl}(p,q) = [p,q]$. The other intervals follows analogously.
Part 3: since (p,q) is open, $\text{int}[p,q] \supseteq (p,q)$. If N is a neighborhood of p, then either

1. $p \in (a,b) \subseteq N$ for some $a,b \in X$: then there exists $r \in (a,p)$. Thus, N is not a subset of $[p,q]$ since it contains r.
2. $p \in (a,\infty) \subseteq N$ for some $a \in X$: then there exists $r \in (a,p)$. Thus, N is not a subset of $[p,q]$ since it contains r.
3. $p \in (-\infty,b) \subseteq N$ for some $b \in X$: if there does not exist $r < p$, then $p = \min L$ (since X is totally ordered), a contradiction. Thus, N is not a subset of $[p,q]$ since it contains r.

In any case, N cannot be a subset of $[p,q]$, so p cannot be an interior point of $[p,q]$. Similarly, q cannot be an interior point of $[p,q]$. We conclude $\text{int}[p,q] = (p,q)$. The other intervals follows analogously. □

Theorem 4.82 (Monotone Convergence Theorem)**.** Let X be a Dedekind-complete ordered set. If $x : A \to X$ is a bounded above, increasing net, i.e.

1. $x_\alpha \leq x_\beta$ whenever $\alpha \leq \beta$.

2. There exists some $p \in X$ such that $x_\alpha \leq p$ for all $\alpha \in A$.

Then it converges to $\sup_{\alpha \in A} x_\alpha$. Similarly, if $x : A \to X$ is a bounded below, decreasing net, then it converges to $\inf_{\alpha \in A} x_\alpha$.

Proof. We only show the supremum version, the infimum version is similar. Since $(x_\alpha)_{\alpha \in A}$ is bounded above, $p = \sup_{\alpha \in A} x_\alpha$ exists. For each neighborhood N of p, there either exists

1. $u, v \in X$ such that $p \in (u, v) \subseteq N$: by definition of supremum, there exist some $\eta \in A$ such that $u < x_\eta \leq p$. Since $(x_\alpha)_{\alpha \in A}$ is increasing, we have $u < x_\eta \leq x_\alpha \leq p$ for all $\alpha \geq \eta$. In particular, $x_\alpha \in (u, v) \subseteq N$ for all $\alpha \geq \eta$.

2. $u \in X$ such that $p \in (u, \infty) \subseteq N$: similarly, to 1st case, there exists some $\eta \in A$ such that $x_\alpha \in N$ for all $\alpha \geq \eta$.

3. $v \in X$ such that $p \in (-\infty, v) \subseteq N$: Since $x_\alpha \leq p$ by definition, we have $x_\alpha \in N$ for all $\alpha \in A$. In particular, pick some $\eta \in A$ and we get $x_\alpha \in N$ for all $\alpha \geq \eta$.

In any case, there always some $\eta \in A$ such that $x_\alpha \in N$ for all $\alpha \geq \eta$. □

Proposition 4.83 (Monotonicity)**.** Let $(x_\alpha)_{\alpha \in A}$ converges to p, $(y_\alpha)_{\alpha \in A}$ converges to q. If $(x_\alpha)_{\alpha \in A}$ is eventually dominated by $(y_\alpha)_{\alpha \in A}$ (i.e. there exists some $\lambda \in A$ such that $x_\alpha \leq y_\alpha$ for all $\alpha \geq \lambda$), then $p \leq q$.

Proof. Suppose otherwise, that $p > q$. Then $(-\infty, p)$ is a neighborhood of q, so there exists some $\gamma \in A$ such that $y_\alpha < p$ for all $\alpha \geq \gamma$. We also have (q, ∞) is a neighborhood of p, so there exists another $\xi \in A$ such that $x_\alpha > q$ for all $\alpha \geq \xi$. Since A is directed, there exists $\eta \in A$ greater than or equal to γ, ξ, and λ. Therefore, $q < x_\alpha \leq y_\alpha < p$ for all $\alpha \geq \eta$. In particular, (q, p) is not empty, so let $r \in (q, p)$. Then $(-\infty, r)$ is a neighborhood of q, so there exists some $\nu \in A$ such that $y_\alpha \in (-\infty, r)$ for all $\alpha \geq \nu$. Vice versa, (r, ∞) is a neighborhood of p, so there exists some $\mu \in A$ such that $x_\alpha \in (r, \infty)$ for all $\alpha \geq \mu$. Let $\lambda \in A$ be an upper bound of η, μ, ν, we then have $x_\alpha > r$ and $x_\alpha \leq y_\alpha < r$ for all $\alpha \geq \lambda$, a contradiction. □

Theorem 4.84 (Squeeze theorem)**.** Let A be a directed set, and $x, y, z : A \to L$ be nets on L such that $x_\alpha \leq y_\alpha \leq z_\alpha$ for all $\alpha \in A$. If $x_\alpha \to p$ and $z_\alpha \to p$, then $y_\alpha \to p$.

Proof. Given any basic open neighborhood N of p (i.e. N is either an interval, or a ray), there must exists $\eta_x, \eta_z \in A$ such that $x_\alpha \in N$ for all $\alpha \geq \eta_x$ and $z_\alpha \in N$ for all $\alpha \geq \eta_z$. Since A is directed, there exists some $\eta \in A$ such that $\eta \geq \eta_x, \eta_z$. Hence, $x_\alpha, z_\alpha \in N$ for all $\alpha \geq \eta$. However, since N is basic (and in particular, convex), $[x_\alpha, z_\alpha] \subseteq N$. Therefore, we also have $y_\alpha \in N$ for all $\alpha \geq \eta$. Since this holds for any basic open neighborhood of p, and the basic open sets generate the topology, we conclude that $y_\alpha \to p$. □

Proposition 4.85. Order-isomorphism is a homeomorphism between order topologies.

Proof. Suppose $f : (P, \leq_P) \to (Q, \leq_Q)$ is an order-isomorphism. In particular, f^{-1} is monotone, so $f^{-1}((b_1, b_2)) = (f^{-1}(b_1), f^{-1}(b_2))$. Similarly, $f^{-1}(b, \infty) = (f^{-1}(b), \infty)$ and $f^{-1}(-\infty, b) = (-\infty, f^{-1}(b))$. We just test on all basic open sets whose inverse images are also open, so f must be continuous. Vice versa, we can show analogously that f^{-1} is continuous. Therefore, f is a homeomorphism. □

Proposition 4.86.

1. If X is not bounded above, any dense subset is cofinal.
2. Dually, if X is not bounded below, any dense subset of coinitial.

As a result, $\operatorname{cf} X \leq d(X)$.

Proof. Let S be a dense set. If it is not cofinal, then there exists some $p \in X$ such that $s \leq p$ for all $s \in S$. Thus, $\operatorname{cl}(S) \subseteq (-\infty, p]$. But since X is not bounded above, there exists some $q > p$, so $q \notin \operatorname{cl}(S)$, a contradiction. The dual version is proved similarly. □

Theorem 4.87. Let L, X be totally ordered sets, with L unbounded (so the cofinality of L is at least ω). If $(x_i)_{i \in L}$ is an increasing sequence in X, then either

1. $(x_i)_{i \in L}$ is eventually constant, i.e. there exists some $p \in X$ and $k \in L$ such that $x_i = p$ for all $i \geq k$.
2. There exists a *strictly* increasing, cofinal subsequence $(y_\alpha)_{\alpha < \kappa}$ (where $\kappa = \operatorname{cf} L$). Note, however, that such subsequence may not be necessarily unbounded in X.

Proof. Denote $S = \{x_i : i \in L\} \subseteq X$. For each $s \in S$, let j_s to be some element in L such that $x_{j_s} = s$. We show that $(j_s)_{s \in S}$ is strictly increasing. By definition, we know that $j_s \neq j_t$ whenever $s \neq t$. If $s < t$ yet $j_s > j_t$, then, by monotonicity of $(x_i)_{i \in L}$, $s = x_{j_s} \geq x_{j_t} = t$, a contradiction.

Now, if $(j_s)_{s \in S}$ is bounded, then let k be an upper bound of it. Denote $p = x_k$, then $p \geq x_{j_s} = s$ for all $s \in S$ by monotonicity. By definition, $p = \max S$, and so, for all $i \geq k$, we get $x_i = p$ by monotonicity.

Otherwise, $(j_s)_{s \in S}$ is unbounded in L, so there exists a cofinal subsequence $(\mu_\alpha)_{\alpha < \kappa}$ of $(j_s)_{s \in S}$, where κ is the cofinality of L. Since an increasing function preserves cofinality, $(x_{\mu_\alpha})_{\alpha < \kappa}$ is a cofinal subsequence of $(x_{j_s})_{s \in S}$. By definition of j_s, $(x_{\mu_\alpha})_{\alpha < \kappa}$ is a strictly increasing, cofinal subsequence of $(x_i)_{i \in L}$. □

Proof Note: this uses axiom of choice to pick out a representative $x_\eta = s$ for each $s \in S$ (and also to pick out a cofinal subsequence of $(j_s)_{s\in S}$). However, if L is well-ordered, then there is no need for choice.

Corollary 4.88. If $\operatorname{cf} L > \operatorname{cf} S$ for all $S \subseteq X$, then any increasing function $f : L \to X$ is eventually constant.

Proof. If it is not eventually constant, then there exists a strictly increasing, cofinal subsequence $(f(i_\alpha))_{\alpha<\kappa}$ (with $\kappa = \operatorname{cf} L$) of $(f(i))_{i\in L}$. But $\kappa > \lambda = \operatorname{cf} f(L)$, so there also exists a strictly increasing, cofinal subsequence $(f(j_\beta))_{\beta<\lambda}$ of $(f(i_\alpha))_{\alpha<\kappa}$. Hence, $(f(j_\beta))_{\beta<\lambda}$ is a strictly increasing, cofinal subsequence of $(f(i))_{i\in L}$.

However, $(j_\beta)_{\beta<\lambda}$ cannot be cofinal in L (since it would contradict the minimality), so there must be an upper bound k of it. Since f is increasing, $f(k) \geq f(j_\beta)$ for all $\beta < \lambda$. But $(f(j_\beta))_{\beta<\lambda}$ is cofinal in $f(L)$, so for each $i \in L$, there exists some $\beta < \lambda$, such that $f(i) \leq f(j_\beta) \leq f(k)$. This implies that f is eventually constant, a contradiction. □

4.6.1 Discrete ordering

Given $a < b$ in X, we say a is the *predecessor* of b (or b is the successor of a) if the open interval (a, b) is empty. We then denote $a <^* b$.

Remark 4.89. The predecessor of $p \in X$, if exists, is unique. Similarly, the successor of p, if exists, is unique.

Proof. Suppose $a, a' <^* p$ (in particular, $a, a' < p$), then $(a, p) = (a', p) = \varnothing$. WLOG, we can assume that $a \leq a'$. If $a < a'$, then $a' \in (a, p)$, a contradiction to $(a, x) = \varnothing$. Therefore, $a = a'$. By similar argument, we can show that the successor of x, if exists, is unique. □

Lemma 4.90. Let X be totally ordered. If we define $\sim$ as the relation such that $a \sim b$ if there is only finitely many points in X between a and b, then

1. $\sim$ is an equivalence relation.
2. $X/\sim$ is totally ordered by: $C < D$ if and only if $x < y$ for all $x \in C, y \in D$.
3. X is the ordered sum of these equivalence classes (the index set is, of course, $X/\sim$).

Such relation $\sim$ is called the $\mathbb{Z}$-*modulo* relation (not to be confused with $\mathbb{Z}$-module, which refers to abelian group).

Proof.
Part 1:

(a) Since (a, a) is empty, $a \sim a$.

(b) If $a \sim b$ then $b \sim a$, since we do not need to know whether $a \leq b$ or $b \geq a$, we just need to count the points in between a and b (so it will be $|(a,b)|$ if $a \leq b$, and $|(b,a)|$ if $a \geq b$).

(c) Suppose $a \sim b$ and $b \sim c$: consider the following case

i. $a \leq b \leq c$: then $(a,c) = (a,b) \cup \{b\} \cup (b,c)$ is finite (since it is a finite union of finite sets).

ii. $a \leq c \leq b$: then $(a,c) = (a,b) \setminus [(c,b) \cup \{c\}]$ is finite (since it is a subset of a finite set).

iii. $b \leq a \leq c$: then $(a,c) = (b,c) \setminus [(b,a) \cup \{a\}]$ is finite.

iv. $b \leq c \leq a$: then $(c,a) = (b,a) \setminus [(b,c) \cup \{c\}]$ is finite.

v. $c \leq b \leq a$: then $(c,a) = (c,b) \cup \{b\} \cup (b,a)$ is finite.

vi. $c \leq a \leq b$: then $(c,a) = (c,b) \setminus [(a,b) \cup \{a\}]$ is finite.

Therefore, $a \sim c$.

Part 2:

(a) Irreflexivity: $C \not< C$ since $x \not< x$ for any $x \in C$.

(b) Transitivity: if $C < D$ and $D < E$, then $x < y < z$ for all $x \in C, y \in D, z \in E$. By transitivity of the order on X, $C < E$.

(c) Totality: given $a \in C$ and $b \in D$, then either $a < b$ or $b > a$ since we assume $C \neq D$ (and so $C \cap D = \varnothing$). WLOG, let's say $a < b$, we will show $C < D$. Suppose otherwise, that $x > y$ for some $x \in C$ and $y \in D$, we consider the following cases

i. $x \leq a$: then $b > a \geq x > y$. By definition, (b,y) contains finitely many points, so (b,a) also contains finitely many point, so $b \sim a$, or equivalently, $C = D$, a contradiction.

ii. $x > a$, $y \geq b$: then $x > y \geq b > a$. By the same argument as in previous case, we derive a contradiction.

iii. $x > a$, $y < b$: note that both (a,x) and (y,b) are finite sets such that $(a,x) \cup (y,b) = (a,b)$ (since $x > y$), so (a,b) is also finite, contradicting $a \not\sim b$.

Part 3: the ordered sum of $C \in X/\sim$ is $\sum_{C \in X/\sim} C = \{(x,C) : x \in C \in X/\sim\}$ (which is a subset of $X \times (X/\sim)$). The ordering is $(x,C) < (y,D)$ if and only if either $C = D$ and $x < y$, or $C < D$. However, $C < D$ implies that $x < y$ by part (2), so the ordering is simplified to $(x,C) < (y,D)$ iff $x < y$. The isomorphism from X to $\sum_{C \in X/\sim} C$ is $f(x) = (x,[x])$, and its inverse is $g(x,C) = x$. Each of them is order-preserving since $x < y$ if and only if $(x,[x]) < (y,[y])$ by definition. $\square$

Lemma 4.91. If X is discrete under the order topology, and p is a point in X, then *exactly* one of the following holds (unless X is a singleton)

1. p is the minimum of X, and the successor of p exists.
2. The predecessor and successor of X exist.
3. p is the maximum of X, and the predecessor of p exists.

Proof. Every $p \in X$ is isolated, so there exists a neighborhood N of p such that $N \cap X = \{p\} = N$. There also exists some $a, b \in X$ such that either

1. $p \in (a, b) \subseteq N$: then we have $(a, b) = \{p\}$. In particular, since (a, b) is the disjoint union of (a, p), $\{p\}$, (p, b), (a, p) and (p, b) must be empty. In other words, $a <^* p$ and $b >^* p$.
2. $p \in (a, \infty) \subseteq N$: then we have $(a, \infty) = \{p\}$. Again, (a, ∞) is the disjoint union of (a, p), $\{p\}$, and (p, ∞), we get $(a, p) = \varnothing$ (i.e. $a <^* p$) and $(p, \infty) = \varnothing$ (i.e. $p = \max X$).
3. $p \in (-\infty, b) \subseteq N$: by similar argument in case 2, we get $p = \min X$ and $b >^* p$.

The minimum cannot have a predecessor (an element before the minimum), and the maximum cannot have a successor. This excludes points that either satisfies (1) and (2), or (2) and (3). If there exists a point that satisfies both (1) and (3), then indeed X is just a singleton (otherwise the maximum and the minimum have to be distinct). □

Theorem 4.92 (Classification of discrete order)**.** If X is discrete under order topology, then it is (or rather, is isomorphic to) either

1. A finite set.
2. $I \times \mathbb{Z}$ (with lexicographical ordering).
3. A sum of $\mathbb{Z}_{\geq 0}$ and $I \times \mathbb{Z}$.
4. A sum of $I \times \mathbb{Z}$ and $\mathbb{Z}_{\leq 0}$.
5. A sum of $\mathbb{Z}_{\geq 0}$, $I \times \mathbb{Z}$, and $\mathbb{Z}_{\leq 0}$.

Furthermore

1. The converse also holds, i.e. the order topology on each of these sets is discrete.
2. These ordered sets are pairwise non-isomorphic, except only the following:
   (a) $I \times \mathbb{Z} \cong J \times \mathbb{Z}$ iff $I \cong J$.
   (b) $\mathbb{Z}_{\geq 0} + I \times \mathbb{Z} \cong \mathbb{Z}_{\geq 0} + J \times \mathbb{Z}$ iff $I \cong J$.
   (c) $I \times \mathbb{Z} + \mathbb{Z}_{\leq 0} \cong J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$ iff $I \cong J$.

(d) $\mathbb{Z}_{\geq 0} + I \times \mathbb{Z} + \mathbb{Z}_{\leq 0} \cong \mathbb{Z}_{\geq 0} + J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$ iff $I \cong J$.

Proof. To reduce the cases that we need to consider, we can assume that X is infinite. With respect to the $\mathbb{Z}$-modulo relation in Lemma 4.90, we shall prove each $C \in X/\sim$ is isomorphic to either $\mathbb{Z}, \mathbb{Z}_{\geq 0}, \mathbb{Z}_{\leq 0}$. Note that

1. If $\min C$ exists: then there is no predecessor of $\min C$ in X (that will contradicts the minimality). According to Lemma 4.91, $\min C$ has to be $\min X$.

2. If $\max C$ exists: then there is no successor of $\max C$ in X. Thus, $\max C$ has to be $\max X$.

So if both $\min C$ and $\max C$ exist, then $C = [\min C, \max C] = [\min X, \max X] = X$ and C is finite. This contradicts X being infinite.

Therefore, an equivalence class C either

1. Has $\min C$, but not $\max C$: let $\mu : C \to \mathbb{Z}_{\geq 0}$ which sends each $x \in C$ to $|[\min C, x)|$ - the number of elements between $\min C$ inclusively, and x exclusively.

   (a) μ is surjective: if $x \in C$ does not have a successor, then $x = \max X$, which contradicts $\max C$ does not exist (Lemma 4.91). Suppose x^+ is the successor of x, then $[\min C, x^+) = [\min C, x) \cup \{x\}$, and so $\mu(x^+) = |[\min C, x)| + |\{x\}| = \mu(x) + 1$. By induction
      i. The base step $n = 0$: $\mu(\min C) = |[\min C, \min C)| = 0$.
      ii. The inductive step: suppose $\mu(x) = n$ for some $x \in C$, then $\mu(x^+) = \mu(x) + 1 = n + 1$.

      we conclude μ is surjective.

   (b) Next, we will show that μ is injective: suppose $\mu(x) = \mu(y)$ but $x < y$. Then $x^+ \leq y$ since there is no point in (x, x^+) by definition of successor. Thus, $[\min C, y)$ contains $[\min C, x^+)$ and so $\mu(y) \geq \mu(x^+) > \mu(x)$, a contradiction.

   (c) Finally, μ is order-preserving (thus, an order-isomorphism): this is in the proof of injection.

   We conclude C is isomorphic to $\mathbb{Z}_{\geq 0}$.

2. Has $\max C$, but $\min C$: similar argument, but use $\mu : C \to \mathbb{Z}_{\leq 0}$ that maps x to $-|(x, \max C]|$ (notice the minus sign) as the isomorphism between C and $\mathbb{Z}_{\leq 0}$.

3. Has neither $\min C$ nor $\max C$: again, according to Lemma 4.91, each $x \in C$ has predecessor and successor. Fix some $p \in C$, and let $\mu : C \to \mathbb{Z}$ send each $x \in C$ to either $|[p, x)|$ if $x \geq p$, or $-|(x, p]|$ if $x \leq p$. Follow the proofs of 1st and 2nd cases, we can show that

(a) $C_{\leq p} = \{x \in C : x \leq p\}$ is order-isomorphic to $\mathbb{Z}_{\leq 0}$: just restrict μ to $C_{\leq p}$.

(b) $C_{\geq p} = \{x \in C : x \geq p\}$ is order-isomorphic to $\mathbb{Z}_{\geq 0}$: just restrict μ to $C_{\geq p}$.

Since $\mu(p)$ is always 0 regardless of either definition (as $(p,p] = [p,p) = \varnothing$), we can patch $C_{\leq p}$ and $C_{\geq p}$ into C to show that μ is an order-isomorphism between C and $\mathbb{Z}$.

*Finally, we show that X is order-isomorphic to either type 2, 3, 4, or 5. Based on whether X is bounded above or below (or both, or neither), we divide into 4 cases, .

1. Neither $\min X$ nor $\max X$ exists: each C is then cannot have maximum or minimum. Therefore, each equivalence class must be isomorphic to $\mathbb{Z}$, and so
$$X = \sum_{C \in X/\sim} C \cong (X/\sim) \times \mathbb{Z}$$

2. $\min X$ exists, but not $\max X$: there is only one (equivalence) class C_i^* with minimum in it, and no class with maximum in it. C_i^* must be less than any other class since $\min X \in C_i^*$. However, C_i^* is isomorphic to $\mathbb{Z}_{\geq 0}$, while other classes are isomorphic to $\mathbb{Z}$. Hence,
$$X = C_i^* + \sum_{\substack{C \in X/\sim \\ C \neq C_i^*}} C \cong \mathbb{Z}_{\geq 0} + [(X/\sim) \setminus \{C_i^*\}] \times \mathbb{Z}$$

3. $\max X$ exists, but not $\min X$: similar to the 2nd case, if we denote C_f^* to be the (only) equivalence class with maximum, then
$$X = \sum_{\substack{C \in X/\sim \\ C \neq C_i^*}} C + C_f^* \cong [(X/\sim) \setminus \{C_f^*\}] \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$$

4. Both $\max X$ and $\min X$ exist: then there is only class C_i^* with minimum, C_f^* with maximum, and every class in between can neither have minimum nor maximum. Thus,
$$X = C_i^* + \sum_{\substack{C \in X/\sim \\ C \neq C_i^*, C_f^*}} C + C_f^* \cong \mathbb{Z}_{\geq 0} + [(X/\sim) \setminus \{C_i^*, C_f^*\}] \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$$

*Converse:

1. If X is finite: after relabeling, we can assume that $X = \{1, 2, 3, \ldots, n\}$ with natural ordering. Thus, $\{1\} = (-\infty, 2)$, $\{n\} = (n-1, \infty)$, and $\{i\} = (i-1, i+1)$ are open.

2. If $X = I \times \mathbb{Z}$: then for any $(i, n) \in I \times \mathbb{Z}$, then (i, n) is sandwiched between $(i, n-1)$ and $(i, n+1)$ so $\{(i, n)\}$ is open.

3. If $X = \mathbb{Z}_{\geq 0} + I \times \mathbb{Z}$, then each point in X is either in $\mathbb{Z}_{\geq 0}$ or in $I \times \mathbb{Z}$. The first case is just $\mathbb{N}$ with order topology, and it is discrete. The second case is proven above.

4. Similar argument for the remaining case.

*Uniqueness:

1. A finite (totally) ordered set is determined uniquely by its cardinality, and it cannot be isomorphic to the other sets in the list since they are all infinite.

2. If $f : I \times \mathbb{Z} \to J \times \mathbb{Z}$ is an order-isomorphism: observe that f has to map $\{i\} \times \mathbb{Z}$ to some $\{j\} \times \mathbb{Z}$. Indeed, for each $n \in \mathbb{Z}$, if $f(i, n+1)$ is not a successor of $f(i, n)$, then there is some (j, m) in between them. However, f^{-1} is also order-preserving, so $f^{-1}(j, m)$ has to be between (i, n) and $(i, n+1)$, which there is none. In other words, $f(i, n+1)$ has to be the successor of $f(i, n)$ for all $i \in I$ and $n \in \mathbb{Z}$. Simply put, $f(i, n+1) = (j, m+1)$ whenever $f(i, n) = (j, m)$. Fix $f(i, 0) = (j, m)$, then by induction, one can show that $f(i, n) = (j, m+n)$ and $f(i, -n) = (j, m-n)$ for all $n \in \mathbb{N}$. Hence, $f(i, n) = (j, m+n)$ for all $n \in \mathbb{Z}$. This shows that $f(\{i\} \times \mathbb{Z}) = \{j\} \times \mathbb{Z}$.

   Define $\varphi : I \to J$ which maps i to j such that $f(\{i\} \times \mathbb{Z}) = \{j\} \times \mathbb{Z}$. This map is

   (a) Well-defined: if $\{j_1\} \times \mathbb{Z} = \{j_2\} \times \mathbb{Z}$, then $(j_1, m_1) = (j_2, m_2)$ for some $m_1, m_2 \in \mathbb{Z}$, and so $j_1 = j_2$.

   (b) Injective: if $\varphi(i_1) = \varphi(i_2) = j$, then $f^{-1}(j, 0) \in \{i_1\} \times \mathbb{Z}$ and $f^{-1}(j, 0) \in \{i_2\} \times \mathbb{Z}$. In particular, $f^{-1}(j, 0) = (i_1, n_1) = (i_2, n_2)$ for some $n_1, n_2 \in \mathbb{Z}$. Thus, $i_1 = i_2$.

   (c) Surjective: given $j \in J$, note the symmetry between f and f^{-1}, we must have $f^{-1}(j \times \mathbb{Z}) = \{i\} \times \mathbb{Z}$ for some $i \in I$. We then get $j \times \mathbb{Z} = f(\{i\} \times \mathbb{Z})$ and $\varphi(i) = j$.

   (d) Order-preserving: if $i_1 < i_2$, let $\varphi(i_1) = j_1, \varphi(i_2) = j_2$. In particular, $f(i_1, 0) = (j_1, m_1)$ and $f(i_2, 0) = (j_2, m_2)$ for some $m_1, m_2 \in \mathbb{Z}$. Since $i_1 < i_2$, $(i_1, 0) < (i_2, 0)$, and so $(j_1, m_1) < (j_2, m_2)$. But $j_1 \neq j_2$ because φ is injective, so we must have $j_1 < j_2$ (due to lexicorgraphical ordering).

Therefore, I and J are order-isomorphic. In other words, $I \times \mathbb{Z}$ is not isomorphic to $J \times \mathbb{Z}$ whenever I is not isomorphic to J.

3. Suppose $f : \mathbb{Z}_{\geq 0} + I \times \mathbb{Z} \to \mathbb{Z}_{\geq 0} + J \times \mathbb{Z}$ is an isomorphism: we will use $n \in \mathbb{Z}$ to denote the elements of initial segment $\mathbb{Z}_{\geq 0}$, and (i, n) (or (j, m) to denote the elements of $I \times \mathbb{Z}$ (or $J \times \mathbb{Z}$).

   Since 0 is the minimum of $\mathbb{Z}_{\geq 0} + I \times \mathbb{Z}$, $f(0)$ has to be the minimum of $\mathbb{Z}_{\geq 0} + J \times \mathbb{Z}$. Otherwise, if $a < f(0)$ for some $a \in \mathbb{Z}_{\geq 0} + J \times \mathbb{Z}$, then $f^{-1}(a) < 0$, a contradiction. With the same successor argument as before, we can show by induction that $f(n) = n$ for all $n \in \mathbb{Z}_{\geq 0}$. Hence, by cutting out the initial segment, we obtain an isomorphism between $I \times \mathbb{Z}$ and $J \times \mathbb{Z}$. Therefore, I is isomorphic to J.

4. Similarly, we can show that

   (a) If $I \times \mathbb{Z} + \mathbb{Z}_{\leq 0} \cong J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$, then $I \cong J$.
   (b) If $\mathbb{Z}_{\geq 0} + I \times \mathbb{Z} + \mathbb{Z}_{\leq 0} \cong \mathbb{Z}_{\geq 0} + J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$, then $I \cong J$.

   by first prove the initial segments are isomorphic and final segments are isomorphic, then cut out those to obtain an isomorphism of $I \times \mathbb{Z}$ and $J \times \mathbb{Z}$.

5. Suppose $f : I \times \mathbb{Z} \to \mathbb{Z}_{\geq 0} + J \times \mathbb{Z}$ is an isomorphism. Then $f^{-1}(0) = (i, n)$ has to be the minimum of $I \times \mathbb{Z}$, which is not since $(i, n-1) < (i, n)$. Similar argument for

   (a) $f : I \times \mathbb{Z} \to J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$.
   (b) $f : I \times \mathbb{Z} \to \mathbb{Z}_{\geq 0} + J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$.
   (c) $f : \mathbb{Z}_{\geq 0} + I \times \mathbb{Z} \to J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$.
   (d) $f : \mathbb{Z}_{\geq 0} + I \times \mathbb{Z} \to \mathbb{Z}_{\geq 0} + J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$.
   (e) $f : I \times \mathbb{Z} + \mathbb{Z}_{\leq 0} \to \mathbb{Z}_{\geq 0} + J \times \mathbb{Z} + \mathbb{Z}_{\leq 0}$.

   by looking at minimum/maximum element.

□

Note: some author may use $\mathbb{N}$ instead of $\mathbb{Z}_{\geq 0}$ and $\mathbb{N}^*$ (the reverse order of $\mathbb{N}$) instead of $\mathbb{Z}_{\leq 0}$.

4.6.2 One-sided limit

Let $f : X \to Y$ be a function from a (totally) ordered set X to a topological space Y. Suppose further, for the sake of convenience, that p is neither the minimum nor the maximum of X. Recall from Example 4.77 that

1. q is a left-sided limit of f at p, if for all neighborhood N of q, there exists some $a < p$ such that $f(x) \in N$ for all $a < x < p$.

2. q is a right-sided limit of f at p, if for all neighborhood N of q, there exists some $b > p$ such that $f(x) \in N$ for all $p < x < b$.

3. q is a limit of f at p, if for a neighborhood N of $f(p)$, there exists some $a < p < b$ such that $f(x) \in N$ for all $a < x < b$, $x \neq p$.

   Note that this definition is the same one as in Example 4.77 because of

   (a) The open intervals and open rays form the basis of X.

   (b) p is neither the maximum nor the maximum of X, so there is always some open interval lies inside an open ray.

Proposition 4.93 (Existence of limit)**.** q is a limit of f at p if and only if it is a left-sided and a right-sided limit at p.

Proof. The forward direction is straight from definition. So suppose q is the left-sided and right-sided limit of f at p. Then for all neighborhood N of q, there exists some $a < p$ and $b > p$ such that $f(x) \in N$ for all $x \in (a, p) \cup (b, p) = (a, b) \setminus \{p\}$. By definition, q is the limit of f at p. □

4.6.3 Limit inferior and limit superior

Let $(x_\alpha)_{\alpha \in A}$ be a net in a complete ordered set X (i.e. Dedekind-complete ordered set with maximum and minimum).

1. The *limit inferior* of $(x_\alpha)_{\alpha \in A}$ is defined as

$$\liminf_{\alpha \in A} x_\alpha = \lim_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta$$

2. The *limit inferior* of $(x_\alpha)_{\alpha \in A}$ is defined as

$$\limsup_{\alpha \in A} x_\alpha = \lim_{\alpha \in A} \sup_{\beta \geq \alpha} x_\beta$$

Remark 4.94. Both limits

1. Are well-defined because of the completeness of X.

2. Exist because of the monotone convergence theorem. In fact,

   (a) $\liminf_{\alpha \in A} x_\alpha = \sup_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta$.

   (b) $\limsup_{\alpha \in A} x_\alpha = \inf_{\alpha \in A} \sup_{\beta \geq \alpha} x_\beta$.

   So the limit inferior and the limit superior are unique.

Proof. Since $\{x_\beta : \beta \geq \alpha_1\} \subseteq \{x_\beta : \beta \geq \alpha_2\}$ whenever $\alpha_1 \geq \alpha_2$, we get $\inf_{\beta \geq \alpha_1} x_\beta \geq \inf_{\beta \geq \alpha_2} x_\beta$ and $\sup_{\beta \geq \alpha_1} x_\beta \leq \sup_{\beta \geq \alpha_2} x_\beta$. We also have $\inf_{\beta \geq \alpha} x_\alpha$ and $\sup_{\beta \geq \alpha} x_\alpha$ sandwiched between the minimum and the maximum of X. □

Proposition 4.95 (Equivalent definition)**.** Let $(x_\alpha)_{\alpha\in A}$ be a net and p be a point in X

1. TFAE
   (a) p is a limit inferior of $(x_\alpha)_{\alpha\in A}$.
   (b) p is the supremum of all eventual lower bounds of the net.
   (c) $x_\alpha > r$ eventually for all $r < p$, and $x_\alpha < s$ frequently for all $s > p$.
2. TFAE
   (a) p is a limit superior of $(x_\alpha)_{\alpha\in A}$
   (b) p is the infimum of all eventual upper bounds of the net.
   (c) $x_\alpha > r$ frequently for all $r < p$, and $x_\alpha < s$ eventually for all $s > p$.

Proof. Part (2) is just the dual version of part (1), so the proof is similar once we show part (1) is true.
Part 1: we prove the following direction $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$.

*Suppose p is a limit inferior $(x_\alpha)_{\alpha\in A}$, or $p = \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta$. If s is an eventual lower bound, then there exists $\eta \in A$ such that $x_\beta \geq s$ for all $\beta \geq \eta$. In other words, $s \leq \inf_{\beta\geq\eta} x_\beta$, so $s \leq \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta = p$. On the other hand, any $s < p$ is an eventual lower bound. Indeed, by definition of supremum, there exists some $\eta \in A$ such that $s < \inf_{\beta\geq\eta} x_\beta$. So p is the supremum of all eventual lower bounds.

*Suppose p is the supremum of all eventual lower bounds of the net. If $r < p$, then there is some eventual lower bound l of the net such that $r < l \leq p$. Equivalently, there is some $\lambda \in A$ such that $x_\alpha \geq l > r$ for all $\alpha \geq \lambda$. In particular, $x_\alpha > r$ eventually. On the other hand, suppose $s > p$, yet $x_\alpha < s$ not frequently. In other words, $x_\alpha \geq s$ eventually (Lemma 4.39), so $s \leq p$ by definition of p, a contradiction.

*Suppose p is such that $x_\alpha > r$ eventually for all $r < p$, and $x_\alpha < s$ frequently for all $s > p$. If $p < \liminf_{\alpha\in A} x_\alpha = \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta$, then there exists some $\eta \in A$ such that $p < \inf_{\beta\geq\eta} x_\beta$. If $s = \inf_{\beta\geq\eta} x_\beta$, then $x_\alpha < s$ frequently. In particular, there exists some $\beta \geq \eta$ such that $x_\beta < s$. But $x_\beta \geq \inf_{\beta\geq\eta} x_\eta$ by definition, so this is a contradiction.

Vice versa, if $p > \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta$, let $r = \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta$. Then $x_\alpha > r$ eventually, i.e. there exists some $\eta \in A$ such that $x_\alpha > \sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta$ for all $\alpha \geq \eta$. Consider the following cases

1. If (r, p) is empty: then $x_\alpha \geq p$ for all $\alpha \geq \eta$. Hence, $\inf_{\beta\geq\eta} x_\beta \geq p$, and so $\sup_{\alpha\in A} \inf_{\beta\geq\alpha} x_\beta \geq p$, contradicting the assumption.

2. If (r, p) is not empty: then let $s \in (r, p)$, and $\theta \in A$ be such that $x_\alpha > s$ for all $\alpha \geq \theta$. Thus, $\inf_{\beta \geq \theta} x_\beta \geq s$, and so $\sup_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta \geq s$, contradicting $s \in (r, p)$.

□

Proposition 4.96. Let $(x_\alpha)_{\alpha \in A}$ be a net in X

1. $\inf_{\alpha \in A} x_\alpha \leq \liminf_{\alpha \in A} x_\alpha \leq \limsup_{\alpha \in A} x_\alpha \leq \sup_{\alpha \in A} x_\alpha$
2. If E is the set of cluster points of $(x_\alpha)_{\alpha \in A}$ (or rather cluster points of the eventuality filter base), then $\liminf_{\alpha \in A} x_\alpha = \min E$ and $\limsup_{\alpha \in A} x_\alpha = \max E$. In particular, E cannot be empty (as long as the space is order-complete).
3. If $(y_\alpha)_{\alpha \in A}$ is another net (with the same directed set) such that it dominates $(x_\alpha)_{\alpha \in A}$ (i.e. $y_\alpha \geq x_\alpha$ eventually), then
$$\limsup_{\alpha \in A} y_\alpha \geq \limsup_{\alpha \in A} x_\alpha$$
$$\liminf_{\alpha \in A} y_\alpha \geq \liminf_{\alpha \in A} x_\alpha$$
4. $(x_\alpha)_{\alpha \in A}$ converges to p if and only if $p = \limsup_{\alpha \in A} x_\alpha = \liminf_{\alpha \in A} x_\alpha$.

Proof.
Part 1: since $\inf_{\alpha \in A} x_\alpha$ is a lower bound of $(x_\alpha)_{\alpha \in A}$, we have $\liminf_{\alpha \in A} x_\alpha \geq \inf_{\alpha \in A} x_\alpha$ (Proposition 4.95). Similarly, we get $\limsup_{\alpha \in A} \leq \sup_{\alpha \in A} x_\alpha$. As for the remaining inequality $\liminf_{\alpha \in A} x_\alpha \leq \limsup_{\alpha \in A} x_\alpha$, we can instead show $\sup_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta \leq \inf_{\alpha \in A} \sup_{\beta \geq \alpha} x_\beta$. For any $\eta, \gamma \in A$, there exists $\theta \in A$ such that $\theta \geq \eta, \gamma$. Hence,
$$\inf_{\beta \geq \eta} x_\beta \leq \inf_{\beta \geq \theta} x_\beta \leq \sup_{\beta \geq \theta} x_\beta \leq \sup_{\beta \geq \gamma} x_\beta$$
Since η, γ are arbitrary, we get $\sup_{\eta \in A} \inf_{\beta \geq \eta} x_\beta \leq \inf_{\gamma \in A} \sup_{\beta \geq \gamma} x_\beta$.

Part 2: let p be a cluster point of $(x_\alpha)_{\alpha \in A}$, or equivalently, for any neighborhood N of p and $\alpha \in A$, there exists some $\eta \geq \alpha$ such that $x_\eta \in N$. If $p < \liminf_{\alpha \in A} x_\alpha$, then there exists some $\lambda \in A$ such that $r = \inf_{\beta \geq \lambda} x_\beta > p$, or $x_\beta > p$ for all $\beta \geq \lambda$. Set $N = (-\infty, r)$ and $\alpha = \lambda$, we get some $\eta \geq \lambda$ such that $x_\eta \in N$. In other words, $x_\eta < \inf_{\beta \geq \lambda} x_\beta$ for some $\eta \geq \lambda$, a contradiction. Similarly, asssuming $p > \limsup_{\alpha \in A} x_\alpha$ will lead to a contradiction, so p must be between the limit inferior and the limit superior of $(x_\alpha)_{\alpha \in A}$.

*To finish, we need to show that those limits are also cluster points of $(x_\alpha)_{\alpha \in A}$. Let $m = \liminf_{\alpha \in A} x_\alpha$, N be a neighborhood of m, and $\alpha \in A$

1. If N contains (r, ∞) for some $r < m$: then there exists some $\lambda \in A$ such that $r < \inf_{\beta \geq \lambda} x_\beta$, or $x_\beta > r$ for all $\beta \geq \lambda$. Since A is directed, there must be some $\eta \geq \alpha$ and $\eta \geq \lambda$, so we have $x_\eta > r$. In other words, $x_\eta \in N$ for some $\eta \geq \alpha$.

2. If N contains $(-\infty, s)$ for some $s > m$: then $s > \inf_{\beta \geq \alpha} x_\beta$ in particular. By definition of infimum, there exists some $\eta \geq \alpha$ such that $x_\eta < s$, or $x_\eta \in N$.
3. Finally, if N contains (r, s) for some $r < m < s$: first, let $\lambda \in A$ such that $r < \inf_{\beta \geq \lambda} x_\beta$. Again A is directed, so there exists some $\theta \geq \alpha$ and $\theta \geq \lambda$. On the other hand, $s > \sup_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta$, so we get $s > \inf_{\beta \geq \theta} x_\beta$. By definition, there is some $\eta \geq \theta \geq \alpha$ such that $x_\eta < s$. But since $\eta \geq \theta \geq \lambda$, we also have $x_\eta > r$ (recall $r < \inf_{\beta \geq \lambda} x_\beta$ at the beginning). In summary, we get $x_\eta \in N$ for some $\eta \geq \alpha$.

In any case, we get $x_\eta \in N$ for some $\eta \geq \alpha$. As N and α varies, $m = \liminf_{\alpha \in A} x_\alpha$ is a cluster point of $(x_\alpha)_{\alpha \in A}$. Similarly, $M = \limsup_{\alpha \in A} x_\alpha$ is also a cluster point of $(x_\alpha)_{\alpha \in A}$.

Part 3: let $\eta \in A$ be such that $y_\alpha \geq x_\alpha$ for all $\alpha \geq \eta$. From there, we get $\inf_{\beta \geq \alpha} y_\beta \geq \inf_{\beta \geq \alpha} x_\beta$ for all $\alpha \geq \eta$. Hence the net $(\inf_{\beta \geq \alpha} y_\beta)_{\alpha \in A}$ dominates $(\inf_{\beta \geq \alpha} x_\alpha)_{\alpha \in A}$. By Proposition 4.83, $\lim_{\alpha \in A} \inf_{\beta \geq \alpha} y_\beta \geq \lim_{\alpha \in A} \inf_{\beta \geq \alpha} x_\beta$. Similarly, $\limsup_{\alpha \in A} y_\alpha \geq \limsup_{\alpha \in A} x_\alpha$.

Part 4: if $p = \limsup_{\alpha \in A} x_\alpha = \liminf_{\alpha \in A} x_\alpha$, then for any $r < p < s$, we have $x_\alpha > r$ and $x_\alpha < s$ eventually (Proposition 4.95). Since (r, s) is a basic open neighborhood of p (with $(r, s) \subseteq (r, \infty), (-\infty, s)$), we have $x_\alpha \in N$ eventually for each neighborhood N of p. By definition, $(x_\alpha)_{\alpha \in A}$ converges to p.

If $(x_\alpha)_{\alpha \in A}$ converges to p, then for all $r < p < s$, x_α is eventually in (r, s). Since eventual implies frequent (Lemma 4.39), we know that $x_\alpha > r$ eventually, and $x_\alpha < s$ frequently. Thus, as r, s vary, p must be the limit inferior of $(x_\alpha)_{\alpha \in A}$. Similarly, p is the limit superior of $(x_\alpha)_{\alpha \in A}$. □

The *oscillator* of the net $(x_\alpha)_{\alpha \in A}$ is defined to be the closed interval with limit supremum and limit infimum as the endpoints.

Remark 4.97. The oscillator is always non-empty (as long as X is completely ordered). It contains exactly 1 element if and only if the net is convergent.

We can also define limit supremum and infimum for a function $f : X \to L$ from a topological space X to a complete totally ordered L. Let $\mathcal{B} \to p$, then

1. The limit supremum of f at $p \in X$ with respect to $\mathcal{B}$ (or $\mathcal{B}$-limit supremum) is the limit of the net $\sup f(B)$ indexed by $B \in \mathcal{B}$. Notationally
$$\limsup_{\mathcal{B} \to p} f(\mathcal{B}) = \lim_{B \in \mathcal{B}} \sup_{x \in B} f(x)$$
2. The limit infimum of f at $p \in X$ with respect $\mathcal{B}$ (or $\mathcal{B}$-limit supremum) is the limit of the net $\inf f(N)$ indexed by neighborhoods N at p. Notationally
$$\liminf_{\mathcal{B} \to p} f(\mathcal{B}) = \lim_{B \in \mathcal{B}} \inf_{x \in B} f(x)$$

Again, one can verify that these are well-defined and exist (hint: use monotone convergence theorem). In fact, $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$ and $\liminf_{\mathcal{B}\to p} f(\mathcal{B}) = \sup_{B\in\mathcal{B}} \inf_{x\in B} f(x)$.

Proposition 4.98.

1. TFAE

   (a) u is the limit supremum of f at p with respect to $\mathcal{B}$.

   (b) u is the infimum of upper bounds of $f(B)$ for all $B \in \mathcal{B}$.

   (c) For all $s > u$, there exists $B \in \mathcal{B}$ such that $f(x) < s$ for all $x \in B$. And for all $r < u$, and for all $B \in \mathcal{B}$, there exists some $x \in B$ such that $f(x) > r$.

   Dually, TFAE

   (a) l is the limit infimum of f at p.

   (b) l is the supremum of lower bounds of $f(B)$ for all $B \in \mathcal{B}$.

   (c) For all $r < l$, there exists $B \in \mathcal{B}$ such that $f(x) > r$ for all $x \in B$. And for all $s > l$, and for all $B \in \mathcal{B}$, there exists some $x \in B$ such that $f(x) < s$.

2. $\inf_{x\in X} f(x) \le \liminf_{\mathcal{B}\to p} f(\mathcal{B}) \le \limsup_{\mathcal{B}\to p} f(\mathcal{B}) \le \sup_{x\in X} f(x)$

3. If $E_{\mathcal{B}}$ denotes the set of cluster points of $f(\mathcal{B})$, then $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \max E_{\mathcal{B}}$ and $\liminf_{\mathcal{B}\to p} f(\mathcal{B}) = \min E_{\mathcal{B}}$.

4. If $\mathcal{C}$ refines $\mathcal{B}$ (and both converges at p), then

$$\limsup_{\mathcal{C}\to p} f(\mathcal{C}) \le \limsup_{\mathcal{B}\to p} f(\mathcal{B})$$
$$\liminf_{\mathcal{C}\to p} f(\mathcal{C}) \ge \liminf_{\mathcal{B}\to p} f(\mathcal{B})$$

5. If $f(x) \le g(x)$ for all $x \in X$, then

$$\limsup_{\mathcal{B}\to p} f(\mathcal{B}) \le \limsup_{\mathcal{B}\to p} g(\mathcal{B})$$
$$\liminf_{\mathcal{B}\to p} f(\mathcal{B}) \le \liminf_{\mathcal{B}\to p} g(\mathcal{B})$$

6. TFAE

   (a) f is continuous at p.

   (b) $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \liminf_{\mathcal{B}\to p} f(\mathcal{B})$

   (c) $\limsup_{\mathcal{N}_p\to p} f(\mathcal{N}_p) = \liminf_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$, where $\mathcal{N}_p$ is the neighborhood filter at p.

   If one of these holds, then we also have $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \liminf_{\mathcal{B}\to p} f(\mathcal{B}) = f(p)$

Proof.
Part 1: we only show the limit supremum part, with $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$ direction.

*Suppose u is the limit supremum of f at p with respect to $\mathcal{B}$. Then $u = \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$. If s is an upper bound of $f(B)$, then $s \geq \sup_{x\in B} f(x)$, and so $s \geq u$ by definition of infimum. On the other hand, whenever $s > u$, then there exists some $B_s \in \mathcal{B}$ such that $s > \sup_{x\in B_s} f(x)$. Thus, $r = \sup_{x\in B_s} f(x) < s$ is an upper bound of $f(B_s)$. By definition, u is the infimum of upper bounds of $f(B)$ for all $B \in \mathcal{B}$.

*Suppose u is the infimum of upper bounds of $f(B)$ for all $B \in \mathcal{B}$. If $s > u$, then there is some $s' \in L$ and $B \in \mathcal{B}$ such that $s > s' \geq f(x)$ for all $x \in B$. On the other hand, if $r < u$, yet there exists $B_0 \in \mathcal{B}$ such that $f(x) \leq r$ for all $x \in B_0$. Equivalently $\sup_{x\in B_0} f(x) \leq r < u$, which contradicts the hypothesis since r is smaller than u while being an upper bound of $f(B_0)$.

*Suppose

1. For all $s > u$, there exists $B \in \mathcal{B}$ such that $f(x) < s$ for all $x \in B$.
2. For all $r < u$, and for all $B \in \mathcal{B}$, there exists some $x \in B$ such that $f(x) > r$.

If $u > \limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$, then there exists some $B_0 \in \mathcal{B}$ such that $u > \sup_{x\in B_0} f(x)$. Let $r = \sup_{x\in B_0} f(x)$, then by the hypothesis, there must be some $x_0 \in B_0$ such that $f(x_0) > r$, which contradicts $\sup_{x\in B_0} f(x) \geq f(x_0)$.

Vice versa, if $u < \limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$, let $s = \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$. Then there exists some $B_0 \in \mathcal{B}$ such that $f(x) < s$ for all $x \in B_0$. Consider the following cases.

1. If (u, s) is empty: then $f(x) \leq u$ for all $x \in B_0$. Hence, $\sup_{x\in B_0} f(x) \leq u$, which means $\inf_{B\in\mathcal{B}} \sup_{x\in B} f(x) \leq u$, contradicting the assumption.
2. If (u, s) contains some $t \in L$: by the hypothesis, there exists some $B_0 \in \mathcal{B}$ such that $f(x) < t$ for all $x \in B_0$. Thus, $\sup_{x\in B_0} f(x) \leq t$, so $u < \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x) \leq t$, contradicting $t < s = \limsup_{\mathcal{B}\to p} f(\mathcal{B})$.

Each case leads to a contradiction, so we must have $u \geq \limsup_{\mathcal{B}\to p} f(\mathcal{B})$. Combining with the prior result, we get $u = \limsup_{\mathcal{B}\to p} f(\mathcal{B})$.

Part 2: since $B \subseteq X$, $\inf_{x\in B} f(x) \geq \inf_{x\in X} f(x)$. Hence, since $\mathcal{B}$ is non-empty, $\liminf_{\mathcal{B}\to p} f(\mathcal{B}) \geq \inf_{x\in B_0} f(x) \geq \inf_{x\in X} f(x)$ (fix some $B_0 \in \mathcal{B}$). Similarly, $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) \leq \sup_{x\in X} f(x)$. As for the remaining inequality, for any $B_1, B_2 \in \mathcal{B}$, let $B \in \mathcal{B}$ be such that $B \subseteq B_1 \cap B_2$. From there, we get $\sup_{x\in B_1} f(x) \geq \sup_{x\in B} f(x) \geq \inf_{x\in B} f(x) \geq \inf_{x\in B_2} f(x)$. As B_1 and B_2 vary, we must have

$\inf_{B\in\mathcal{B}} \sup_{x\in B} f(x) \geq \sup_{B\in\mathcal{B}} \inf_{x\in B} f(x)$.

Part 3: let p be a cluster point of $f(\mathcal{B})$. If $p < \liminf_{\mathcal{B}\to p} f(\mathcal{B})$, then there exists some $B_0 \in \mathcal{B}$ such that $p < r = \inf_{x\in B_0} f(x)$. Let $N = (-\infty, r)$ be a neighborhood of p, and $B = B_0$, then $f(B_0)\cap(-\infty, r)$ is not empty. Thus, there is some $x_0 \in B_0$ such that $f(x_0) < r$, which contradicts $\inf_{x\in B_0} f(x) \geq f(x_0)$. Similarly, we can show $p \leq \limsup_{\mathcal{B}\to p} f(\mathcal{B})$ through contradiction.

*To finish, we show that these limit extrema are also cluster point of $f(\mathcal{B})$. Denote $u = \limsup_{\mathcal{B}\to p} f(\mathcal{B})$, N be a neighborhood of m, and $B \in \mathcal{B}$

1. If N contains $(-\infty, s)$ for some $s > u$: there exists some $B_0 \in \mathcal{B}$ such that $s > \sup_{x\in B_0} f(x)$, or $f(x) < s$ for all $x \in B_0$. Since $\mathcal{B}$ is a filter base, there is some $B' \in \mathcal{B}$ such that $B' \subseteq B \cap B_0$, so $f(x) < s$ for all $x \in B'$. In other words, $f(B') \subseteq N$, so $f(B) \cap N \supseteq f(B')$.

2. If N contains (r, ∞) for some $r < u$: then $r < \sup_{x\in B} f(x)$ in particular. By definition of supremum, there is some $z \in B$ such that $r < f(z)$, so $f(B) \cap N$ contains $f(z)$.

3. If N contains (r, s) for some $r < u < s$: there exists some $B_0 \in \mathcal{B}$ such that $s > \sup_{x\in B_0} f(x)$, or $f(x) < s$ for all $x \in B_0$. There also exists some $B' \in \mathcal{B}$ such that $B' \subseteq B \cap B_0$, so $f(x) < s$ for all $x \in B'$. On the other hand, $r < \inf_{B\in\mathcal{B}} \sup_{x\in B} f(x)$, so $r < \sup_{x\in B'} f(x)$ in particular. Hence, there exists some $x' \in B'$ such that $f(x') > r$. Since $f(x') < s$ (as $x' \in B' \subseteq B_0$) We conclude that $f(B) \cap N$ contains $f(x')$.

In any case, $f(B) \cap N$ is always non-empty, so we conclude $\limsup_{\mathcal{B}\to p} f(\mathcal{B})$ is a cluster point of $f(\mathcal{B})$. Similarly, $\liminf_{\mathcal{B}\to p} f(\mathcal{B})$ is a cluster point of $f(\mathcal{B})$.

Part 4: for each $B \in \mathcal{B}$, there exists $C \in \mathcal{C}$ such that $C \subseteq B$. This means $\sup_{x\in C} f(x) \leq \sup_{x\in B} f(x)$. Let $\mathcal{C}_0 = \{C \in \mathcal{C} : C \subseteq B \text{ for some } B \in \mathcal{B}\}$, we then get $\inf_{B\in\mathcal{B}} \sup_{x\in B} f(x) \geq \inf_{C\in\mathcal{C}_0} \sup_{x\in C} f(x) \geq \inf_{C\in\mathcal{C}} \sup_{x\in C} f(x)$. Similarly, $\liminf_{\mathcal{B}\to p} f(\mathcal{B}) \leq \limsup_{\mathcal{C}\to p} f(\mathcal{C})$.

Part 5: by the hypothesis, $\sup_{x\in B} f(x) \leq \sup_{x\in B} g(x)$ for each $B \in \mathcal{B}$. Thus, $\inf_{B\in\mathcal{B}} \sup_{x\in B} f(x) \leq \inf_{B\in\mathcal{B}} \sup_{x\in B} g(x)$. Similarly, $\liminf_{\mathcal{B}\to p} f(\mathcal{B}) \leq \liminf_{\mathcal{B}\to p} g(\mathcal{B})$.

Part 6: we prove $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$.

*If f is continuous at p: then $f(\mathcal{B}) \to p$ whenever $\mathcal{B} \to p$ (Proposition 4.74).

1. For the neighborhood $(-\infty, s)$, there exists $B_s \in \mathcal{B}$ such that $f(B_s) \subseteq (-\infty, s)$.

2. For the neighborhood (r, ∞), there exists $B_r \in \mathcal{B}$ such that $f(B_r) \subseteq (r, \infty)$. Given $B \in \mathcal{B}$, there exists some $B_0 \in \mathcal{B}$ such that $B_0 \subseteq B_r \cap B$. From there, $f(B_0) \subseteq f(B_r \cap B) \subseteq f(B_r) \cap f(B) \subseteq (r, \infty)$. Since $B_0 \neq \varnothing$, $f(B) \subseteq (r, \infty) \neq \varnothing$ for arbitrary $B \in \mathcal{B}$.

These 2 points characterize $f(p) = \limsup_{\mathcal{B}\to p} f(\mathcal{B})$ (part 1). Similarly, $f(p) = \liminf_{\mathcal{B}\to p} f(\mathcal{B})$.

*If $\limsup_{\mathcal{B}\to p} f(\mathcal{B}) = \liminf_{\mathcal{B}\to p} f(\mathcal{B})$ for all $\mathcal{B} \to p$, then $\limsup_{\mathcal{N}_p\to p} f(\mathcal{N}_p) = \liminf_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$ in particular (as $\mathcal{N}_p \to p$).

*If $\limsup_{\mathcal{N}_p\to p} f(\mathcal{N}_p) = \liminf_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$, denote such limit to be q. If $f(p) < q$, then there exists a neighborhood N_f of p such that $f(x) > f(p)$ for all $x \in N_f$ (since $q = \liminf_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$). In particular, $f(p) > f(p)$ (since $p \in N_f$), a contradiction. Similarly, if $f(p) > q = \limsup_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$, then there exists a neighborhood N_f of p such that $f(x) < f(p)$ for all $x \in N_f$, which also leads to a contradiction. We conclude that $q = f(p)$, and so for an arbitrary neighborhood U of $f(p)$

1. If U contains (r, ∞) for some $r < f(p)$: there exists a neighborhood N of p such that $f(N) \subseteq (r, \infty) \subseteq U$ (again, $f(p) = \liminf_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$). In other words, $N \subseteq f^{-1}(U)$ and so $f^{-1}(U)$ is also a neighborhood of p.

2. If U contains $(-\infty, s)$ for some $s > f(p)$: similarly, because of $f(p) = \limsup_{\mathcal{N}_p\to p} f(\mathcal{N}_p)$, there exists a neighborhood N of p such that $f(N) \subseteq (-\infty, s) \subseteq U$. We then also have $f^{-1}(U)$ as a neighborhood of p.

3. If U contains (r, s) for some $r < f(p) < s$: there exists neighborhoods N_r, N_s of p such that $f(N_r) \subseteq (r, \infty)$, and $f(N_s) \subseteq (-\infty, s)$. Then $N = N_r \cap N_s$ is a neighborhood of p such that $f(N) \subseteq f(N_r) \cap f(N_s) \subseteq (r, \infty) \cap (-\infty, s) = (r, s) \subseteq U$. So $f^{-1}(U)$ is a neighborhood of p.

In any case, $f^{-1}(U)$ is a neighborhood of p. As U varies, f is continuous at p by definition. □

4.7 Subspace Topology

Let S be a subset of a topological space X, then the *subspace topology* on S (relative to X) is the following collection $\mathcal{T}_S = \{U \cap S : U \text{ is open in } X\}$

Remark 4.99. It is indeed a topology on S.

Proof.

1. Any set in $\mathcal{T}_S$ is of the form $U \cap S$, so it must be a subset of S.
2. $X \cap S = S \in \mathcal{T}_S$ and $\varnothing \cap S = \varnothing \in \mathcal{T}_S$.
3. If U, V are open set in X, then $U \cap V$ is open in X and so $(U \cap S) \cap (V \cap S) = (U \cap V) \cap S \in \mathcal{T}_S$.
4. If $\{U_i\}_{i \in I}$ is a family of open sets in X, then $\bigcup_{i \in I} U_i$ is open in X and so $\bigcup_{i \in I}(U \cap S) = \left(\bigcup_{i \in I} U_i\right) \cap S \in \mathcal{T}_S$ by distributivity.

□

Proposition 4.100 (Basic properties)**.**

1. A set C is closed in S if and only if there exists some closed set D in X such that $C = D \cap S$.
2. If S is open, then U is open in S if and only if it is open in X.
   Vice versa, if S is closed, then C is closed in S if and only if it is closed in X.
3. If $\mathcal{T}_1 \subseteq \mathcal{T}_2$ are topologies on X, and $\mathcal{S}_1, \mathcal{S}_2$ are respectively the subspace topologies on X relative to $\mathcal{T}_1, \mathcal{T}_2$, then $\mathcal{S}_1 \subseteq \mathcal{S}_2$.
4. If $p \in S$ and N is a neighborhood of p with respect to X, then $N \cap S$ is a neighborhood of p with respect to S.
5. If $\mathcal{B}$ is a basis of X, then $\mathcal{B}_S = \{B \cap S : B \in \mathcal{B}\}$ is a basis of S.
   As a result, $w(S) \leq w(X)$.
6. If $\mathcal{G}$ generates X, then $\mathcal{G}_S = \{G \cap S : G \in \mathcal{G}\}$ generates X.

Proof. Part 1: if C is closed in S, then $U = S \setminus C$ is open in S, so there exists some V open in X such that $U = V \cap S$. Let $D = X \setminus V$ be a closed set in X, then $D \cap S = (X \setminus V) \cap S = S \setminus V = S \setminus (V \cap S) = S \setminus U = C$. If $C = D \cap S$ for some closed set D in X, then $S \setminus C = S \setminus (D \cap S) = S \setminus D = S \cap (X \setminus D)$, so $S \setminus C$ must be open in S. Equivalently, C is closed in S.
Part 2: the closed case is proven similarly to the open case (with the help of part 1), so we only prove the latter one. Given S open, if U is open in S, then there exists some open V in X such that $U = V \cap S$. But the intersection of 2 open sets is open, so U is open in X. Vice versa, if U is open in X, then $U = U \cap S$ (since $U \subseteq S$), so U is open in S.

Part 3: if $U \in \mathcal{S}_1$, then there exists some $V \in \mathcal{T}_1$ such that $U = V \cap S$. But $\mathcal{T}_1 \subseteq \mathcal{T}_2$, so $U = V \cap S$ for some $V \in \mathcal{T}_2$. In other words, $U \in \mathcal{S}_2$.
Part 4: if N is a neighborhood of p with respect to X, then there exists some open subset U of N in X containing p. Then $p \in U \cap S \subseteq N \cap S$ and $U \cap S$ is open in S, so $N \cap S$ is a neighborhood of p with respect to S.
Part 5: by definition of subspace topology, any set in $\mathcal{B}_S$ is open. On the other hand, given any open neighborhood V of p in S, there exists some U open in X such that $V = U \cap S$. In particular, U is an open neighborhood of p, so there exists some $B \in \mathcal{B}$ such that $p \in B \subseteq U$. But then $B \cap S \in \mathcal{B}_S$ and $p \in B \cap S \subseteq U \cap S = V$. Therefore, $\mathcal{B}_S$ is a basis of S.
Part 6: by Theorem 4.18, $\mathcal{G}$ generates X if and only if $\mathcal{B} = \mathcal{G}^{FI,M} = \{\bigcap_{i=1}^n G_i : G_i \in \mathcal{G}\} \cup \{\varnothing, X\}$ is a basis on X. Since $\mathcal{B}_S = \{H \cap S : H \in \mathcal{G}^{FI,M}\} = \{(\bigcap_{i=1}^n G_i) \cap S : G_i \in \mathcal{G}\} \cup \{\varnothing, S\} = \{\bigcap_{i=1}^n (G_i \cap S) : G_i \in \mathcal{G}\} \cup \{\varnothing, S\} = \mathcal{G}_S^{FI,M}$, we conclude (by the same theorem mentioned previously) that $\mathcal{G}_S$ generates S. □

Proposition 4.101 (Operator properties)**.** Let $A \subseteq S \subseteq X$. We will subscript the operators with either S or X to distinguish between the topologies.

1. If S is open (in X), $\text{int}_S(A) = \text{int}_X(A)$.
2. $\text{cl}_S(A) = \text{cl}_X(A) \cap S$.

   As a result, A is dense in S (with respect to S) if and only if $S = \text{cl}_X(A)$ (i.e. A is dense in S with respect to X).
3. $\text{ext}_S(A) = \text{ext}_X(A) \cap S$.
4. $\text{lp}_S(A) = \text{lp}_X(A) \cap S$.
5. $\text{ip}_S(A) = \text{ip}_X(A)$.

   As a result, S is a discrete subset of X if and only if S is a discrete space (under subspace topology). In short term, the 2 notions of discreteness are the same, so there will be no confusion when we mention discrete subspace of a space.

Proof. Part 1: if S is open, then $\text{int}_S(A) \subseteq A$ is open in X by Proposition 4.100. Thus, $\text{int}_S(A) \subseteq \text{int}_X(A)$. On the other hand, $\text{int}_X(A) \cap S$ is open in S by definition, so $\text{int}_X(A) \cap S \subseteq \text{int}_S(A)$ by maximality. However $\text{int}_X(A) \subseteq A \subseteq S$, so $\text{int}_X(A) \subseteq \text{int}_S(A)$.
Part 2: the inclusion $\text{cl}_X(A) \cap S \supseteq \text{cl}_S(A)$ holds by minimality since $\text{cl}_X(A)$ is a closed set containing A. For the reverse direction, there exists a closed set C in X such that $\text{cl}_S(A) = C \cap S$. Since $\text{cl}_S(A)$ contains A, C contains A and so C contains $\text{cl}_X(A)$. In other words, $\text{cl}_X(A) \cap S \subseteq C \cap S = \text{cl}_S(A)$.
Part 3: applying part (2), we get $\text{ext}_S(A) = S \setminus \text{cl}_S(A) = S \setminus (\text{cl}_X(A) \cap S) = (S \setminus \text{cl}_X(A)) \cup (S \setminus S) = S \cap (X \setminus \text{cl}_X(A)) = S \cap \text{ext}_X(A)$.
Part 4: suppose p is a limit point of A relative to S. If N is a neighborhood

of p (relative to X), then $N \cap S$ is a neighborhood of p (relative to S) by Proposition 4.100. Therefore, $(N \cap S) \cap A$ contains another point other than p. Since $(N \cap S) \cap A \subseteq N \cap A$, $N \cap A$ contains another point other than p, and so p is a limit point of A relative to X.

Vice versa, suppose $p \in S$ is a limit point of A relative to X. Given V an open neighborhood of p relative to S, there exists an open set U in X such that $V = U \cap S$, since $p \in V$, U is an open neighborhood of p relative to X. By assumption, $U \cap A$ contains another point different from p. Since $A \subseteq S$, $U \cap A = V \cap A$, so as V varies, p should be a limit point of A relative to S.
Part 5:

$$\begin{aligned}\mathrm{ip}_S(A) &= \mathrm{cl}_S(A) \setminus \mathrm{lp}_S(A) = \mathrm{cl}_X(A) \cap S \cap [X \setminus (\mathrm{lp}_X(A) \cap S)] \\ &= \mathrm{cl}_X(A) \cap S \cap [(X \setminus \mathrm{lp}_X(A)) \cup (X \setminus S)] \\ &= [\mathrm{cl}_X(A) \cap S \cap (X \setminus \mathrm{lp}_X(A))] \cup [\mathrm{cl}_X(A) \cap S \cap (X \setminus S)] \\ &= [(\mathrm{cl}_X(A) \setminus \mathrm{lp}_X(A)) \cap S] \cup \varnothing = \mathrm{ip}_X(A) \cap S = \mathrm{ip}_X(A)\end{aligned}$$

The final equality holds since $\mathrm{ip}_X(A) \subseteq A \subseteq S$.

As a result, if S is a discrete subset of X, then $\mathrm{ip}_S(S) = \mathrm{ip}_X(S) = S$. Therefore, with respect to S, every point in S must be open, and thus, S is a discrete space. Vice versa, if S is discrete then every point in S must be open relative to S, so by definition $S \subseteq \mathrm{ip}_S(S) = \mathrm{ip}_X(S) \subseteq S$, i.e. S is a discrete subset of X. □

Proposition 4.102 (Categorical properties)**.**

1. If $A \subseteq S \subseteq X$, then the subspace topology on A relative to S is the same as the subspace topology on A relative X.
2. Characteristic property: the subspace topology on S is unique in that
   (a) The inclusion $\iota_S : S \to X$ is continuous
   (b) $f : Z \to S$ is continuous if and only if $\iota_S \circ f : Z \to X$ is continuous.
3. The subspace topology on S is also the smallest topology such that $\iota_S : S \to X$ is continuous.

Proof.
Part 1: if U is open in A relative to S, then there exists V open in S and W open in X such that $U = V \cap A$ and $V = W \cap S$. Combine these 2 equalities, we get $U = W \cap S \cap A = W \cap A$, so U is open in A relative to X. Vice versa, if U is open in A relative X, there exists W open in X such that $U = W \cap A$. Let $V = W \cap S$, which is open in S relative to X, we have $U = W \cap A = W \cap S \cap = V \cap A$ (since $A \subseteq S$, or $A \cap S = A$), so U is open in A relative to S.

Part 2: we first verify the 2 properties. Since $\iota_S^{-1}(U) = \{x \in S : \iota_S(x) = x \in U\} = S \cap U$ is open in S whenever U is open in X, ι_S is continuous by definition.

If $f : Z \to S$ is continuous, then $\iota_S \circ f : Z \to X$ is continuous (composition of continuous functions is continuous). Vice versa, suppose $\iota_S \circ f$ is continuous. For each V open in S, there exists U open in X such that $V = U \cap S$, so $\iota_S^{-1}(U) = V$, Therefore, $f^{-1}(V) = f^{-1}(U \cap S) = f^{-1}(\iota_S^{-1}(U)) = (\iota_S \circ f)^{-1}(U)$ is open in Z. We conclude f is continuous.

Next, suppose $\mathcal{T}$ is another topology on S that satisfies the 2 conditions. Then $U \cap S = \iota_S^{-1}(U) \in \mathcal{T}$ for all U open in X, so $\mathcal{T}$ is finer than the subspace topology. On the other hand, let $\mathrm{id}_S : S \to (S, \mathcal{T})$ where the former one is endowed with subspace topology. Since we already show that $\iota_S \circ \mathrm{id}_S : S \to X$ is continuous, id_S must be continuous. By Proposition 4.68, the subspace topology must be finer than $\mathcal{T}$. We conclude $\mathcal{T}$ is the subspace topology.
Part 3: this follows from the uniqueness proof of part (2). □

A subset L of a linearly ordered set X is *convex* if $(a, b) \subseteq L$ whenever $a, b \in L$.

Proposition 4.103 (Order vs. Subspace)**.** Let L be a subset of a linearly ordered set X, endowed with order topology

1. Let $\mathcal{T}_L$ and $\mathcal{T}_\leq$ be the subspace topology and order topologies on L (with respect to the topology and order on X), then $\mathcal{T}_L$ is finer $\mathcal{T}_\leq$
2. Equality happens (between the 2 topologies) when L is convex.

Proof. We will compare the topologies based on their corresponding bases (recall Lemma 4.14): the basis for $\mathcal{T}_L$ consists of $(a, b) \cap L$, $(a, \infty) \cap L$, $(-\infty, b) \cap L$ $(a, b \in X)$ (Proposition 4.100), and the basis for $\mathcal{T}_\leq$ consists of $(p, q \in L)$

1. $(p, q)_L = \{x \in L : p < x < q\} = (p, q) \cap L$
2. $(p, \infty)_L = \{x \in L : x > p\} = (p, \infty) \cap L$
3. $(-\infty, q)_L = \{x \in L : x < q\} = (-\infty, q) \cap L$

Part 1: from the above, the basis for $\mathcal{T}_\leq$ is a subcollection of the basis for $\mathcal{T}_L$, so $\mathcal{T}_\leq \subseteq \mathcal{T}_L$.
Part 2: to show the reverse inclusion, we consider the following basic open sets of $\mathcal{T}_L$, and show that there exists $U \in \mathcal{T}_\leq$ around each point p in the sets. Note that every mention of neighborhood is relative to $\mathcal{T}_\leq$.

1. $z \in (a, \infty) \cap L$: divide into the several cases
   (a) $L = \{z\}$: then $L \cap (a, \infty) = \{z\}$ and so L is the neighborhood of z such that $L \subseteq L \cap (a, \infty)$.
   (b) $p < z < q$ for some $p, q \in L$, and $p \leq a \leq q$: then $a \in L$ and $(a, q) \subseteq L$ (i.e. $(a, q)_L = (a, q)$). Hence, (a, q) is an open neighborhood of z, such that $(a, q) \subseteq (a, \infty) \cap L$.
   (c) $p < z < q$ for some $p, q \in L$, and $a < p$: then (p, q) is a neighborhood of z such that $(p, q) \subseteq (a, \infty) \cap L$ (since L is convex).

(d) $p < z < q$ for some $p, q \in L$ and $a > q$: this never happen, since it would mean $a < z < q < a$, a contradiction.

(e) $z = \max L$, and $a \geq p$ for some $p \in L \setminus \{z\}$ (note we already examine the case $L = \{z\}$ in the beginning): then $a \in L$ (as $z > a \geq p$) and $(a, \infty)_L = (a, z] \subseteq (a, \infty)$. Thus, $(a, z]$ is a neighborhood of z such that $(a, z] \subseteq (a, \infty) \cap L$.

(f) $z = \max L$, and $a < p$ for some $p \in L \setminus \{z\}$: then $(p, \infty)_L = (p, z]$ is non-empty since $p \neq z$, and $(p, z] \subseteq (a, \infty)$. Therefore, $(p, z]$ is a neighborhood of z such that $(p, z] \subseteq (a, \infty) \cap L$.

(g) $z = \min L$, and $a \geq q$ for some $q \in L \setminus \{z\}$: this cannot happen since we would get $q \leq a < z \leq q$, a contradiction.

(h) $z = \min L$, and $a < q$ for some $q \in L \setminus \{z\}$: then $(-\infty, q)_L = [z, q)$ is not empty since $q \neq z$, and $[z, q) \subseteq (a, \infty)$. Hence, $[z, q)$ is a neighborhood of z such that $[z, q) \subseteq (a, \infty) \cap L$.

2. $z \in (-\infty, b) \cap L$: similar to $z \in (a, \infty) \cap L$

3. $z \in (a, b) \cap L$: let N^+ be a neighborhood of z such that $N^+ \subseteq (a, \infty) \cap L$ (1st type), and N^- be a neighborhood of z such that $N^- \subseteq (-\infty, b) \cap L$ (2nd type). Denote $N = N^+ \cap N^-$, then we get a neighborhood of z such that $N \subseteq (a, \infty) \cap L \cap (-\infty, b) \cap L = (a, b) \cap L$.

We conclude that $\mathcal{T}_L \subseteq \mathcal{T}_\leq$, or $\mathcal{T}_L = \mathcal{T}_\leq$ simply. □

4.7.1 Continuous function on a subspace

Proposition 4.104 (Restriction)**.** Let $f : X \to Y$ be a continuous function, $S \subseteq X$, $T \supseteq f(X)$

1. The (domain) restriction $f|_S : S \to Y$ is continuous.

2. Vice versa, the codomain restriction $f|^T : X \to T$ is continuous.

Proof. Part 1: $f|_S$ is just $f \circ \iota_S$, so $f|_S$ is continuous because it is the composition of continuous functions.
Part 2: if V is open in T, then there exists U open in Y such that $V = T \cap U$. We have $f^{-1}(U)$ open in X, so $(f|^T)^{-1}(V) = f^{-1}(V) = f^{-1}(T \cap U) = f^{-1}(T) \cap f^{-1}(U) = X \cap f^{-1}(U) = f^{-1}(U)$ is also open in X. By definition, $f|^T$ is continuous. □

Given a continuous map $f : X \to Y$

1. If the restriction $f_X : X \to f(X)$ is a homeomorphism, then we say f is an *topological embedding*. Additionally, if the image $f(X)$ is open (resp. closed), we say that f is an open (resp. closed) embedding.

   Note: by definition, an embedding is always injective.

2. If the image $f(X)$ is dense in Y, we say f is *dominant*.
   Note: a surjective (continuous) map is always dominant, but not vice versa.

3. If, for each point p in X there exists an *open* neighborhood U of p such that $f_U : U \to f(U)$ is a homeomorphism, then we say f is a *local homeomorphism*.

4. If $Y \subseteq X$, and $f \circ \iota_Y = \mathrm{id}_Y$, then we say f is a *retract* of X into Y.
   Note: a retract always surjective (since it is a left inverse of some function).

Proposition 4.105 (Functional properties)**.**

1. If $f : X \to Y$ is a homeomoprhism, and $S \subseteq X$, then the restriction $f_S : S \to f(S)$ is also a homeomorphism.

2. Given $f : X \to Y$ an injective continuous map.

   (a) f is an open map iff f is an open embedding.

   (b) f is a closed map iff f is a closed embedding.

3. Composition of embeddings is an embedding.
   Composition of open embeddings is an open embedding.
   Composition of closed embeddings is a closed embedding.
   Composition of dominant maps is dominant.
   Composition of local homeomorphism is a local homeomorphism.
   Composition of retracts is another retract.

4. The map $\iota_S : S \hookrightarrow X$ is an embedding. It is open iff S is open, and it is closed iff S is closed.

Proof. Part 1: note that f_S and similarly $(f_S)^{-1} = f^{-1}_{f^{-1}(S)}$ is continuous by Proposition 4.104.
Part 2:

(a) If f is open, then $f(X)$ is open, so we only need to show $f_X : X \to f(X)$ is a homeomorphism. Denote $g : f(X) \to X$ to be the inverse of f_X. If U is open in X, then $g^{-1}(U) = f(U)$ is also open by assumption. By definition, g is continuous. Vice versa, if f is an open embedding, then $f(X)$ is open and $f(U)$ is open in $f(X)$ whenever U is open in X (since $f_X : X \to f(X)$ is a homeomorphism). Therefore, $f(U)$ is open in Y because of Proposition 4.100. Since U is arbitrary, f must be open.

(b) Similar to part (2a).

Part 3: given $f : X \to Y$ and $g : Y \to Z$

(a) If f, g are embedding, then $f_X : X \to f(X) \subseteq Y$ and $g_Y : Y \to g(Y) \subseteq Z$ are homeomorphisms. So by part (1), $g_{f(X)} : f(X) \to g(f(X))$ is a homeomorphism, and $(g \circ f)_X : X \to g(f(X)) \subseteq Z$ (which is the composition of f_X and $g_{f(X)}$) is a homeomorphism. By definition, $g \circ f$ is an embedding.

(b) If f, g are open embedding, then $g \circ f$ is an embedding by part (3a). By part (2a), f, g are open maps, so $g \circ f$ is also open by Proposition 4.73. By part (2a), $g \circ f$ is an open embedding.

(c) Similar to part (3b).

(d) If f, g are dominant, then $\mathrm{cl}_Y(f(X)) = Y$, and $\mathrm{cl}_Z(g(Y)) = Z$. By continuity, we have $\mathrm{cl}_Z(g(f(X))) \supseteq g(\mathrm{cl}_Y f(X)) = g(Y)$. Applying the closure on both side we get $\mathrm{cl}_Z(g(f(X))) \supseteq \mathrm{cl}_Z(g(Y)) = Z$. Thus, $g \circ f$ is dominant.

(e) If f, g are local homeomorphism, then for each $p \in X$, there exists open neighborhood U of p and open neighborhood V of $f(p)$ such that $f_U : U \to f(U)$ and $g_V : V \to g(V)$ are homeomorphism. Let $W = f^{-1}(V) \cap U$ be an open neighborhood of p such that $W \subseteq U$ and $f(W) \subseteq V$. The restriction $f_W : W \to f(W)$ (from U to W) of the homeomorphism f_U is still a homeomorphism by part (1), and $g_{f(W)} : f(W) \to g(f(W))$ (from V to $f(W)$) is similarly a homeomorphism. Therefore, $(g \circ f)_W : W \to g(f(W))$ is also a homeomorphism, which concludes $g \circ f$ is a local homeomorphism as p varies in X.

(f) If f, g are retracts, then $f \circ \iota_{YX} = \mathrm{id}_Y$ and $g \circ \iota_{ZY} = \mathrm{id}_Z$ (note that $Z \subseteq Y \subseteq X$, and $\iota_{ZY} : Z \hookrightarrow Y$, $\iota_{YX} : Y \hookrightarrow X$). We then have $\iota_{YX} \circ \iota_{ZY}$ is the inclusion ι_{ZX} of Z into X and

$$(g \circ f) \circ \iota_{ZX} = g \circ (f \circ \iota_{YX}) \circ \iota_{ZY} = g \circ \mathrm{id}_Y \circ \iota_{ZY} = \mathrm{id}_Z$$

Part 4: ι_S is injective and continuous, and given U open in $\iota_S(S) = S$, then $(\iota_S^{-1})^{-1}(U) = \iota_S(U) = U$ is also open in S. By definition, the inverse ι_S^{-1} is continuous, and so ι_S is an embedding.

If S is open (in X), then $\iota_S(S) = S$ is open in X, thus ι_S is an open embedding by definition. If ι_S is an open embedding, then $\iota_S(S)$ is open. The closed case follows similarly □

Proposition 4.106 (Categorical properties)**.**

1. A (continuous) map is a monomorphism iff it is injective.
2. A map is a epimorphism iff it is surjective.
3. A map is an extremal monomorphism iff it is an embedding.
   In fact, all extremal monomorphisms in the **Top** category are also regular (which is stronger).

Proof. Part 1: suppose $f : X \to Y$ is a monomorphism, i.e. for any continuous map $g_1, g_2 : Z \to X$, $f \circ g_1 = f \circ g_2$ implies $g_1 = g_2$. Given $f(x) = f(y)$, let $g_x : \{p\} \to X, p \mapsto x$ and $g_y : \{p\} \to X, p \mapsto y$. These are continuous (Theorem 4.67), and $f \circ g_1(p) = f(x) = f(y) = f \circ g_2(p)$, so $g_x = g_y$, i.e. $x = y$. By

definition, f is injective. The reverse direction is just a basic fact of set theory.
Part 2: suppose $f : X \to Y$ is not surjective, i.e. there exists some $y_0 \in Y$ such that $f(x) \neq y_0$ for all $x \in X$. Endow $Z = \{0, 1\}$ with indiscrete topology, and let $h_1 : Y \to Z$ to be the constant map at 1, $h_f : Y \to Z$ to be the characteristic function of $f(X) \subset Y$ (i.e. $h_f(y) = 1$ if $y \in f(X)$, and 0 otherwise). We then get $h_1(f(x)) = 1 = h_f(f(x))$, yet $h_1 \neq h_f$ since $h_1(y_0) = 1$ yet $h_f(y_0) = 0$. By definition, f is not an epimorphism. The converse, again, is just a basic fact of set theory.
Part 3: since regular monomorphism is extremal, we can show instead every extremal monomorphism is an embedding, and every embedding is a regular monomorphism.

Recall that an extremal monomorphism $f : X \to Y$ is a monomorphism such that whenever $f = g \circ e$ for an epimorphism $e : X \to Z$ (with $g : Z \to Y$), then e is an isomorphism. Let $e = f_X : X \to f(X)$ and $g = \iota_{f(X)} : f(X) \hookrightarrow Y$, then e is surjective. By part (2), e is an epimorphism, so it must be an isomorphism. In other words, f must be an embedding.

Next, if $f : X \to Y$ is an embedding, let $h_1, h_f : Y \to Z$ be defined as in part (2). We then have $h_1(f(x)) = 1 = h_f(f(x))$. In addition, if $e : S \to Y$ is a continuous function such that $h_1 \circ m = h_f \circ m$, then we have $e(s) \in f(X)$ for all $s \in S$ (since $h_f(e(s)) = h_1(e(s)) = 1$). If $u : S \to X$ is a function such that $e = f \circ u$, then $e(s) = f(u(s))$. Since f is injective u is uniquely defined by sending s to $f^{-1}(e(s))$ (it is well-defined since we already show that $e(S) \subseteq f(X)$). If u is continuous, then we have f as the equalizer between h_1 and h_f (in simple term, f is a regular monomorphism). If W is open in X, then

1. $f(W) = O \cap f(X)$ for some open set O in Y.
2. $e^{-1}(f(W)) = e^{-1}(O)$ is open in S.
3. $u^{-1}(W) = \{s \in S : u(s) \in W\} = \{s \in S : f^{-1}(e(s)) \in W\} = \{s \in S : e(s) \in f(W)\} = e^{-1}(f(W))$ which is also open in S by the previous point.

□

4.7.2 Deleted Limit

Let X be a topological space, p be a point in X, and $\mathcal{B}$ be a filter base converging to p. We can then talked about the deleted filter base $\tilde{\mathcal{B}} = \{B \setminus \{p\} :\}$ on the subspace $X \setminus \{p\}$. We will assume from now on that $B \setminus \{p\}$ is non-empty for all $B \in \mathcal{B}$.

Remark 4.107.

1. $\tilde{\mathcal{B}}$ also converges to p
2. If $\mathcal{B} = \mathcal{N}_p$ is the neighborhood filter at p, then $\tilde{\mathcal{N}_p}$ is a non-trivial filter base (in fact, a filter) exactly when p is a limit point of X.

Proof. Part 1: straight from $B \setminus \{p\} \subseteq B$, so $\tilde{\mathcal{B}}$ refines $\mathcal{B}$.
Part 2: $\tilde{\mathcal{N}_p}$ is non-trivial if and only if $N \setminus \{p\} \neq \varnothing$ for all neighborhood N of p. Equivalently, $N \cap X = N$ contains another point different from p. This, by definition, shows that p is a limit point of X. □

Given a function $f : X \setminus \{p\} \to Y$, then a deleted limit at p with respect to $\mathcal{B}$ is an element of $\lim_{\tilde{\mathcal{B}} \to p} f(\tilde{\mathcal{B}})$.

Proposition 4.108. If $f : X \setminus \{p\} \to Y$ is continuous, and $g : X \to Y$ is a *continuous extension* of f to the whole space (i.e. g is continuous and $g(x) = f(x)$ for all $x \neq p$), then $g(p)$ is a deleted limit of f at p (with respect to the neighborhood filter $\mathcal{N}_p$).

Proof. Let N be a neighborhood of $g(p)$, then $g^{-1}(N)$ is a neighborhood of p, or $g^{-1}(N) \setminus \{p\}$ is the deleted neighborhood of p. But $g^{-1}(N) \setminus \{p\} \in \tilde{\mathcal{N}_p}$, and N is arbitrary, so we conclude that $g(p) \in \lim_{\tilde{\mathcal{N}_p} \to p} f(\tilde{\mathcal{N}_p})$. □

Remark 4.109. When the deleted limit is unique, such (continuous) extension will be unique, as long as it exists. We will discuss later whether such extension exists.

4.8 Product topology

To prevent degenerate case, we assume axiom of choice for the product to be non-empty. Given a family $\{X_i\}_{i\in I}$ of topological spaces, there are 2 topologies on the product $\prod_{i\in I} X_i$

1. The product topology, generated by $\pi_i^{-1}(U_i)$ for U_i open in X_i and $i \in I$. Here $\pi_i : \prod_{i\in I} X_i \to X_i$ is the projection onto the i-th coordinate.
2. The box topology, with the basis consists of $\prod_{i\in I} U_i$ where each U_i is open in X_i.

Remark 4.110. The collection described in the box topology is indeed a basis.

Proof. This follows from $\prod_{i\in I} A_i \cap \prod_{i\in I} B_i = \prod_{i\in I} A_i \cap B_i$ (and any topology is closed under finite intersection). □

Proposition 4.111.

1. The box topology is finer the product topology.
2. They are equal if and only if cofinitely many X_i is indiscrete. In particular, this holds when the product is finite.
3. The product of indiscrete spaces is indiscrete in either topologies. The product of discrete spaces is discrete in box topology (but not product topology necessarily if the product is infinite).

Proof. Part 1: since $\pi_i^{-1}(U_i)$ is just product of X_j $(j \neq i)$ and U_i, and X_j is already open, $\pi_i^{-1}(U_i)$ is open in the box topology. By minimality, the product topology is coarser than the box topology.
Part 2: suppose cofinitely many X_i is indiscrete, we denote $J = \{i \in I : X_i \text{ is indiscrete }\}$, and relabel $I \setminus J = \{1, 2, \ldots, n\}$. Thus, each *non-empty* $\prod_{i\in I} U_i$ (U_i is open in X_i) turns into $\prod_{j\in J} X_j \times \prod_{k=1}^n U_k$. This coincides with $\bigcap_{k=1}^n \pi_k^{-1}(U_k)$, and so every basic open set in the box topology is also open in the product topology. By Lemma 4.14, the box topology is coarser than the product topology. Combining part (1), we get the box topology is the same as the product topology.

Vice versa, suppose infinitely many X_i is not indiscrete. We can then pick non-empty U_i open in X_i such that $U_i \neq X_i$ infinitely many $i \in I$. If the box topology is the same as the product topology, there must exists some $V_{i_k} \in X_{i_k}$, $1 \leq k \leq n$, such that $\bigcap_{k=1}^n \pi_{i_k}^{-1}(V_{i_k}) = \prod_{j\neq i_k} X_j \times \prod_{k=1}^n V_{i_k} \subseteq \prod_{i\in I} U_i$ (recall Theorem 4.18). Since $U_i \neq X_i$ for infintiely many $i \in I$, we can pick $j \in I \setminus \{i_1, i_2, \ldots, i_n\}$ such that $U_j \neq X_j$. Yet, looking at index j, we can see X_j is not a subset of U_j, a contradiction.
Part 3: since the only non-empty open set in X_i is X_i, the only non-empty open set in $\prod_{i\in I} X_i$ is the space itself.

If X_i is discrete for each $i \in I$: since $\{p_i\}$ is open in X_i, $\{(p_i)_{i\in I}\}$ is also open. We conclude that $\prod_{i\in I} X_i$, with box topology, is discrete. □

Proposition 4.112 (Universal properties of product topology)**.**

1. By definition, the projection maps are continuous with respect to product topology. Since the box topology is finer than the product topology, the maps are also continuous with respect to box topology

2. The product topology is the unique topology such that

   (a) π_i is continuous for each $i \in I$.

   (b) If $f_i : Y \to X_i$ is continuous for each $i \in I$, then there exists a unique continuous map $f : Y \to \prod_{i\in I} X_i$ such that $f_i = \pi_i \circ f$. We write $f = (f_i)_{i\in I}$.

   In categorical terms, product topology on $\prod_{i\in I} X_i$ is the categorical product of X_i.

3. The product topology is the smallest topology such that π_i is continuous for all $i \in I$.

Proof.
Part 2: we already know π_i is continuous in part (1). Given $f_i : Y \to X_i$, if $f : \prod_{i\in I} X_i \to Y$ is such that $f_i = \pi_i \circ f$ for all $i \in I$, then $f_i(y) = \pi_i(f(y))$. In other words, $f(y) = (f_i(y))_{i\in I}$, so f is uniquely defined. We just need to show f is continuous. Note that $f^{-1}(\pi_i^{-1}(U_i)) = (\pi_i \circ f)^{-1}(U_i) = f_i^{-1}(U_i)$ is continuous for all U_i open in X_i and $i \in I$. By Proposition 4.66, f is continuous.

Now, suppose $\mathcal{T}$ is a topology on $\prod_{i\in I}$ satisfies those conditions. Then $\pi_i^{-1}(U_i) \in \mathcal{T}$ for all U_i open in X_i and $i \in I$. By minimality, $\mathcal{T}$ contains the product topology. On the other hand, let f_i be the projection map from $\prod_{i\in I} X_i$ to X_i (with respect to product topology). They are continuous so the identity map $\mathrm{id}_X : \prod_{i\in I} X_i \to (\prod_{i\in I} X_i, \mathcal{T})$ (note $f_i = \mathrm{id}_X \circ \pi_i$) is also continuous. By Proposition 4.68, the product topology contains $\mathcal{T}$, so they must be equal.
Part 3: follows from the uniqueness proof of part (2). □

Note: from now on, when we mention product space, we will use the product topology on it by default.

Proposition 4.113 (Basic properties)**.**

1. If C_i is closed in X_i, then $\prod_{i\in I} C_i$ is closed in $\prod_{i\in I} X_i$ with respect to both box and product topologies.

2. If $\mathcal{T}_i \subseteq \mathcal{S}_i$ are topologies on X_i, and $\mathcal{T}, \mathcal{S}$ are respectively the product topologies on $\prod_{i\in I}(X_i, \mathcal{T}_i), \prod_{i\in I}(X_i, \mathcal{S}_i)$, then $\mathcal{T} \subseteq \mathcal{S}$. Same for the box topology.

3. If $\mathcal{B}_i$ is a basis of X_i, then $\prod_{i\in I}\mathcal{B}_i$ is a basis for the box topology on $\prod_{i\in I}X_i$.

4. If $\mathcal{A}_i$ generates the topology on X_i for each $i\in I$, then

   (a) $\{\pi_i^{-1}(A_i) : A_i \in \mathcal{A}_i, i \in I\}$ generates the product topology on $\prod_{i\in I}X_i$.

   (b) $\prod_{i\in I}\mathcal{A}_i$ generates the box topology on $\prod_{i\in I}X_i$.

Proof. Part 1: since π_i is continuous (in product topology), $\pi_i^{-1}(C_i)$ is closed in $\prod_{i\in I}X_i$ (with product topology). Intersection of a family of closed sets is still closed, so $\prod_{i\in I}C_i = \bigcap_{i\in I}\pi_i^{-1}(C_i)$ is closed in product topology. It is thus also closed in box topology as the box topology is finer than the product topology.

Part 2: we first show the inclusion for product topologies. A basic open set in $(\prod_{i\in I}X_i, \mathcal{T})$ is of the form $\bigcap_{k=1}^n \pi_{i_k}^{-1}(U_{i_k})$ for some $i_k \in I$ and $U_{i_k}\in\mathcal{T}_{i_k}$. Since $\mathcal{T}_i \subseteq \mathcal{S}_i$ for all $i\in I$, we thus have $\bigcap_{k=1}^n \pi_{i_k}^{-1}(U_{i_k}) \in \mathcal{S}$. By Lemma 4.14, $\mathcal{T}\subseteq\mathcal{S}$.

As for box topologies, it is even more straightforward. A basic open set in $\mathcal{T}$ is of the form $\prod_{i\in I}U_i$ where $U_i\in\mathcal{T}_i$ for each $i\in I$. Since $\mathcal{T}_i\subseteq\mathcal{S}_i$, we get $\prod_{i\in I}U_i\in\mathcal{S}$. By the same Lemma 4.14, we conclude $\mathcal{T}\subseteq\mathcal{S}$.

Part 3: if V is an open set of $\prod_{i\in I}X_i$, and $(x_i)_{i\in I}\in V$, then by definition of box topology, there exists some open set U_i in X_i for each $i\in I$ such that $(x_i)_{i\in I}\in\prod_{i\in I}U_i\subseteq V$. In particular $x_i\in U_i$. Since $\mathcal{B}_i$ is a basis of X_i, there also exists $B_i\in\mathcal{B}_i$ such that $x_i\in B_i\subseteq U_i$. We then have $(x_i)_{i\in I}\in\prod_{i\in I}B_i\subseteq\prod_{i\in I}U_i\subseteq V$. As x and V varies, $\prod_{i\in I}\mathcal{B}_i$ is a basis of the box topology on $\prod_{i\in I}X_i$.

Part 4:

(a) Since any set in $\mathcal{A}_i$ is open in X_i, the inverse of it is open in $\prod_{i\in I}X_i$ (with respect to the product topology). Given any topology $\mathcal{T}$ on $\prod_{i\in I}X_i$ containing every $\pi_i^{-1}(A_i)$ for $A_i\in\mathcal{A}_i$ and $i\in I$, we need to show that it also contains $\pi_i^{-1}(U_i)$ for every open set U_i in X_i, and for all $i\in I$. By Theorem 4.18, for each $U_i\in X_i$, there exists a family $\{\mathcal{F}_\lambda\}_{\lambda\in L}$ of finite families of sets in $\mathcal{A}_i$ such that $U_i = \bigcup_{\lambda\in L}\bigcap\mathcal{F}_\lambda$. Since inverse image preserves union and intersection, we get $\pi_i^{-1}(U_i) = \bigcup_{\lambda\in L}\bigcap_{A\in\mathcal{F}_\lambda}\pi_i^{-1}(A)$. By assumption, $\pi_i^{-1}(A)\in\mathcal{T}$ for each $A\in\mathcal{F}_\lambda$ and $\lambda\in L$. But $\mathcal{T}$ is a topology, so we must have $\bigcap_{A\in\mathcal{F}_\lambda}\pi_i^{-1}(A)\in\mathcal{T}$ (as $\mathcal{F}_\lambda$ is finite) for each $\lambda\in L$, and so $\pi_i^{-1}(U_i)\in\mathcal{T}$. Hence, $\mathcal{T}$ contains the product topology, and by definition, the topology generated by $\{\pi_i^{-1}(A_i) : A_i\in\mathcal{A}_i, i\in I\}$ is the same as the product topology.

□

Proposition 4.114 (Operator properties)**.**

1. Denote $X = \prod_{i\in I} X_i$
   (a) If we use the box topology, then $\operatorname{int}_X \left(\prod_{i\in I} A_i\right) = \prod_{i\in I} \operatorname{int}_{X_i}(A_i)$.
   (b) If we use the product topology, and $J = \{i \in I : A_i \neq X_i\}$, then
      i. $\operatorname{int}_X \left(\prod_{i\in I} A_i\right) = \varnothing$ if J is infinite.
      ii. $\operatorname{int}_X \left(\prod_{i\in I} A_i\right) = \prod_{i\in I\setminus J} X_i \times \prod_{j\in J} \operatorname{int}_{X_j}(A_j)$ if J is finite.
2. $\operatorname{cl}_X \left(\prod_{i\in I} A_i\right) = \prod_{i\in I} \operatorname{cl}_{X_i}(A_i)$ with respect to both box and product topologies.

   As a result, if A_i is dense in X_i for each $i \in I$, then $\prod_{i\in I} A_i$ is dense in $\prod_{i\in I} X_i$.
3. $\operatorname{ext}_X \left(\prod_{i\in I} A_i\right) = \bigcup_{i\in I} \left[\prod_{j\neq i} X_j \times \operatorname{ext}_{X_i}(A_i)\right]$ with respect to both box and product topologies.
4. Limit point
5. Isolated point
6. Density
7. Regular
8. Discrete/Dense-in-itself/Perfect

Proposition 4.115 (Functional properties)**.**

1. Convergence
2. Continuity

Proposition 4.116 (Lexicographical vs. Box)**.** If $\{X_i\}_{i\in I}$ is an well-ordered family of ordered sets, then the order topology on $\prod_{i\in I} X_i$ (with respect to the lexicographical ordering) is finer than the box topology.

Proposition 4.117 (Maximum and minimum are continuous)**.** If X is a linearly ordered set, and $X \times X$ is endowed with product topology, then the maximum function $\max : X \times X \to X$ and the minimum function $\min : X \times X \to X$ are continuous.

4.8.1 The topology of pointwise convergence

Theorem 4.118. Let X be a set, and S be the Sierpiński space. We endow the set of functions S^X with the topology of pointwise convergence. Then given any directed family of subsets $\{A_i\}_{i\in I}$

1. If $A = \lim_{i\in I} A_i$ exists, then 1_A is the limit of $(1_{A_i})_{i\in I}$ in S^X. Here, $1_A : X \to S$ is the indicator function of A: $1_A(x) = 1$ iff $x \in A$.
2. If $f = \lim_{i\in I} 1_{A_i}$ in S^X, then $f = 1_A$ and $A = \lim_{i\in I} A_i$ (where $A = f^{-1}(1)$).

4.9 Initial Topology

4.10 Cauchy spaces

Theorem 4.119. If $f : \omega_\alpha \to M$ is a continuous function between an ordinal (with $\alpha \geq 1$) and a metric space M, then it is eventually constant

4.11 Hypertopology

Topology of pointwise convergence, Fell topology, Vietoris topology, Hausdorff metric, Wijsman convergence

4.12 Separation Axiom

4.12.1 Hausdorff space

Theorem 4.120 (Closed map lemma)**.** If $f : X \to Y$ is a continuous map from a compact space X to a Hausdorff space Y, then f is closed and proper.

Proposition 4.121. The epimorphisms in the category of Hausdorff spaces are exactly the dominant maps (instead of surjections).

4.13 Axiom of Countability

[Reference: Set-Theoretic Topology]

Remark 4.122. If there is only finitely many open neighborhood at p, then the character of X at p is 1.

Proposition 4.123. Let L be a totally ordered space, and X be a topological space such that the cofinality of L is strictly larger than the character of X, then every convergent sequence $(x_i)_{i\in L}$ in X is eventually constant.

Remark 4.124. In a sense, only countably infinite sequences in $\mathbb{R}$ are of considerable interest, otherwise, we get sequences that are eventually constant.

4.14 Connectedness

A function $f : X \to Y$ is *locally constant* if f is

4.15 Convergent series & Integral

4.15.1 Unconditional convergence

Let $(a_i)_{i\in I}$ be a family of points in a topological abelian group G, and define

1. $[I]^{<\omega}$ is the collection of all finite subsets of I.
2. $a_\Sigma : [I]^{<\omega} \to G$ is the summation net (of $(a_i)_{i\in I}$)

$$a_\Sigma(S) = \sum_{i\in S} a_i$$

We say that the series $\sum_{i\in I} a_i$ *converge unconditionally* to b if the net a_Σ converges to b.

4.15.2 Directed convergence

Suppose $(g_\alpha)_{\alpha<\kappa}$ is a transfinite sequence of elements in a Hausdorff topological abelian group G. We then construct another transfinite sequence $g_\Omega : \kappa \to G$, called the sequence of *partial sums*

1. $g_\Omega(0) = g_0$.
2. $g_\Omega(\alpha+1) = g_\Omega(\alpha) + g_{\alpha+1}$.
3. If λ is a limit ordinal, $g_\Omega(\lambda) = \lim_{\alpha<\lambda} g_\Omega(\alpha)$.

We say that the series $\sum_{\alpha<\kappa} g_\alpha$ *converges (conditionally)* to h if $g_\Omega \to h$.

Chapter 5

Mathematical Analysis I

Theorem 5.1 (Silverman-Toeplitz theorem)**.**

Theorem 5.2 (Lebesgue criterion for Riemann integrability)**.** $f : [a, b] \to \mathbb{R}$ is Riemann integrable if the set of discontinuities is of Lebesgue measure 0.

Theorem 5.3 (Lebesgue differentiation theorem)**.** Let f be a locally integrable real or complex-valued function on $\mathbb{R}^n$, then for almost every point $x \in \mathbb{R}^n$, we get

$$\lim_{U \to x} \frac{1}{\lambda(U)} \int_U f d\lambda = f(x)$$

where the limit is taken over the open neighborhoods of x, and λ is the Lebesgue measure on $\mathbb{R}^n$. In fact, since

$$\left| \frac{1}{\lambda(U)} \int_U f d\lambda - f(x) \right| \leq \frac{1}{\lambda(U)} \int_U |f(t) - f(x)| d\lambda(t)$$

A stronger statement holds: that the RHS tends to 0.

Theorem 5.4 (Existence of FHK integral)**.** Let $F : U \subseteq X \to L(X, Y)$ be a continuous function such that U is an open subset of the normed space X, and $L(X, Y)$ is the space of continuous linear map. Let $\gamma : [a, b] \to U$ be a rectifiable curve (i.e. continuous with finite length). Then the Frechet-Henstock-Kurzweil (FHK) integral exists.

Theorem 5.5 (First Fundamental Theorem of Banach Calculus)**.** Suppose $F : U \subseteq X \to L(X, Y)$ is a FHK integrable function such that

1. U is an open subset of a normed space X, and Y is Banach.
2. For any rectifiable loop in U, the FHK integral of F is 0.

Then there exists some continuous map $f : U \to Y$ such that $F = df$.

Theorem 5.6 (Second Fundamental Theorem of Banach Calculus)**.** If $f : U \subseteq X \to Y$ is a differentiable function between an open set U of a normed space X, and a Banach space Y, then

$$\int_\gamma df = f(\gamma(b)) - f(\gamma(a))$$

for any rectifiable curve $\gamma : [a, b] \to U$.

5.0.1 Curl

We assume from now on that V, W are normed spaces, W is Banach, and U is an open subset of V. We also denote $L(V, W)$ to be the space of continuous linear map from X to Y (which is Banach given Y is Banach).

Given $F : U \subseteq V \to L(V, W)$ a continuous function, we say that F is *Frechet-curlable* at $p \in U$ if there exists a multilinear map $T : V \times V \to W$ such that

$$\lim_{h,k\to 0} \frac{\left\| \int_{\gamma_p(h,k)} F - T(h, k) \right\|}{\|h\| \cdot \|k\|} = 0$$

where $\gamma_p(h, k)$ is the loop consisting of 4 line segments: $p \to p + h$, $p + h \to p + h + k$, $p + h + k \to p + k$, and $p + k \to p$. We then denote $K = (\operatorname{curl} F)_p$.

Theorem 5.7. If F is Frechet differentiable at p, then it is curlable at p, and $(\operatorname{curl} F)_p(u, v) = dF_p(u, v) - dF_p(v, u)$. In other words, the curl of F at p is the alternization of the differential of F at p.

Corollary 5.8. The curl is always an alternating map.

Theorem 5.9 (Exterior Derivative)**.** If F is the Frechet derivative of another map $f : U \subseteq V \to W$, then it is curlable on U, and $\operatorname{curl} F = 0$ (or $(\operatorname{curl} F)_p = 0$ for all $p \in U$). In short form, $\operatorname{curl} \circ d = 0$.

Proof. Follows directly from the definition and the fundamental theorem of Banach calculus. □

Theorem 5.10 (Exactness)**.** If the curl of F is 0 at every point in U, then there exists some Frechet differentiable $f : U \to W$ such that $F = df$.

Let $F : U \to L(V, L(V, W))$ be continuous, which can also be interpreted as $F : U \to L^2(V, W)$, where $L^2(V, W)$ is the space of continuous multilinear maps $V \times V \to W$. If $\varphi : [0,1]^2 \to U$ is a surface in U (i.e. a continuous map), we say that F is FHK integrable to φ if there exists some $w_0 \in W$ such that, for all $\varepsilon > 0$, there exists some gauges δ_1 and δ_2 so that

$$\begin{aligned}&\|S(F, \mathcal{P}_1 \times \mathcal{P}_2) - w_0\| \\ &\quad = \left\| \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} F(\varphi(t_i, s_j))(\varphi(x_{i+1}, s_j) - \varphi(x_i, s_j), \varphi(t_i, y_{j+1}) - \varphi(t_i, y_j)) - w_0 \right\| \\ &\quad < \varepsilon\end{aligned}$$

whenever $\mathcal{P}_1 = \{x_0 = 0 < t_0 < x_1 < \cdots < x_{n-1} < t_{n-1} < x_n = 1\}$ is a tagged δ_1-fine partition, and $\mathcal{P}_2 = \{y_0 = 0 < s_0 < y_1 < \cdots < y_{m-1} < s_{m-1} < y_m = 1\}$ is a tagged δ_2-fine partition on $[0,1]$. If so, we denote $w_0 = \iint_\varphi F$.

Theorem 5.11 (Stokes-Banach Theorem)**.** Suppose $F : U \subset V \to L(V, W)$ is curlable on U (i.e. curlable at every point in U), then

$$\iint_\varphi \operatorname{curl} F = \int_{\partial\varphi} F$$

where $\partial\varphi : [0, 4] \to U$ is defined by

$$(\partial\varphi)(t) = \begin{cases} \varphi(0, t) & \text{if } 0 \le t \le 1 \\ \varphi(t - 1, 1) & \text{if } 1 \le t \le 2 \\ \varphi(1, 3 - t) & \text{if } 2 \le t \le 3 \\ \varphi(4 - t, 0) & \text{if } 3 \le t \le 4 \end{cases}$$

Chapter 6

Mathematical Logic II

6.1 Deductive system

Definition 6.1. Let L be a formal language over an alphabet Σ. In logic, every word in L is said to be a *statement* (or *sentence*).

Definition 6.2. If Λ is a *non-empty* subset of L, φ is a statement in L, the pair (Λ, φ) is called a *rule* in L. In this case, statements in Λ are considered to be the *premises*, while φ is considered to be the *conclusion* of the rule.

Note that such rule is an element of $(\mathcal{P}(L) \setminus \{\varnothing\}) \times L$

Definition 6.3. Given a set R of rules, we say a subset Δ of L is *closed under application of rules* in R (or R-*closed*) if, whenever $(\Lambda, \varphi) \in R$ such that $\Lambda \subseteq \Delta$, we must have $\varphi \in \Delta$.

Proposition 6.4.

1. The whole language L and the empty set $\varnothing$ are R-closed.
2. Intersections of R-closed sets is another R-closed sets.

Proof. Part 1: for L, $\varphi \in L$ always holds, so L is R-closed vacuously. For the empty set, note that the premises of any rule is always non-empty, so $\Lambda \subseteq \varnothing$ is always false (whenever $(\Lambda, \varphi) \in R$), and thus, $\varnothing$ is also R-closed vacuously.
Part 2: suppose $\{\Delta_i\}_{i \in I}$ is a family of R-closed sets, let $\Theta = \bigcap_{i \in I} \Delta_i$. Given $(\Lambda, \varphi) \in R$, if $\Lambda \subseteq \Theta$, then $\Lambda \subseteq \Delta_i$ for all $i \in I$. By definition, $\varphi \in \Delta_i$, and so $\varphi \in \Theta$. We conclude that Θ is closed under application of rules in R. □

Definition 6.5. Given $\Gamma \subseteq L$, the set of statements *syntactically entailed* from Γ under R is defined to be the smallest subset of L such that it contains Γ as a subset and it is closed under application of rules in R.

Remark 6.6. Since L is R-closed, such set exists, and it is the intersection of all $\Delta \subseteq L$ such that $\Gamma \subseteq \Delta$ and Δ is R-closed (Proposition 6.4).

If φ is *derivable* from Γ (under R), then we denote $\Gamma \vdash_R \varphi$. If $\Gamma \vdash_R \varphi$ for all $\varphi \in \Delta$, then we succinctly denote $\Gamma \vdash_R \Delta$.

Remark 6.7. $\vdash_R$ is a binary relation on subsets of L, with the convention

1. $\Gamma \vdash_R \varnothing$ for any $\Gamma \subseteq L$.
2. $\varnothing \vdash_R \Gamma$ if and only if Γ is empty.

Regarding the simplification of some notations

1. If $\Gamma \cup \Pi \vdash_R \Delta \cup \Omega$, we can alternatively write $\Gamma, \Pi \vdash_R \Delta, \Omega$ for simplicity.
2. If there are finitely many premises (or conclusions), say $\{\alpha_i : 1 \leq i \leq n\} \vdash_R \{\beta_j : 1 \leq j \leq m\}$, we can instead write

$$\alpha_1, \alpha_2, \ldots, \alpha_n \vdash_R \beta_1, \beta_2, \ldots, \beta_m$$

Proposition 6.8.

1. Monotone (on the left): if $\Gamma \vdash_R \Delta$ and $\Gamma \subseteq \Gamma'$, then $\Gamma' \vdash_R \Delta$.
2. Antitone (on the right): if $\Gamma \vdash_R \Delta$ and $\Delta \supseteq \Delta'$, then $\Gamma \vdash_R \Delta'$.
3. Montone (in the middle): if $\Gamma \vdash_R \Delta$ and $R \subseteq R'$, then $\Gamma \vdash_{R'} \Delta$
4. Reflexive: $\Gamma \vdash_R \Gamma$.
5. Transitive: if $\Gamma \vdash_R \Delta$ and $\Delta \vdash_R \Pi$, then $\Gamma \vdash_R \Pi$.
6. Gentzen cut-elimination: if $\Gamma \vdash_R A, \Delta$ and $\Pi, A \vdash_R \Omega$, then $\Gamma, \Pi \vdash_R \Delta, \Omega$.

Definition 6.9. A *deductive system* $\mathcal{S}$ (over L) is a pair (Ω, R) where

1. Ω is a set of statements in L, called *axioms* of $\mathcal{S}$ (or *laws* in logic).
2. R is a collection of pairs (Λ, φ) (where Λ is a non-empty subset of L and φ is a statement in L), called *inference rules* of $\mathcal{S}$.

Definition 6.10. With respect to $\mathcal{S}$, we define $\Gamma \vdash_{\mathcal{S}} \varphi$ (Γ semantically entails φ relative to $\mathcal{S}$) iff $\Gamma, \Omega \vdash_R \varphi$.

Definition 6.11. A formal system $\mathcal{S}$ is *finitary* if there is only finitely many premises for each inference rule. Otherwise, we say $\mathcal{S}$ is *infinitary*.

6.2 Formal proof

A *proof* (or *derivation*) of a statement φ from Γ is represented by an inverted tree, with each node consists of a single statement, such that

1. The root (at the bottom) is φ.
2. For each node γ
   (a) If it is a leaf (i.e. no children): then γ is either an axiom of $\mathcal{S}$ or is in Γ.
   (b) Otherwise, γ is the conclusion of an inference rule in $\mathcal{S}$, with the premises to be the node's children.

Remark 6.12.

1. Such tree only has finite height since there are only finitely many applications of inference rules before we reach the conclusion φ. That is the basis of inductive definition/construction.
2. Hence, if the system $\mathcal{S}$ is finitary, then a proof tree has finitely many nodes. Using post-order traversal (where immediate children of a node are visited before their parent), such tree can be transformed into a sequence of statements $(\gamma_1, \gamma_2, \ldots, \gamma_n = \varphi)$ where, for each γ_i, either
   (a) γ_i is an axiom of $\mathcal{S}$.
   (b) γ_i belongs to Γ.
   (c) γ_i is the conclusion of some inference rule, with the premises to be those among previous γ_j $(j < i)$.
3. Compactness: as a result, if $\mathcal{S}$ is finitary, and $\Gamma \vdash_{\mathcal{S}} \varphi$, then there exists a finite subcollection Δ of Γ such that $\Delta \vdash_{\mathcal{S}} \varphi$. In fact, this precisely characterizes the finitary property of a deductive system.

Proposition 6.13.

1. If (Γ, φ) is an inference rule in $\mathcal{S}$, then $\Gamma \vdash_{\mathcal{S}} \varphi$. If φ is an axiom of $\mathcal{S}$, then it is a theorem of $\mathcal{S}$.
2. Monotonicity: if Δ contains all the statements in Γ, and Γ (syntactically) entails φ (relative to $\mathcal{S}$), then Δ also entails φ. In particular $\Gamma \vdash_{\mathcal{S}} \varphi$ whenever φ is a theorem of $\mathcal{S}$
3. Preordering:
   (a) $\Gamma \vdash_{\mathcal{S}} \Gamma$
   (b) If $\Gamma \vdash_{\mathcal{S}} \Delta$ and $\Delta \vdash_{\mathcal{S}} \Theta$, then $\Gamma \vdash_{\mathcal{S}} \Theta$.

Proof. Part 1: since φ is the conclusion of Γ through the inference rule (Γ, φ), and $\Gamma \vdash_{\mathcal{S}} \gamma$ for all $\gamma \in \Gamma$ (2nd condition), we get $\Gamma \vdash_{\mathcal{S}} \varphi$ (3rd condition).
Part 2: if $\Gamma \vdash_{\mathcal{S}} \varphi$, then there exists a proof tree of it. The leaves are either axioms of $\mathcal{S}$ or in Γ. Since $\Gamma \subseteq \Delta$, each leaf is also either an axiom, or in Δ. Thus, such tree is also a proof for $\Delta \vdash_{\mathcal{S}} \varphi$.
Part 3:

(a) By definition, $\Gamma \vdash_{\mathcal{S}} \lambda$ for each $\lambda \in \Gamma$, so $\Gamma \vdash_{\mathcal{S}} \Gamma$.

(b) Since $\Delta \vdash_{\mathcal{S}} \theta$, for each $\theta \in \Theta$, there exists a proof tree of θ with each leaf is either an axiom or in Δ. Since $\Gamma \vdash_{\mathcal{S}} \Delta$ and Γ entails every axiom of $\mathcal{S}$, Γ entails each leaf of the proof tree. Recursively, we know that Γ entails θ.

□

A formal system $\mathcal{S}$ is said to be *inconsistent* over L if every statement in L is a theorem of $\mathcal{S}$.

Proposition 6.14. $\mathcal{S}$ is inconsistent if and only if $\Gamma \vdash_{\mathcal{S}} \varphi$ for arbitrary Γ and φ.

Proof. If $\mathcal{S}$ is inconsistent, then φ is a theorem of $\mathcal{S}$, or $\vdash_{\mathcal{S}} \varphi$. By Proposition 6.13, $\Gamma \vdash_{\mathcal{S}} \varphi$. Vice versa, if $\Gamma \vdash_{\mathcal{S}} \varphi$ for arbitrary Γ and φ, then $\vdash_{\mathcal{S}} \varphi$ in particular. By definition, $\mathcal{S}$ is inconsistent. □

A consistent system is said to be *maximal* if adding any *new* statement (i.e. not an axiom) to it as an axiom will render the system inconsistent. Equivalently, $\mathcal{S}$ is maximally consistent if $\varphi \vdash_{\mathcal{S}} \psi$ for any 2 statements φ, ψ in L such that φ is not an axiom of $\mathcal{S}$.

Proposition 6.15. Every theorem in a maximally consistent system is an axiom.

Proof. □

6.3 Interpretation

With respect to a language L, an *interpretation* of L is pair $\mathfrak{I} = (\mathcal{U}, \vDash)$ where

1. $\mathcal{U}$ is a collection of *structures*.
2. $\vDash$ is a relation between structures in $\mathcal{U}$ and statements in L: if $M \vDash \varphi$, then we say M *satisfies* φ.

If $M \vDash \lambda$ for all $\lambda \in \Gamma$, then we can write succinctly $M \vDash \Gamma$. In model theory, we say that M is a *model* for Γ.

If $M \vDash \varphi$ whenever $M \vDash \Gamma$, we say that Γ *semantically entails* φ, and denote $\Gamma \vDash_{\mathfrak{I}} \varphi$.

Remark 6.16. If Γ is empty, then by convention, we say that $\vDash_{\mathfrak{I}} \varphi$ whenever $M \vDash \varphi$ for all structure M. Such φ is called a *tautology* (under $\mathcal{I}$).

Definition 6.17. A collection Γ of statements is *satisfiable* if there exists a model for it, i.e. there exists some structure M such that $M \vDash \Gamma$

Let M be a structures in $\mathcal{U}$, and denote $\mathrm{Thm}_{\mathfrak{I}}(M)$ be the collection of sentences in L that M satisfies.

1. If $\mathrm{Thm}_{\mathfrak{I}}(M) = \mathrm{Thm}_{\mathfrak{I}}(N)$, i.e. M and N satisfies the same sentences, then we say M and N are *elementarily equivalent*.
2. $\mathfrak{I}$ is nontrivial, if $\mathrm{Thm}_{\mathfrak{I}}(M) \neq L$ for each structure M. In simple term, there is always some sentence in L that a structure cannot satisfiable.

Let Γ be a collection of statements in L, and denote $\mathrm{Mod}_{\mathfrak{I}}(\Gamma)$ be the class of structures that satisfy Γ. Note that Γ is satisfiable if and only if $\mathrm{Mod}_{\mathfrak{I}}(\Gamma)$ is non-empty.

1. $\mathfrak{I}$ is Boolean, if
   (a) For each $\alpha \in L$ there exists some $\beta \in L$ such that $M \vDash \alpha$ if and only if $M \nvDash \beta$.
   (b) For each $\alpha, \beta \in L$, there exists some $\gamma, \delta \in L$ such that $\mathrm{Mod}_{\mathfrak{I}}(\gamma) = \mathrm{Mod}_{\mathfrak{I}}(\alpha) \cap \mathrm{Mod}_{\mathfrak{I}}(\beta)$ and $\mathrm{Mod}_{\mathfrak{I}}(\delta) = \mathrm{Mod}_{\mathfrak{I}}(\alpha) \cup \mathrm{Mod}_{\mathfrak{I}}(\beta)$.
2. $\mathfrak{I}$ is compact, if $\mathrm{Mod}_{\mathfrak{I}}(\Gamma) \neq \emptyset$ whenever $\mathrm{Mod}_{\mathfrak{I}}(\Gamma_0) \neq \emptyset$ for each finite subcollection $\Gamma_0 \subseteq \Gamma$.
3. $\mathfrak{I}$ is pointwise-satsifiable, if $\mathrm{Mod}_{\mathfrak{I}}(\varphi)$ is not empty for each statement $\varphi \in L$.

4. Equivalently, $\mathfrak{I}$ is nontrivial if and only if $\mathrm{Mod}_{\mathfrak{I}}(L) = \emptyset$, that is, there does not exists a structure that satisfies every sentence in L.

Remark 6.18. If $\mathfrak{I}$ is pointwise-satsifiable, then $\mathfrak{I}$ is compact if and only if $\mathcal{U}$ is compact, with the topology generated by $\{\mathrm{Mod}_{\mathfrak{I}}(\varphi) : \varphi \in L\}$ as closed sets.

6.4 Soundness - Completeness

A formal system $\mathcal{S}$ together with a formal interpretation $\mathfrak{I}$ over the same language L is then called a *formal logic* over L.

A formal logic $(\mathcal{S}, \mathfrak{I})$ is called

1. *Coherent*, if
   (a) Every structure of the interpretation satisfies every law of the deductive system.
   (b) If (Γ, φ) is an inference rule, then a structure M satisfies φ whenever it satisfies Γ.
2. *Weakly sound*, if $\vDash_{\mathfrak{I}} \varphi$ whenever $\vdash_{\mathcal{S}} \varphi$.
3. *Sound* (or *strongly sound*), if $\Gamma \vDash_{\mathfrak{I}} \varphi$ whenever $\Gamma \vdash_{\mathcal{S}} \varphi$.
4. *Weakly complete*, if $\vdash_{\mathcal{S}} \varphi$ whenever $\vDash_{\mathfrak{I}} \varphi$.
5. *Complete* (or *strongly complete*), if $\Gamma \vdash_{\mathcal{S}} \varphi$ whenever $\Gamma \vDash_{\mathfrak{I}} \varphi$.
6. *T-complete*, if $\Gamma \vdash_{\mathcal{S}} L$ whenever $\Gamma \vDash_{\mathfrak{I}} L$.

Proposition 6.19.

1. Coherence implies soundness, which in turn implies weak soundness.
2. Completeness implies weak completeness and T-completeness

Proposition 6.20.

1. It (the logic) is sound if and only if whenever Γ is satisfiable, then it is consistent.
2. It is complete if and only if whenever Γ is satisfiable, then it is consistent.
3. If it is complete, and the deductive system is finitary, then the interpretation has compactness property.

6.4.1 Conjunctive Normal Form - Towards the full Completeness Theorem

Corollary 6.21. If Γ is a collection of formulas with finitely many propositional variables involved, and $\Gamma \vDash \varphi$, then $\Gamma \vdash \varphi$.

Sketch. Use truth table to selectively pick out rows where some formula in Γ is false (if there is none, then φ is actually a tautology, and so the theorem reduces to the first completeness theorem). Note that there are only finitely many of them since there are finitely many propositional variable involved in Γ. From there, we get a proof of the formula (in conjunctive normal form) where Γ is false. Use $\varphi \vee \psi, \neg\psi \vdash \varphi$ to show that Γ proves the disjunction of rows where every formula in Γ holds. A combination of $(\varphi \to \xi) \wedge (\psi \to \xi) \equiv (\varphi \vee \psi) \to \xi$ and $\overline{p_1}, \overline{p_2}, \ldots, \overline{p_n} \vdash \overline{\varphi}$ is then suffice to get a proof of φ. □

Lemma 6.22.

$$\neg(\psi \wedge \lambda) \vdash \neg\psi$$

Corollary 6.23. Assuming the weak Konig's lemma, if $\Gamma \vDash \varphi$, then $\Gamma \vdash \varphi$.

6.5 Translation

Definition 6.24. Let $f : L \to K$ be a translation map between languages (i.e. a function). The translation of a statement φ in L through f is denote as $\varphi^f := f(\varphi)$.

Let $\mathcal{S}$ be a deductive system over L, and $\mathcal{T}$ be a deductive system over K.

Definition 6.25. Relative to the translator f

1. $\mathcal{S}$ is *interpretable* in $\mathcal{T}$ if

$$\Gamma^f \vdash_{\mathcal{T}} \varphi^f \text{ whenever } \Gamma \vdash_{\mathcal{S}} \varphi.$$

   In particular, f translates every theorem of $\mathcal{S}$ into a theorem of $\mathcal{T}$.

2. $\mathcal{S}$ is *co-interpretable* in $\mathcal{T}$ if

$$\Gamma \vdash_{\mathcal{S}} \varphi \text{ whenever } \Gamma^f \vdash_{\mathcal{T}} \varphi^f.$$

   In particular, if the translation of a statement of L is a theorem of $\mathcal{T}$, then such statement is a theorem of $\mathcal{S}$.

3. $\mathcal{S}$ is *weakly interpretable* in $\mathcal{T}$ if

$$\Gamma^f \cup \{\varphi^f\} \text{ is consistent relative to } \mathcal{T} \text{ whenever } \Gamma \vdash_{\mathcal{S}} \varphi$$

   In particular, the translation of every theorem of $\mathcal{S}$ is consistent with $\mathcal{T}$.

Given a formal system $\mathcal{S}$, the (deductive) *completion* $\mathcal{S}^{\vdash}$ of $\mathcal{S}$ is another formal system such that

1. Every theorem of $\mathcal{S}$ is an axiom of $\mathcal{S}^{\vdash}$, and vice versa.
2. Whenever $\Gamma \vdash_{\mathcal{S}} \varphi$, then (Γ, φ) is an inference rule of $\mathcal{S}^{\vdash}$, and vice versa.

Remark 6.26. φ is a theorem of $\mathcal{S}$ if and only if $(\varnothing, \varphi)$ is an inference rule of $\mathcal{S}^{\vdash}$ (here, $\varnothing$ denotes the empty collection).

Given 2 formal systems $\mathcal{S}_1$ and $\mathcal{S}_2$ over the same language, we say

1. $\mathcal{S}_1$ is a *subsystem* of $\mathcal{S}_2$ if
   (a) Every axiom of $\mathcal{S}_1$ is an axiom of $\mathcal{S}_2$.
   (b) Every inference rule of $\mathcal{S}_1$ is an inference rule of $\mathcal{S}_2$.
2. $\mathcal{S}_1 = \mathcal{S}_2$ if they are subsystem of each other.

3. $\mathcal{S}_1$ is *deductively weaker* than $\mathcal{S}_2$ (or $\mathcal{S}_2$ is deductively stronger than $\mathcal{S}_1$) if $\mathcal{S}_1^{\vdash}$ is a subsystem of $\mathcal{S}_2^{\vdash}$.
4. $\mathcal{S}_1$ and $\mathcal{S}_2$ are *deductively equivalent* if they are deductively weaker than each other. In other words, 2 systems are deductively equivalent if their completion are the same.

Proposition 6.27.

1. Monotonicity: if $\mathcal{S}_1$ is a subsystem of $\mathcal{S}_2$, then $\mathcal{S}_1^{\vdash}$ is a subsystem of $\mathcal{S}_2^{\vdash}$. Stated otherwise, $\mathcal{S}_1$ is deductively weaker than $\mathcal{S}_2$ whenever $\mathcal{S}_1$ is a subsystem of $\mathcal{S}_2$.
2. Idempotency: $(\mathcal{S}^{\vdash})^{\vdash}$ is the same as $\mathcal{S}^{\vdash}$.
3. Conservativity:
   (a) If $\mathcal{S} \cup \Gamma$ denotes the formal system where inference rules stays the same, but we add all statements in Γ as axioms, then the inference rules in $\mathcal{S}^{\vdash}$ are the same as those in $(\mathcal{S} \cup \Gamma)^{\vdash}$.
   (b) The completion of $\mathcal{S}$ is (deductively) equivalent to $\mathcal{S}$.

Proof. Part 1: we first show that if $\Gamma \vdash_{\mathcal{S}_1} \varphi$, then $\Gamma \vdash_{\mathcal{S}_2} \varphi$. Indeed,

1. If φ is an axiom of $\mathcal{S}_1$: then it is also an axiom of $\mathcal{S}_2$, and so $\Gamma \vdash_{\mathcal{S}_2} \varphi$.
2. If φ belongs to Γ: then $\Gamma \vdash_{\mathcal{S}_2} \varphi$ vacuously.
3. If (Λ, φ) is an inference rule of $\mathcal{S}_1$ such that $\Gamma \vdash_{\mathcal{S}_1} \lambda$ for each $\lambda \in \Lambda$. By induction hypothesis, $\Gamma \vdash_{\mathcal{S}_2} \lambda$. Finally, note that $\mathcal{S}_2$ contains all inference rules of $\mathcal{S}_1$, so we must have $\Gamma \vdash_{\mathcal{S}_2} \varphi$ by definition.

By induction principle, $\Gamma \vdash_{\mathcal{S}_2} \varphi$ whenever $\Gamma \vdash_{\mathcal{S}_1} \varphi$. By construction, every inference rule of $\mathcal{S}_1^{\vdash}$ is an inference rule of $\mathcal{S}_2^{\vdash}$. On the other hand, if we let Γ to be empty, then every theorem of $\mathcal{S}_1$ is a theorem of $\mathcal{S}_2$ (Remark 6.26). In other words, every axiom of $\mathcal{S}_1^{\vdash}$ is an axiom of $\mathcal{S}_2^{\vdash}$. We conclude that $\mathcal{S}_1^{\vdash}$ is a subsystem of $\mathcal{S}_2^{\vdash}$.
Part 2: for axioms

1. If φ is an axiom of $\mathcal{S}^{\vdash}$, then it is a theorem of $\mathcal{S}^{\vdash}$, and so an axiom of $(\mathcal{S}^{\vdash})^{\vdash}$.
2. If φ is an axiom of $(\mathcal{S}^{\vdash})^{\vdash}$, then it is a theorem of $\mathcal{S}^{\vdash}$, so every leaf of a proof tree of φ represent a theorem of $\mathcal{S}$. By Remark **??**, φ is then a theorem of $\mathcal{S}$, and so an axiom of $\mathcal{S}^{\vdash}$.

For inference rules

1. If (Γ, φ) is an inference rules of $\mathcal{S}^{\vdash}$, then $\Gamma \vdash_{\mathcal{S}^{\vdash}} \varphi$ by Proposition 6.13. Thus, (Γ, φ) is an inference rule of $(\mathcal{S}^{\vdash})^{\vdash}$.

2. If (Γ, φ) is an inference rule of $(\mathcal{S}^{\vdash})^{\vdash}$, then $\Gamma \vdash_{\mathcal{S}^{\vdash}} \varphi$ by definition of completion. Thus, there exists a proof tree of φ with the leaves represent either axioms of $\mathcal{S}^{\vdash}$ or statements in Γ. Since every axiom of $\mathcal{S}^{\vdash}$ is a theorem of $\mathcal{S}$, we can construct another proof tree of φ such that the leaves represent either axioms of $\mathcal{S}$ or statements in Γ (Remark **??**). Thus, $\Gamma \vdash_{\mathcal{S}} \varphi$, and so (Γ, φ) is an inference rule of $\mathcal{S}^{\vdash}$.

Part 3: just another phrasing of part (2). □

Proposition 6.28 (Local property)**.** $\mathcal{S}_1$ is deductively weaker than $\mathcal{S}_2$ if and only if

1. Every axiom of $\mathcal{S}_1$ is a theorem of $\mathcal{S}_2$.
2. $\Gamma \vdash_{\mathcal{S}_2} \varphi$ whenever (Γ, φ) is an inference rule of $\mathcal{S}_1$.

Proof. If $\mathcal{S}_1$ is deductively weaker than $\mathcal{S}_2$, then

1. Every theorem of $\mathcal{S}_1$ is a theorem of $\mathcal{S}_2$. In particular, every axiom of $\mathcal{S}_1$ is a theorem of $\mathcal{S}_2$.
2. $\Gamma \vdash_{\mathcal{S}_2} \varphi$ whenever $\Gamma \vdash_{\mathcal{S}_1} \varphi$. In particular, from Proposition 6.13, we know that $\Gamma \vdash_{\mathcal{S}_2} \varphi$ whenever (Γ, φ) is an inference rule of $\mathcal{S}_1$.

Vice versa, suppose every axiom of $\mathcal{S}_1$ is a theorem of $\mathcal{S}_2$. If $\Gamma \vdash_{\mathcal{S}_1} \varphi$, and $\Gamma \vdash_{\mathcal{S}_2} \varphi$ for each inference rule (Γ, φ) of $\mathcal{S}_1$. Inductively, for each $\Gamma \vdash_{\mathcal{S}_1} \varphi$

1. If φ belongs to Γ: then $\Gamma \vdash_{\mathcal{S}_2} \varphi$ trivially.
2. If φ is an axiom of $\mathcal{S}_1$: then it is a theorem of $\mathcal{S}_2$, or $\vdash_{\mathcal{S}_2} \varphi$ simply. By Proposition 6.13, $\Gamma \vdash_{\mathcal{S}_2} \varphi$.
3. If (Λ, φ) is an inference rule of $\mathcal{S}_1$ such that Γ entails each statement in Λ relative to $\mathcal{S}_1$:
   (a) By induction hypothesis, Γ entails each statement in Λ relative to $\mathcal{S}_2$.
   (b) Since (Λ, φ) is an inference rule of $\mathcal{S}_1$, Λ entails φ relative to $\mathcal{S}_2$.

   From these 2 points, we conclude Γ entails φ relative to $\mathcal{S}_2$ by definition and by.

Hence, every inference rule of $\mathcal{S}_1^{\vdash}$ is an inference rule of $\mathcal{S}_2^{\vdash}$. Also, if φ is an axiom of $\mathcal{S}_1^{\vdash}$, then it is a theorem of $\mathcal{S}_1$. Thus, $(\varnothing, \varphi)$ is an inference rule of $\mathcal{S}_1^{\vdash}$ (Remark 6.26), so by the previous observation, it is also an inference rule of $\mathcal{S}_2$. In other words, φ is a theorem of $\mathcal{S}_2$, or equivalently, φ is an axiom of $\mathcal{S}_2^{\vdash}$. We conclude that $\mathcal{S}_1^{\vdash}$ is a subsystem of $\mathcal{S}_2^{\vdash}$, or equivalently, $\mathcal{S}_1$ is deductively weaker than $\mathcal{S}_2$. □

6.6 Semantic Categoricity for Propositional Logic

Let F_0 be the set of all wffs of propositional logic, or simply, the language of propositional logic. A valuation on F_0 is defined to be a function $\mathfrak{v} : F_0 \to \{0, 1\}$. We say a valuation is

1. *Implicational* if
   (a) $\mathfrak{v}(\alpha) = 1$ for every axiom α.
   (b) If $\mathfrak{v}(\varphi) = \mathfrak{v}(\varphi \to \psi) = 1$, then $\mathfrak{v}(\psi) = 1$
2. *Negational* if $\mathfrak{v}(\varphi) \neq \mathfrak{v}(\neg\varphi)$ for each wff φ.
3. *Propositional* if it is both implicational and negational.
4. *Coherent* (with respect to propositional logic) if
   (a) $\mathfrak{v}(\alpha \to \beta) = \mathfrak{v}(\varphi \to \psi)$ whenever $\mathfrak{v}(\alpha) = \mathfrak{v}(\varphi)$ and $\mathfrak{v}(\beta) = \mathfrak{v}(\psi)$.
   (b) $\mathfrak{v}(\neg\varphi) = \mathfrak{v}(\neg\psi)$ whenever $\mathfrak{v}(\varphi) = \mathfrak{v}(\psi)$

Remark 6.29. An implicational valuation will always assign 1 to every theorem. Additionally, if $\varphi \equiv \psi$, then $\mathfrak{v}(\varphi) = \mathfrak{v}(\psi)$

Proof. The first statement follows by induction on proof's length. As for second statement, by definition, $\varphi \to \psi$ and $\psi \to \varphi$ are theorems. Hence, suppose, WLOG, that $\mathfrak{v}(\varphi) = 1$ and $\mathfrak{v}(\psi) = 0$ (since there are only 2 values that can be assigned). Since $\mathfrak{v}(\varphi) = \mathfrak{v}(\varphi \to \psi) = 1$, we get $\mathfrak{v}(\psi) = 1$, a contradiction. □

Proposition 6.30. If $\mathfrak{v}$ is negational, then

1. $\mathfrak{v}(\varphi) = \mathfrak{v}(\neg\neg\varphi)$ for each wff φ.
2. $\mathfrak{v}(\neg\varphi) = \mathfrak{v}(\neg\psi)$ whenever $\mathfrak{v}(\varphi) = \mathfrak{v}(\psi)$.

Proof. This follows from the fact that there are only 2 possible values to assign each formula to. □

Lemma 6.31. Suppose $\mathfrak{v}$ is propositional, then

1. If $\mathfrak{v}(\varphi) = 1$, then $\mathfrak{v}(\varphi \to \gamma) = \mathfrak{v}(\gamma)$ and $\mathfrak{v}(\alpha \to \varphi) = 1$.
2. If $\mathfrak{v}(\varphi) = 0$, then $\mathfrak{v}(\alpha \to \varphi) = \mathfrak{v}(\neg\alpha)$ and $\mathfrak{v}(\varphi \to \gamma) = 1$.

Proof.
Part 1: as $\mathfrak{v}$ is implicational, we get $\mathfrak{v}(\varphi \to (\alpha \to \varphi)) = 1$, and so $\mathfrak{v}(\alpha \to \varphi) = 1$. Now, suppose $\mathfrak{v}(\varphi \to \gamma) \neq \mathfrak{v}(\gamma)$, we consider the following cases

1. If $\mathfrak{v}(\varphi \to \gamma) = 1$ and $\mathfrak{v}(\gamma) = 0$: since $\mathfrak{v}(\varphi) = \mathfrak{v}(\varphi \to \gamma) = 1$, we get $\mathfrak{v}(\gamma) = 1$, a contradiction.

2. If $\mathfrak{v}(\varphi \to \gamma) = 0$ and $\mathfrak{v}(\gamma) = 1$: since $\mathfrak{v}(\gamma) = \mathfrak{v}(\gamma \to (\varphi \to \gamma)) = 1$, we get $\mathfrak{v}(\varphi \to \gamma) = 1$, which again is a contradiction.

Part 2: as $\mathfrak{v}$ is negational, we get $\mathfrak{v}(\neg\varphi) = 1$. Apply part 1, we get $\mathfrak{v}(\neg\gamma \to \neg\varphi) = 1$ and $\mathfrak{v}(\neg\varphi \to \neg\alpha) = \mathfrak{v}(\neg\alpha)$. With NP1, we obtain $\mathfrak{v}(\alpha \to \varphi) = \mathfrak{v}(\neg\alpha)$ and $\mathfrak{v}(\varphi \to \gamma) = 1$ □

Note: part 1 only use implicational property.

Theorem 6.32. A propositional valuation is always coherent.

Proof.
*Note that the 2nd condition is a direct result from negational property, so we only need to prove the 1st one. However, we first need to show that, whenever $\mathfrak{v}(\alpha) = \mathfrak{v}(\beta)$, we have

1. $\mathfrak{v}(\alpha \to \gamma) = \mathfrak{v}(\beta \to \gamma)$
2. $\mathfrak{v}(\lambda \to \alpha) = \mathfrak{v}(\lambda \to \beta)$

Consider the following 2 cases

1. If $\mathfrak{v}(\alpha) = \mathfrak{v}(\beta) = 1$: then $\mathfrak{v}(\alpha \to \gamma) = \mathfrak{v}(\gamma) = \mathfrak{v}(\beta \to \gamma)$, and $\mathfrak{v}(\lambda \to \alpha) = 1 = \mathfrak{v}(\lambda \to \beta)$.
2. If $\mathfrak{v}(\alpha) = \mathfrak{v}(\beta) = 0$: then $\mathfrak{v}(\alpha \to \gamma) = \mathfrak{v}(\gamma) = 1 = \mathfrak{v}(\beta \to \gamma)$, and $\mathfrak{v}(\lambda \to \alpha) = \mathfrak{v}(\neg\lambda) = \mathfrak{v}(\lambda \to \beta)$.

*Hence, assuming $\mathfrak{v}(\alpha) = \mathfrak{v}(\varphi)$ and $\mathfrak{v}(\beta) = \mathfrak{v}(\psi)$, we then have

$$\mathfrak{v}(\alpha \to \beta) = \mathfrak{v}(\alpha \to \psi) = \mathfrak{v}(\varphi \to \psi)$$

□

A valuation $\mathfrak{v}$ is *trivial* if it assigns 1 to every wffs.

Remark 6.33. The trivial valuation is implicational and coherent, but not negational.

Theorem 6.34. If $\mathfrak{v}$ is implicational, coherent and non-trivial, then

1. $\mathfrak{v}(\varphi \to \psi) = 0$ if and only if $\mathfrak{v}(\varphi) = 1$ and $\mathfrak{v}(\psi) = 0$.
2. $\mathfrak{v}(\neg\varphi) = 0$ if and only if $\mathfrak{v}(\varphi) = 1$.

Proof.
Part 1: follows from compatibility by induction on the length of some deduction sequence of a theorem.
Part 2: Since it is coherent, $\mathfrak{v}(\varphi \to \psi)$ is a function depending only on the value

of $\mathfrak{v}(\varphi)$ and $\mathfrak{v}(\psi)$. In other words, there exists a function $Q : \{0,1\}^2 \to \{0,1\}$ such that $\mathfrak{v}(\varphi \to \psi) = Q(\mathfrak{v}(\varphi), \mathfrak{v}(\psi))$.

By the hypothesis, let p be a wff such that $\mathfrak{v}(p) = 0$, and q be a theorem (i.e. $\mathfrak{v}(q) = 1$). Suppose $Q(1,0) = 1$, we obtain

1. $\mathfrak{v}(p \to (q \to p)) = 1 = Q(0, Q(1,0)) = Q(0,1)$.
2. $\mathfrak{v}(q \to (p \to q)) = 1 = Q(1, Q(0,1)) = Q(1,1)$ (by the previous points).
3. $Q(0,0)$ cannot be 0 since otherwise, $\mathfrak{v}(p \to (p \to p)) = 1 = Q(0, Q(0,0)) = Q(0,0) = 0$, a contradiction.

In other words, $Q(x,y) = 1$ for all $x, y \in \{0,1\}$. This means that $\mathfrak{v}$ always assign 1 to every $\varphi \to \psi$, contradicting the nontriviality.

Therefore, $Q(1,0) = 0$. From there

1. $Q(0,0) = Q(0, Q(1,0)) = \mathfrak{v}(p \to (q \to p)) = 1$.
2. $Q(0,1) = Q(0, Q(0,0)) = \mathfrak{v}(p \to (p \to p)) = 1$.
3. $Q(1,1) = Q(1, Q(0,1)) = \mathfrak{v}(q \to (p \to q)) = 1$

Part 3: since $\mathfrak{v}(p) \neq \mathfrak{v}(q)$, we must have $\mathfrak{v}(\neg p) \neq \mathfrak{v}(\neg q)$, otherwise $\mathfrak{v}(\neg\neg p) = \mathfrak{v}(\neg\neg q) = b$ by coherence. Then, by DN (and part 1), we must have $Q(0,b) = 1 = Q(1,b)$ and $Q(b,0) = 1 = Q(b,1)$. Checking case-by-case on the possible values of b, we conclude impossibility.
Suppose $\mathfrak{v}(\neg p) = 0 = \mathfrak{v}(p)$, and $\mathfrak{v}(\neg q) = 1 = \mathfrak{v}(q)$, then

$$\begin{aligned}\mathfrak{v}((\neg p \to \neg q) \to (q \to p)) &= 1 \\ &= Q(\mathfrak{v}(\neg p \to \neg q), \mathfrak{v}(q \to p)) \\ &= Q(Q(0,1), Q(1,0)) = Q(1,0) = 0\end{aligned}$$

A contradiction! Thus, it must be the case that $\mathfrak{v}(\neg p) = 1$ and $\mathfrak{v}(\neg q) = 0$. By coherence, we conclude in general that $\mathfrak{v}(\neg\varphi) = 0$ if and only if $\mathfrak{v}(\varphi) = 1$. □

Corollary 6.35. Asssuming $\mathfrak{v}$ is implicational and coherent, then $\mathfrak{v}$ is non-trivial if and only if it is negational.

6.7 Pretext for Axiomatization of Predicate Calculus

6.7.1 Henkin's Model Existence Theorem

Theorem 6.36. With respect to first-order logic, let σ be some signature, and Γ be a set of sentences with respect to σ such that

1. It is complete: for each sentence φ, either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.
2. It is consistent: for each sentence φ, it is impossible that $\varphi \in \Gamma$ and $\neg\varphi \in \Gamma$ simultaneously.
3. It is saturated: for each sentence of the form $\exists w_1 \exists w_2 \dots \exists w_n(\varphi(w_1, w_2, \dots, w_n))$, there exists constants $c_1, c_2, \dots, c_n \in \sigma$ such that

$$[\exists w_1 \exists w_2 \dots \exists w_n(\varphi(w_1, w_2, \dots, w_n)) \to \varphi(c_1, c_2, \dots, c_n)] \in \Gamma$$

   We say that $\vec{c}$ is the Henkin's witness of $\exists \vec{w}(\varphi(\vec{w}))$.

Then Γ is satisfiable.

Proof. Let N be the collection of terms (e.g. constants, and functions with constant arguments). Define the equivalence relation $\sim_\Gamma$ on N where

$$t \sim_\Gamma s \text{ iff } (t = s) \in \Gamma$$

In other words, t and s are the same if it is provable that $t = s$. This is an equivalence relation, inductively from equality and Leibniz's substitution of formulas. Denote $\mathcal{M} = \mathcal{N}/\sim_\Gamma$ (note that completeness plays a role in this as derivability is different from hypothesis). We will prove that $\mathcal{M} \vDash \Gamma$.

For each constant, function, and relation symbol in σ, define

1. If c is a constant symbol, let $c^{\mathcal{M}} = [c]_\Gamma$ (this includes the Henkin's witnesses).
2. If R is an n-ary relation, let $R^{\mathcal{M}} \subseteq \mathcal{M}^n$ be such that $([t_1]_\Gamma, \dots, [t_n]_\Gamma) \in R^{\mathcal{M}}$ iff $R(t_1, \dots, t_n) \in \Gamma$.
3. If f is an n-ary function, let $f^{\mathcal{M}} : \mathcal{M}^n \to \mathcal{M}$ be such that $f^{\mathcal{M}}([t_1]_\Gamma, \dots, [t_n]_\Gamma) = [f(t_1, t_2, \dots, t_n)]_\Gamma$.

These are well-defined: suppose $t_i \sim_\Gamma s_i$, or $t_i = s_i \in \Gamma$. By Leibniz's principle of substitution, this means

1. For an n-ary relational symbol R: $\Gamma \vdash R(t_1, \ldots, t_n) \leftrightarrow R(s_1, \ldots, s_n)$. Since Γ is consistent and complete, either both $R(t_1, \ldots, t_n), R(s_1, \ldots, s_n) \in \Gamma$, or $R(t_1, \ldots, t_n), R(s_1, \ldots, s_n) \notin \Gamma$.

$$\begin{aligned}([t_1]_\Gamma, \ldots, [t_n]_\Gamma) \in R^{\mathcal{M}} &\text{ iff } R(t_1, \ldots, t_n) \in \Gamma \\ &\text{iff } R(s_1, \ldots, s_n) \in \Gamma \\ &\text{iff } ([s_1]_\Gamma, \ldots, [s_n]_\Gamma) \in R^{\mathcal{M}}\end{aligned}$$

2. For an n-ary functional symbol f: $\Gamma \vdash f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n)$. Again, as Γ is consistent and complete, so $f(t_1, \ldots, t_n) = f(s_1, \ldots, s_n) \in \Gamma$. In other words, $[f(t_1, \ldots, t_n)]_\Gamma = [f(s_1, \ldots, s_n)]_\Gamma$.

Now, we will do *structural* induction to show that $\mathcal{M} \vDash \Gamma$, i.e. $\mathcal{M} \vDash \varphi$ whenever $\varphi \in \Gamma$.

1. If $\varphi \equiv (t = s)$ (where t, s are terms): then $t \sim_\Gamma s$, and so $[t]_\Gamma = [s]_\Gamma$, i.e. $\mathcal{M} \vDash (t = s)$ (since any interpretation leaves the terms the same, and sends t to $[t]_\Gamma$).
2. If $\varphi \equiv R(t_1, \ldots, t_n)$ (where t_i are terms, *not* variables): by construction, $([t_1]_\Gamma, \ldots, [t_n]_\Gamma) \in R^{\mathcal{M}}$. Therefore, we have $\mathcal{M} \vDash R(t_1, \ldots, t_n)$.
3. If $\varphi \equiv \neg\psi$: since Γ is consistent, $\psi \notin \Gamma$. Under induction hypothesis, $\mathcal{M} \nvDash \psi$. Since satisfaction is complete (in model-theoretical sense), we must have $\mathcal{M} \vDash \neg\psi$.
4. If $\varphi \equiv \psi \wedge \xi$: as $\varphi \in \Gamma$, both $\psi, \xi \in \Gamma$. By induction, $\mathcal{M} \vDash \psi$ and $\mathcal{M} \vDash \xi$. By definition of satisfaction, $\mathcal{M} \vDash \psi \wedge \xi$.
5. If $\varphi \equiv \psi \vee \xi$ or $\varphi \equiv \psi \to \xi$: it follows from the previous 2 points.
6. If $\varphi \equiv \exists x(\psi(x))$: recall Γ is saturated, meaning there exists a term t (possibly some constant symbol) such that $[\exists x(\psi(x)) \to \psi(t)] \in \Gamma$. Combining this with the hypothesis $\varphi \in \Gamma$, we know that $\psi(t) \in \Gamma$ (deduction theorem). By induction, $\mathcal{M} \vDash \psi(t)$.

   Hence, we can construct a variable interpretation function $\mathfrak{I}_\psi : \mathbf{Var} \to \mathcal{M}$ by sending every variable to $[t]_\Gamma$. This should make $\mathcal{M}, \mathfrak{I}_\psi \vDash \psi(x)$. By definition, $\mathcal{M} \vDash \exists x(\psi(x))$.
7. If $\varphi \equiv \forall x(\psi(x))$: first, there exists a term t such that $[\exists x\,(\neg\psi(x)) \to \neg\psi(t)] \in \Gamma$. Since $\alpha \to \beta$ is logically/provably equivalent to $\neg\beta \to \neg\alpha$,

$$\begin{aligned}&[\neg\neg\psi(t) \to \neg\exists x\,(\neg\psi(x))] \in \Gamma \\ \text{or } &[\psi(t) \to \forall x\,(\psi(x))] \in \Gamma\end{aligned}$$

Let $\mathfrak{I} : \mathbf{Var} \to \mathcal{M}$ be an interpretation. Denote $[v]_\Gamma = \mathfrak{I}(x)$, then

□

6.7.2 Semantical Categoricity

The way the Hilbert's system is designed makes any *reasonable* interpretation coincide with the semantical way we think about logical connectives. Therefore, we may try to design predicate calculus in the same vein. But how do we semantically interpret?

Given a signature σ, we say that (M, σ^M) is a σ-structure if, there is one-to-one correspondence between σ and σ^M (assigning φ to φ^M), such that if $\varphi \in \sigma$ is of arity n, then φ^M is an n-ary relation on M (i.e. $\varphi^M \subseteq M^n$).

Given M to be a σ-structure, and $\mathfrak{I} : \mathbf{Var} \to M$ to be a variable assignment. We define $M, \mathfrak{I} \vDash \varphi$ inductively as follows

1. If x_i are the variables associated with $\varphi \in \sigma$, then $M, \mathfrak{I} \vDash \varphi$ iff $(\mathfrak{I}(x_1), \ldots, \mathfrak{I}(x_n)) \in \varphi^M$, and $M, \mathfrak{I} \nvDash \varphi$ iff $(\mathfrak{I}(x_1), \ldots, \mathfrak{I}(x_n)) \in \varphi^M$
2. $M, \mathfrak{I} \vDash \neg\varphi$ iff $M, \mathfrak{I} \nvDash \varphi$.
3. $M, \mathfrak{I} \nvDash \neg\varphi \to \psi$ iff $M, \mathfrak{I} \vDash \varphi$, and $M, \mathfrak{I} \vDash \psi$.
4. $M, \mathfrak{I} \vDash \varphi[x, y := z]$ if $M, \mathfrak{I} \circ \mathcal{S}_z^{x,y} \vDash \varphi$, where $\mathcal{S}_z^{x,y} : \mathbf{Var} \to \mathbf{Var}$ is the variable reassignment
$$\mathcal{S}_z^{x,y}(x) = \mathcal{S}_z^{x,y}(y) = z$$
$$\mathcal{S}_z^{x,y}(w) = w \text{ for all } w \notin \{x, y\}$$
5. $M, \mathfrak{I} \vDash \forall x(\varphi)$ if for any x-variant $\mathfrak{K}$ (i.e. $\mathfrak{I}(y) = \mathfrak{K}(y)$ for any $y \neq x$) including $\mathfrak{I}$, we have $M, \mathfrak{K} \vDash \varphi$

The semantic rules for logical connectives are justified based upon the previous categorical result.

If $M, \mathfrak{I} \vDash \varphi$ for any $\mathfrak{I} \in M^{\mathbf{Var}}$, then we say $M \vDash \varphi$.

Theorem 6.37 (Universal Instantiation & Generalization)**.** Let φ be a wff of arity $n > 0$, then $M \vDash \varphi(\vec{x})$ iff $M \vDash \forall\vec{x}(\varphi(\vec{x}))$.

Proof. Suppose $M \vDash \varphi(\vec{x})$, then for any $\mathfrak{I} \in M^{\mathbf{Var}}$ and $\vec{x}$-variant $\mathfrak{H}$, we should still have $M, \mathfrak{H} \vDash \varphi(\vec{x})$ by definition. However, this is just another way of saying that $M, \mathfrak{I} \vDash \forall\vec{x}(\varphi(\vec{x}))$. Thus, $M \vDash \forall\vec{x}(\varphi(\vec{x}))$.

Next, suppose $M \vDash \forall\vec{x}(\varphi(\vec{x}))$. Let $\mathfrak{L}$ be a variable assignment in M.

Then $M, \mathfrak{L} \vDash \forall \vec{x}(\varphi(\vec{x}))$, so $M, \mathfrak{I} \vDash \varphi(\vec{x})$ for any $\vec{x}$-variant $\mathfrak{I}$ of $\mathfrak{L}$. Now note that $\mathfrak{L}$ is actually $\vec{x}$-variant of itself, so $M, \mathfrak{L} \vDash \varphi(\vec{x})$. As $\mathfrak{L}$ is arbitrary, we conclude $M \vDash \varphi(\vec{x})$ □

Chapter 7

Descriptive Set Theory

7.1 Cantor-Bendixson Rank

Given a set A in a topological space, we can define a transfinite sequence as follows:

1. $\mathrm{lp}_0(A) = A$
2. $\mathrm{lp}_{\alpha+1}(A) = \mathrm{lp}(\mathrm{lp}_\alpha(A))$
3. $\mathrm{lp}_\gamma(A) = \bigcap_{\beta<\gamma} \mathrm{lp}_\beta(A)$ for a limit ordinal γ.

Theorem 7.1 (Cantor-Bendixson Rank)**.**

1. For arbitrary set A, the above sequence is eventually periodic, with the period at most ω.
2. However, for a closed set C, we have a stronger result: the above sequence stabilizes, and the smallest ordinal ρ such that $\mathrm{lp}_\rho(C) = \mathrm{lp}_{\rho+1}(C)$ is called the *Cantor-Bendixson rank* of C.

Proof. Part 1: note that $A \supseteq \mathrm{lp}_\omega(A) \supseteq \mathrm{lp}_{\omega\cdot 2}(A) \supseteq \cdots$, so the sequence $\mathrm{lp}_\gamma(A)$ with γ a limit ordinal is decreasing. Let $\Gamma_A = \bigcap_\alpha \mathrm{lp}_\alpha(A) \subseteq A$. For each $x \in A \setminus \Gamma_A = \bigcup_\alpha A \setminus \mathrm{lp}_\alpha(A)$, there must be some ordinal σ_x such that $x \notin \mathrm{lp}_{\sigma_x}(A)$. Let σ_x be the least such ordinal. By replacement, $\{\sigma_x : x \in A\setminus\Gamma_A\}$ is a set (of ordinals), so let μ be a *limit* ordinal that bounds this set above. By construction, $\mathrm{lp}_\mu(A) \subseteq \mathrm{lp}_{\sigma_x}(A)$, so $x \notin \mathrm{lp}_\mu(A)$ for all $x \in A \setminus \Gamma_A$. Equivalently, $A \setminus \Gamma_A \subseteq A \setminus \mathrm{lp}_\mu(A)$, or $\Gamma_A \supseteq \mathrm{lp}_\mu(A)$. On the other hand, Γ_A is the intersection of all $\mathrm{lp}_\alpha(A)$, so we must get $\Gamma_A = \mathrm{lp}_\mu(A)$. Since μ is a limit ordinal, $\mu + \omega \cdot \alpha$ is also a limit ordinal (for any ordinal α). Hence, $\Gamma_A = \mathrm{lp}_\mu(A) \supseteq \mathrm{lp}_{\mu+\omega\cdot\alpha}(A)$, which again shows that $\Gamma_A = \mathrm{lp}_\mu(A) = \mathrm{lp}_{\mu+\omega}(A) = \mathrm{lp}_{\mu+\omega\cdot 2}(A) = \cdots$. We conclude that the sequence is eventually periodic with period at most ω.

Part 2: if $A = C$ is closed, then $\mathrm{lp}(C) \subseteq C$, and we can prove by induction that

1. $\mathrm{lp}_\alpha(C)$ is closed.
2. $\mathrm{lp}_\alpha(C)$ is decreasing on α.

Based on this, we can show analogously (as with part 1) that $\Gamma_C = \mathrm{lp}_\mu(C) = \mathrm{lp}_{\mu+1}(C) = \mathrm{lp}_{\mu+2}(C) = \cdots$, i.e. the sequence stabilizes. Because of that, Γ_C is also closed. □

Remark 7.2. Let X be an indiscrete space with more than 2 elements, then $\mathrm{lp}(S) = X$ for any subset S of X. Hence, $\mathrm{lp}_n(S) = X$ for any $n \geq 1$. Yet, $\mathrm{lp}_\omega(S) = S$ (since $\mathrm{lp}_0(S) = S$), so the sequence is periodic with the period exactly ω.

Chapter 8

General Topology II

8.1 Kuratowski Monoid

[Reference: The Kuratowski Closure-Complement Theorem]

Theorem 8.1 (Kuratowski 14 sets)**.** Let X be a topological space and $E \subseteq X$. Then, at most 14 distinct subsets of X can be obtained from E by taking closure and complement

8.2 Convergence spaces

A convergence space is a pair $(X, \to)$ where X is a set, and $\to$ is a relation from the collection $\text{Filter}(X)$ of filters on X, to X, such that

1. The principal filter $\mathcal{F}_x = \{A \subseteq X : x \in A\}$ at x converges to x (i.e. $\mathcal{F}_x \to x$).
2. If $\mathcal{F} \subseteq \mathcal{G}$ and $\mathcal{F} \to x$, then $\mathcal{G} \to x$.
3. If $\mathcal{F}, \mathcal{G} \to x$, then there exists some filter $\mathcal{H} \subseteq \mathcal{F} \cap \mathcal{G}$ such that $\mathcal{H} \to x$.

8.2.1 Neighborhood system

A function $\mathcal{N} : X \to \text{Filter}(X)$ from X to the collection of filters on X is called a *neighborhood system* if $p \in N$ for any $N \in \mathcal{N}_p := \mathcal{N}(p)$. A *pretopological space* is a pair $(X, \mathcal{N})$ where $\mathcal{N}$ is a neighborhood system on X. If $N \in \mathcal{N}_p$, we say that N is a neighborhood of p.

A continuous function $f : (X, \mathcal{N}) \to (Y, \mathcal{M})$ between these spaces consists of a function $f : X \to Y$ such that $f^{-1}(V) \in \mathcal{N}_p$ whenever $V \in \mathcal{M}_{f(p)}$. Most of the time, we just use the same letters $\mathcal{N}$ for both neighborhood systems on X and Y for convenience.

Remark 8.2. An equivalent definition of morphism is: for each neighborhood V of $f(p)$, there exists a neighborhood U of p such that $f(U) \subseteq V$.

Proof. Suppose $f : X \to Y$ is continuous in the first sense, then for any $V \in \mathcal{N}_{f(p)}$, we can take $U = f^{-1}(V) \in \mathcal{N}_p$ so that $f(U) = f(f^{-1}(V)) \subseteq V$. Vice versa, suppose $f : X \to Y$ is continuous in the second sense, then for any $V \in \mathcal{N}_{f(p)}$, there exists $U \in \mathcal{N}_p$ such that $f(U) \subseteq V$. This means that $U \subseteq f^{-1}(V)$. Since $\mathcal{N}_p$ is a filter, we have $f^{-1}(V) \in \mathcal{N}_p$. □

Proposition 8.3. The pretopological spaces, along with the continuous functions, form a category, denoted as **PreTop**.

Proof.
Identity: the identity function $\mathrm{id}_X : (X, \mathcal{N}) \to (X, \mathcal{N})$ is continuous. Indeed, for each $V \in \mathcal{N}_{\mathrm{id}_X(p)} = \mathcal{N}_p$, we have $\mathrm{id}_X^{-1}(V) = V \in \mathcal{N}_p$.
Composition: given $f : (X, \mathcal{N}^X) \to (Y, \mathcal{N}^Y)$ and $g : (Y, \mathcal{N}^Y) \to (Z, \mathcal{N}^Z)$ continuous, we show that $gf : (X, \mathcal{N}^X) \to (Z, \mathcal{N}^Z)$ is also continuous. Let $W \in \mathcal{N}^Z_{gf(p)}$, then $g^{-1}(W) \in \mathcal{N}^Y_{f(p)}$, and $(gf)^{-1}(W) = f^{-1}(g^{-1}(W)) \in \mathcal{N}^X_p$. □

Suppose X is a topological space, let $\mathrm{Nbh}_X(p)$ (or simply $\mathrm{Nbh}(p)$) be the collection of all neighborhoods around p.

Theorem 8.4. The forget functor $L : \mathbf{Top} \to \mathbf{PreTop}$ by sending

1. $(X, \mathcal{T})$ to (X, Nbh)
2. $f : X \to Y$ to the same underlying function f

is well-defined, fully faithful, and has a left adjoint G that is also a left inverse of L.

Proof.
Well-defined: first we show that Nbh is a neighborhood system on X

1. By definition, $p \in N$ for all neighborhood N around p.
2. If N is a neighborhood of p, then there exists an open set U such that $p \in U \subseteq N$. Any M containing N will automatically contains U, so it should be also a neighborhood of p.
3. If N_1, N_2 are neighborhoods of p, then there exists open sets U_1, U_2 such that $p \in U_k \subseteq N_k$ $(k = 1, 2)$. Thus $p \in U_1 \cap U_2 \subseteq N_1 \cap N_2$, which shows that $N_1 \cap N_2$ is neighborhood of p.

Fully faithful: straight from equivalent definition of continuity (in topological space).

Adjoint: the adjoint G is given by the following: for each pretopological

space $(X, \mathcal{N})$, construct $\mathcal{T}$ as the collection of U such that it is a neighborhood of itself (i.e. $U \in \mathcal{N}_p$ for all $p \in U$). G then sends $(X, \mathcal{N})$ to $(X, \mathcal{T})$, and $f : X \to Y$ remains the same (just like the functor L).

We first verify that $\mathcal{T}$ is a topology on X.

1. The empty set is trivially in $\mathcal{T}$. X is also in $\mathcal{T}$ since each $\mathcal{N}_p$ is a filter, so $X \in \mathcal{N}_p$.
2. Given a family $\{U_i\}_{i \in I}$ of sets in $\mathcal{T}$, let p be a point in $U = \bigcup_{i \in I} U_i$. By definition, there exist some $i_0 \in I$ such that $p \in U_i$. Thus, $U_i \in \mathcal{N}_p$. Since $\mathcal{N}_p$ is a filter, $U \in \mathcal{N}_p$, which implies that $U \in \mathcal{T}$.
3. Given 2 sets U and V in $\mathcal{T}$, and any $p \in U \cap V$. Then we know that $p \in U, V$, so $U, V \in \mathcal{N}_p$. But $\mathcal{N}_p$, so we get $U \cap V \in \mathcal{N}_p$, and hence, $U \cap V \in \mathcal{T}$.

Given $f : (X, \mathcal{N}) \to (Y, \mathcal{M})$ continuous, we next show that $G(f) = f : (X, \mathcal{T}) \to (Y, \mathcal{S})$ is continuous, where $G(X, \mathcal{N}) = (X, \mathcal{T})$ and $G(Y, \mathcal{M}) = (Y, \mathcal{S})$. Let $V \in \mathcal{S}$, and $p \in U = f^{-1}(V)$, then $f(p) \in V$, so $V \in \mathcal{M}_{f(p)}$. By definition, $U \in \mathcal{N}_p$. But p is an arbitrary point in U, so by construction of $\mathcal{T}$, we have $U \in \mathcal{T}$.

Next, we show that GL is the identity functor on **Top**. Denote $(X, \mathcal{N}) = L(X, \mathcal{T})$ and $(X, \mathcal{T}') = G(X, \mathcal{N})$. Let $U \in \mathcal{T}'$, then $U \in \mathcal{N}_p$ for all $p \in U$. By Remark 4.9, $U \in \mathcal{T}$. Therefore, $\mathcal{T}' \subseteq \mathcal{T}$. Vice versa, let $U \in \mathcal{T}$, then $U \in \mathcal{N}_p$ for each $p \in U$ (by definition of a neighborhood). Thus, $U \in \mathcal{T}'$ by definition. We conclude $\mathcal{T} = \mathcal{T}'$.

Finally, we prove G is the left adjoint to L.
Note that

1. Suppose $f : (X, \mathcal{N}) \to (Y, \mathrm{Nbh})$ is continuous. Let $V \in \mathcal{T}$, we need to show that $f^{-1}(V) \in \mathcal{T}_{\mathcal{N}}$, or equivalently, $f^{-1}(V) \in \mathcal{N}_p$ for all $p \in f^{-1}(V)$. Let $p \in f^{-1}(V)$, then $f(p) \in V$, so $V \in \mathrm{Nbh}(f(p))$. Thus, $f^{-1}(V) \in \mathcal{N}_p$.
2. Vice versa, suppose $f : (X, \mathcal{T}_{\mathcal{N}}) \to (Y, \mathcal{T})$ is continuous. Let $V \in \mathrm{Nbh}(f(p))$, then there exists an open set V' such that $f(p) \in V' \subseteq V$. Thus, $f^{-1}(V')$ is open and $p \in f^{-1}(V') \subseteq f^{-1}(V)$. By definition of $\mathcal{T}_{\mathcal{N}}$, we have $f^{-1}(V') \in \mathcal{N}_p$ and so $f^{-1}(V) \in \mathcal{N}_p$.

Hence, we have established a bijection $j_{X,Y}$ between $\mathrm{Hom}_{\mathbf{PreTop}}(X, LY)$ and $\mathrm{Hom}_{\mathbf{Top}}(GX, Y)$. This is also natural, intuitively because the function itself remains the same between different categories. Concretely, we'll show that j is a natural isomorphism between

$$j : \mathrm{Hom}_{\mathbf{PreTop}}(-, L-) \to \mathrm{Hom}_{\mathbf{Top}}(G-, -)$$

These functors are of the form $\mathbf{PreTop}^{op} \times \mathbf{Top} \to \mathbf{Set}$. For each morphism $(\varphi, \psi) : (X, Y) \to (Z, W)$ in $\mathbf{PreTop}^{op} \times \mathbf{Top}$ (i.e. $\varphi : Z \to X$ is a morphism in $\mathbf{PreTop}$, $\psi : Y \to W$ is a morphism in $\mathbf{Top}$), and $f \in \mathrm{Hom}_{\mathbf{PreTop}}(X, LY)$

$$\begin{aligned} j_{Z,W} \circ \mathrm{Hom}_{\mathbf{PreTop}}(\varphi, L\psi)(f) &= \mathrm{Hom}_{\mathbf{PreTop}}(\varphi, L\psi)(f) \\ &= L\psi \circ f \circ \varphi = \psi \circ f \circ G\varphi \\ &= \mathrm{Hom}_{\mathbf{Top}}(G\varphi, \psi)(f) \\ &= \mathrm{Hom}_{\mathbf{Top}}(G\varphi, \psi) \circ j_{X,Y}(f) \end{aligned}$$

Recall that $L\psi = \psi$ and $G\varphi = \varphi$. □

Theorem 8.5 (Topology in terms of neighborhood system)**.** The image of the functor L consists exactly of the pretopological spaces $(X, \mathcal{N})$ satisfying the following property: each neighborhood N of p contains a neighborhood M of p such that N is a neighborhood of the whole M

$$\forall p \in X, \forall N \in \mathcal{N}_p, \exists M \in \mathcal{N}_p, \forall x \in M : N \in \mathcal{N}_x$$

Proof. Suppose $(\mathcal{N}_p)_{p \in X}$ is a neighborhood system on X, and $\mathcal{T}$ is the topology constructed from it. Given N a neighborhood of p (with respect to $\mathcal{T}$, not $\mathcal{N}_p$), there exists some $U \in \mathcal{T}$ such that $p \in U \subseteq N$ by definition. However, by construction, $U \in \mathcal{N}_p$, so $N \in \mathcal{N}_p$. This shows that $\mathrm{Nbh}(p) \subseteq \mathcal{N}_p$.

On the other hand, suppose $(X, \mathcal{N})$ satisfies the condition. Let $N \in \mathcal{N}_p$, and denote $U = \{x \in N : N \in \mathcal{N}_x\}$. Given $x \in X$, if $N \in \mathcal{N}_x$, then there exists some $M_x \in \mathcal{N}_x$ such that $N \in \mathcal{N}_y$ for all $y \in M_x$. By construction, $M_x \subseteq U$ (note that $x \in N$ whenever $N \in \mathcal{N}_x$). But $\mathcal{N}_x$ is a filter, so $U \in \mathcal{N}_x$. In summary, $N \in \mathcal{N}_x$ implies $U \in \mathcal{N}_x$. In particular,

1. $U \in \mathcal{N}_p$. So we get $p \in U \subseteq N$.
2. For all $x \in U$, $N \in \mathcal{N}_x$, so $U \in \mathcal{N}_x$ also. In other words, $U \in \mathcal{T}$.

Therefore, N is neighborhood of p (with respect to $\mathcal{T}$). This shows that $\mathcal{N}_p \subseteq \mathrm{Nbh}(p)$. □

Remark 8.6. The above result gives us another way to prove the adjunction between L and G: note that if $(X, \mathcal{M}) = LG(X, \mathcal{N})$, then $\mathcal{M}_p \subseteq \mathcal{N}_p$ for each $p \in X$. Therefore, the map $\mathrm{id}_X : (X, \mathcal{N}) \to (X, \mathcal{M})$ (note the reversal of direction) is continuous. This provides the unit $\eta : 1_{\mathbf{PreTop}} \to LG$, while the counit $\varepsilon : GL \to 1_{\mathbf{Top}}$ is just the identity transformation

Example 8.7. Since every set in $\mathcal{N}_p$ has to contain p, the filter $\mathcal{N}_p$ is a subset of the principal filter at p.

1. If, for each $p \in X$, we have $\mathcal{N}_p = \uparrow \{p\}$, then $G(X, \mathcal{N})$ is the discrete space on X.
2. If, for each $p \in X$, we have $\mathcal{N}_p = \{X\}$, then $G(X, \mathcal{N})$ is the indiscrete space on X.

8.2.2 Čech closure operator

Given a set X, a *preclosure operator* (or Čech closure operator) is a function cl on the power set of X such that

1. $\mathrm{cl}(\varnothing) = \varnothing$.
2. $A \subseteq \mathrm{cl}(A)$.
3. $\mathrm{cl}(A \cup B) = \mathrm{cl}(A) \cup \mathrm{cl}(B)$

Such pair (X, cl) is then called the *preclosure space*.

A map $f : (X, \mathrm{cl}_X) \to (Y, \mathrm{cl}_Y)$ consists of a function $f : X \to Y$. We say f is continuous if $f(\mathrm{cl}_X(A)) \subseteq \mathrm{cl}_Y(f(A))$ for any $A \subseteq X$.

Proposition 8.8. Preclosure spaces, along with its continuous functions, form a category, denoted as **CloTop**.

Proof.
Identity: the identity $\mathrm{id}_X : X \to X$ is continuous since for each $A \subseteq X$, $\mathrm{id}_X(\mathrm{cl}(A)) = \mathrm{cl}(A) = \mathrm{cl}(\mathrm{id}_X(A))$.
Composition: if $f : (X, \mathrm{cl}_X) \to (Y, \mathrm{cl}_Y)$ and $g : (Y, \mathrm{cl}_Y) \to (Z, \mathrm{cl}_Z)$ are continuous, then $gf : (X, \mathrm{cl}_X) \to (Z, \mathrm{cl}_Z)$ is also continuous. Indeed, let $A \subseteq X$, we have

$$gf(\mathrm{cl}_X(A)) \subseteq g(\mathrm{cl}_Y(f(A))) \subseteq \mathrm{cl}_Z(gf(A))$$

□

It turns out that this category is actually equivalent with **PreTop**, so we can interchange different viewpoints regarding pretopological spaces. To do this, we introduce the concept of convergence

Theorem 8.9. From (X, cl), we can define

Corollary 8.10 (Kuratowski closure axiom)**.**

Chapter 9

Euclidean Geometry

Theorem 9.1 (Routh Theorem)**.** Let $\triangle ABC$ be a triangle, and D, E, F are respectively points on the lines BC, CA, AB. Denote $x = \overline{AF}/\overline{FB}, y = \overline{BD}/\overline{DC}, z = \overline{CE}/\overline{EA}$.

1. If P is the intersection of BE and CF, Q is the intersection of CF and AD, and R is the intersection of AD and BE, then

$$\frac{S_{PQR}}{S_{ABC}} = \frac{(xyz-1)^2}{(xy+x+1)(yz+y+1)(zx+z+1)}$$

   where S_{ABC} denotes the area of $\triangle ABC$.

2.

$$\frac{S_{DEF}}{S_{ABC}} = \frac{xyz+1}{(x+1)(y+1)(z+1)}$$

Corollary 9.2.

1. Ceva Theorem: AD, BE, CF are concurrent iff $xyz = 1$.
2. Menelaus Theorem: D, E, F are collinear iff $xyz = -1$.

Proposition 9.3. For a given simplex S in $\mathbb{R}^n$ with its set of vertices $\mathcal{V}$.

$$\sum_{A\in\mathcal{V}} \vec{GA} = 0 \text{ iff } G = \frac{\int_S x dV}{\int_S dV} = \frac{\int_S x dV}{\text{vol}(S)}$$

Theorem 9.4. Let A, B, C, D, E, F be points in Euclidean plane that are not all collinear. Denote M to be the intersection of AB and DE, N to be the intersection of BC and EF, and P to be the intersection of CD and FA. Prove that

1. Pascal theorem: if A, B, C, D, E, F lies on a conic, then M, N, P are collinear.
2. Braikenridge-Maclaurin theorem: if M, N, P are collinear, then A, B, C, D, E, F lies on some conic.

Theorem 9.5 (Poncelet Porism)**.**

1. Given a conic $\mathcal{C}$ inside and conic $\mathcal{I}$ such that there some polygon with n sides (for some fixed n) that is inscribed in $\mathcal{I}$, and circumscribed $\mathcal{C}$, then any poiny on $\mathcal{I}$ is a vertex of such polygon (still with n sides).
2. Cayley existence criteria (see *On Cayley's explicit solution to Poncelet's porism*, by Phillip Griffiths and Joseph Harris)
3. [Richelot 1830, Kerawala 1947] Let (I, r) be a circle (centered at I, of radius r) inside another circle (O, R). If
   (a) $d = OI, a = 1/(R+d), b = 1/(R-d), c = 1/r$
   (b) $\lambda = 1 + \frac{2c^2(a^2-b^2)}{a^2(b^2-c^2)}, \omega = \cosh^{-1}\lambda, k^2 = 1 - e^{-2\omega}$

   Then there exists an n-side polygon that is inscribed in (O, R), and circumscribed (I, r) if and only if

$$\mathrm{sc}\left(\frac{K(k)}{n}, k\right) = \frac{c\sqrt{b^2-a^2} + b\sqrt{c^2-a^2}}{a(b+c)}$$

   where $\mathrm{sc}(x, y)$ is the Jacobi elliptic function, and $K(z)$ is the complete elliptic integral of the 1st kind. As a result
   (a) When $n = 3$ (Euler, Chapple): we have the identity $d^2 = R(R-2r)$ between the radii of inscribed and circumscribed circles of a triangle, and the distance between them.
   (b) Similarly, when $n = 4$ (Nicolas Fuss): we have the identity

$$\frac{1}{(R-d)^2} + \frac{1}{(R+d)^2} = \frac{1}{r^2}$$

4. If a polygon is inscribed and circumscribed by circles, then
   (a) If the number of sides is even, then the lines, each of which connects the contact points (to the inscribed circle) of 2 opposite sides, are concurrent.
   (b) If the number of sides is odd, then the lines, each of which connects a vertex to the contact point of the opposite side, are concurrent.

   Additionally, the point of concurrency is the limiting point of the 2 circles.

Theorem 9.6 (Porcupine Theorem)**.** Given a closed, orientable, smooth (or just C^1) manifold M of dimension $n-1$, possibly with rough corners, such that it is embedded in $\mathbb{R}^n$. We then have

$$\int_M \vec{n} dS = \vec{0}$$

Where $\vec{n}$ is the normal vector at point x.

Theorem 9.7 (Abboud-Viviani invariant)**.** Given a convex polygon $A_1A_2 \dots A_n$ (with $A_{n+1} = A_1$) which has vertex angles equal to each other (i.e. an equi-angular polygon), and T is a point on the plane. Let

1. S be the *total width* of the polygon

$$S = \begin{cases} \text{The sum of distances between opposite sides} & \text{if } 2 \mid n \\ \text{The sum of distances between a vertex and its opposite side} & \text{if } 2 \nmid n \end{cases}$$

2. D_n be the *Abboud-Viviani invariant* of n:

$$D_n = \begin{cases} 1 & \text{if } n \text{ is even} \\ \frac{\cos(\pi/n)}{\cos(\pi/n)+1} & \text{if } n \text{ is odd} \end{cases}$$

Then

$$\sum_{k=1}^{n} \overline{d(T, A_kA_{k+1})} = D_n S$$

where the sign of $\overline{d(T,\ell)}$ is positive if T is on the same side as the polygon with respect to the line ℓ, and negative otherwise.

Theorem 9.8 (Viviani conservation principle)**.** For a convex polygon $A_1A_2 \dots A_n$, and T is an arbitrary point inside the polygon, if

$$\textstyle\sum_{k=1}^{n} d(T, A_kA_{k+1}) \text{ is constant, regardless of the position of } T$$

then $\sum_{k=1}^{n} \vec{t_k} = 0$, where $\vec{t_k}$ is the projection *unit* vector of T into A_kA_{k+1} (note that these vectors depends only on the polygon and not T). In fact,

$$\sum_{k=1}^{n} \overline{d(T, A_kA_{k+1})} = \text{const if and only if } \sum_{k=1}^{n} \vec{t_k} = 0$$

Corollary 9.9. When $n = 3$, then the sum is constant if and only if the triangle is equilateral.

Theorem 9.10 (Erdős–Mordell inequality)**.** Let $A_1A_2\dots A_n$ be a convex n-gon, P be a point in the polygon. Denote $r_i = PA_i$, $d_{i,j} = d(P, A_iA_j)$, $w_{i,j}$ to be the bisector length of $\angle A_iPA_j$ from P to its intersection with A_iA_j.

1. [Lenhard]

$$\sum_{i=1}^{n} r_i \geq \sec\frac{\pi}{n}\sum_{i=1}^{n} w_{i,i+1} \geq \sec\frac{\pi}{n}\sum_{i=1}^{n} d_{i,i+1}$$

   Corollary [Barrow]: $r_1 + r_2 + r_3 \geq 2(w_{1,2} + w_{2,3} + w_{3,1})$, with equality holds only for equilateral triangle, and P is its center.

2. [Gueron, Shafrir] For any $\lambda_i \geq 0$

$$\sum_{i=1}^{n} \lambda_i r_i \geq \sec\frac{\pi}{n}\sum_{i=1}^{n} \sqrt{\lambda_i\lambda_{i+1}}\, d_{i,i+1}$$

   Corollary [Weighted Erdős–Mordell]: $\lambda_1 r_1 + \lambda_2 r_2 + \lambda_3 r_3 \geq 2(\sqrt{\lambda_1\lambda_2}d_{1,2} + \sqrt{\lambda_2\lambda_3}d_{2,3} + \sqrt{\lambda_3\lambda_1}d_{3,1})$, with equality holds only for equilateral triangle, P is its center, and $\lambda_1 = \lambda_2 = \lambda_3$.

3. [N. Ozeki]

$$\prod_{i=1}^{n} r_i \geq \sec^n\frac{\pi}{n}\prod_{i=1}^{n} w_{i,i+1}$$

4. [L. Fejes, Tóth]

$$\prod_{i=1}^{n} r_i \geq \sec^n\frac{\pi}{n}\prod_{i=1}^{n} d_{i,i+1}$$

5. Conjecture:

   (a) Generalization of Gueron-Shafrir and Lenhard results:

$$\sum_{i=1}^{n} \lambda_i r_i \geq \sec\frac{\pi}{n}\sum_{i=1}^{n} \sqrt{\lambda_i\lambda_{i+1}}\, w_{i,i+1} \geq \sec\frac{\pi}{n}\sum_{i=1}^{n} \sqrt{\lambda_i\lambda_{i+1}}\, d_{i,i+1}$$

   (b) Weighted N. Ozeki inequality:

$$\prod_{i=1}^{n} r_i^{\lambda_i} \geq \sec^n\frac{\pi}{n}\prod_{i=1}^{n} w_{i,i+1}^{\sqrt{\lambda_i\lambda_{i+1}}}$$

Theorem 9.11 (Kazarinoff)**.** Let $ABCD$ be a tetrahedron, P be a point inside. Respectively, denote

1. d_A, d_B, d_C, d_D to be the distances from P to the plane BCD, CDA, DAB, ABC.
2. R_A, R_B, R_C, R_D to be the length of PA, PB, PC, PD.

Then

$$\frac{R_A + R_B + R_C + R_D}{d_A + d_B + d_C + d_D} > 2\sqrt{2}$$

In fact, $2\sqrt{2}$ is the best (or largest) possible constant.

Theorem 9.12 (Power Summation Theorem)**.**

1. Let $A_1A_2 \dots A_n$ be a regular n-gon inscribed in (O, R) with an arbitrary point M on the circle. Suppose $0 \leq t < n$ is an integer, we have

$$\frac{1}{R^{2t}} \sum_{i=1}^{n} MA_i^{2t} = n\binom{2t}{t}$$

   (special thanks to OEIS - On-Line Encyclopedia of Integer Sequences, for the constant $\binom{2t}{t}$)
   Conjectures:

   (a) If t is not an integer, or is an integer but not in between 0 and $n-1$ (inclusively), then the sum is not constant (i.e. it depends on the position of M). If so, can there be an explicit formula?

   (b) What other polygons (beside regular ones) have such property, i.e. the sum is constant for some power t.

2. Let $A_1A_2 \dots A_n$ be a regular n-gon inscribed in (O, R) with an arbitrary line ℓ going through the center O. Suppose $t \geq 0$ is an integer such that $2t < n$ if n is even, and $t < n$ if n is odd, show that

$$\frac{1}{R^{2t}} \sum_{i=1}^{n} d(A_i, \ell)^{2t} = \frac{n}{4^t}\binom{2t}{t}$$

3. Let $A_1A_2 \dots A_n$ be a regular n-gon inscribed in (O, R) with an arbitrary point M on the circle. Suppose t is an integer such that $0 \leq 2t < n$, we have

$$\sum_{i=1}^{n} d(M, A_iA_{i+1})^{2t} = \text{const, regardless of } M$$

Theorem 9.13 (Van Schooten Theorem)**.** 1. [Bui Quang Tuan] Let $A_1A_2\dots A_n$ be a polygon inscribed in a circle centered at O, with P be a point on the circle. Denote $a_i = A_iA_{i+1}, d_i = d(P, A_iA_{i+1})$ and $t_i = a_i/d_i$. Prove that among $t_1, t_2, \dots, t_n$, there is one t_i that sums up the other t_j (for $j \neq i$)

2. Partial corollary [Van Schooten]: for a triangle $\triangle ABC$, show that

   $\triangle ABC$ is an equilateral triangle if and only if one of PA, PB, PC sums up the other 2 for any point P on the circumcircle of $\triangle ABC$.

Theorem 9.14 (Petr-Douglas-Neumann theorem)**.** Let P_0 be an n-gon ($n \geq 3$), and $(k_i)_{i=0}^{n-3}$ be a permutation of $\{1, 2, \dots, n-2\}$. Construct a sequence of polygon $(P_i)_{i=0}^{n-2}$ (with the intital one to be the same as P_0) recursively as follows

1. Given P_i, draw isoceles triangles with apex angle $k_i 2\pi/n$ on each sides of the polygon P_i such that if the isoceles triangle with base S_hS_{h+1} is on the same side as the side $S_{h+1}S_{h+2}$ with respect to the same base S_hS_{h+1}, then the next isoceles triangle at $S_{h+1}S_{h+2}$ has to be on the same side as S_hS_{h+1} with respect to its base $S_{h+1}S_{h+2}$. Such an arrangement consitute an *orientation* of the polygon, and there are exactly 2 distinct orientations for any polygon (even if it is self-intersecting).

2. The apices of these isoceles triangles form the next polygon P_{i+1}, and the order of vertices follows the same order as the bases of isoceles triangles, i.e. any 2 vertices of P_{i+1} are adjacent if the repective bases (from the isoceles triangles) are the adjacent sides of P_i.

Show that P_{n-2} is a regular n-gon whose centroid coincides with the centroid of the original polygon P_0.

Corollary 9.15.

1. Napoleon theorem: the centers of the equilaterals constructing at the sides of some triangle forms another equilateral.

2. Van Aubel theorem: let $ABCD$ be a quadrilateral. Let P, Q, R, S be the center of the squares constructing on the side AB, BC, CD, DA respectively. Then $PR = QS$ and $PR \perp QS$.

3. Thébault problem I: given any parallelogram, the centers of the squares constructed at the 4 sides form another square.

Theorem 9.16. Let

$$s(x) = \begin{cases} \frac{x}{2} & \text{for Euclidean geometry} \\ \sinh \frac{x}{2} & \text{for hyperbolic geometry} \\ \sin \frac{x}{2} & \text{for spherical geometry} \end{cases}$$

1. Conjecture: suppose $(C_1, r_1), \ldots, (C_n, r_n)$ $(n \geq 3)$ are n circles tangent to some other circle (O, R). Define

$$t_{ij} = \begin{cases} \text{The segment of exterior common tangent of } (C_i) \text{ and } (C_j) \\ \qquad \text{if they are on the same side of } (O) \\ \text{The segment of interior common tangent of } (C_i) \text{ and } (C_j) \\ \qquad \text{if they are on the different sides of } (O) \end{cases}$$

   If

   (a) E_n is the set of polynomials f of $n(n+1)/2$ variables such that $f(|t_{ij}|_e, |r_i|_e) = 0$ when the circles (C_i) varies, where $|AB|_e$ is the Euclidean distance between 2 points.

   (b) H_n is the set of polynomials f of $n(n+1)/2$ variables such that $f(\sinh \frac{|t_{ij}|_h}{2}, \frac{1}{2} \sinh |r_i|_h) = 0$ when the circles (C_i) varies, where $|AB|_h$ is the hyperbolic distance between 2 points.

   (c) S_n is the set of polynomials f of $n(n+1)/2$ variables such that $f(\sin \frac{|t_{ij}|_s}{2}, \frac{1}{2} \sin |r_i|_s) = 0$ when the circles (C_i) varies, where $|AB|_s$ is the spherical distance between 2 points.

   Then $E_n = H_n = S_n$, and every polynomial in either of them is homogeneous.

2. [Gregorac, Ren Guo, Nilgun Sonmez] Same as above, but when the circles degenerate into points.

3. [Gregorac, Ren Guo, Nilgun Sonmez] Let $A_0, A_1, \ldots, A_{2n-1}$ be points on a circle (in one the 3 geometries described above). Denote l_{ij} to be the length of the geodesic segment A_iA_j (for $i \neq j$). Then $\det(a_{ij}) = 0$ where $a_{ij} = (-1)^{\delta_{ij}}[s(l_{2i-2,2j-1})s(l_{2j-1,2i})]^{-1}$ (and δ_{ij} is the Kronecker delta).

4. [Gregorac, Ren Guo, Nilgun Sonmez] Let $P_0, P_1, \ldots, P_n$ be points on a circle (in one the 3 geometries described above). Denote l_{ij} to be the length of the geodesic segment P_iP_j (for $i \neq j$).

   (a) For each $|j - i| \geq 2$, there exists some polynomial F_{ij} such that $s(l_{ij}) = F_{ij}(s(l_{12}, l_{23}, \ldots, l_{n1}))$.

(b) If R is the radius of the circle, there exists some polynomial G such that $s(2R) = G(s(l_{12}, l_{23}, \ldots, l_{n1}))$.

5. [Fuhrmann] Let $(C_1), (C_2), (C_3), (C_4), (C_5), (C_6)$ (in that order, clockwise) be 6 circles tangent to another circles (O). Define t_{ij} as in part (1) previously, but replace with the length of the tangent. Then

$$t_{14}t_{25}t_{36} = t_{12}t_{45}t_{36} + t_{23}t_{56}t_{14} + t_{34}t_{16}t_{25} + t_{12}t_{34}t_{56} + t_{16}t_{23}t_{45}$$

6. Casey theorem: let $\alpha, \beta, \gamma, \delta$ be circles. Let $t_{\alpha\beta}$ be the length of a common tangent (either interior of exterior) of α and β. Same for $t_{\alpha\gamma}, t_{\alpha\delta}, t_{\beta\gamma}, t_{\beta\delta}, t_{\gamma\delta}$. Show that

   There exists a way of choosing tangents and signs so that the equation $\pm t_{\alpha\beta}t_{\gamma\delta} \pm t_{\beta\gamma}t_{\delta\alpha} \pm t_{\gamma\alpha}t_{\beta\delta} = 0$ holds, if and only if all 4 circles $\alpha, \beta, \gamma, \delta$ are tangent to another circle.

7. Ptolemy inequality: let $ABCD$ be a quardrilateral, then $AC \cdot BD \leq AB \cdot CD + BC \cdot DA$, with equality happens only if the quadrilateral is cyclic.

8. [McDougall] Let $P_0, P_1, \ldots, P_n$ (with n even) be points on a circle in the counter-clockwise order, and denote $R_k = \prod_{i \neq k} P_k P_i$. Show that

$$\sum_{k=1}^{n} (-1)^{k-1} \frac{1}{R_k} = 0$$

   Conjecture: does it still hold in hyperbolic/spherical geometry?

Chapter 10

Abstract Algebra

Proposition 10.1.

1. $xyz = x + y + z + 2 \Leftrightarrow \frac{1}{1+x} + \frac{1}{1+y} + \frac{1}{1+z} = 1$
2. If x, y, z are non-negative, then $x^2 + y^2 + z^2 + 2xyz = 1$ if and only if there exists $\alpha, \beta, \gamma \geq 0$ with $\alpha + \beta + \gamma = \pi$ such that $x = \cos\alpha$, $y = \cos\beta$, and $z = \cos\gamma$.
3. Suppose $x, y, z > 0$ and $\max\{x, y, z\} \geq 2$, then $x^2+y^2+z^2 = xyz+4$ if and only if there exists $a, b, c > 0$ such that $x = a+1/a, y = b+1/b, z = c+1/c$.

Proposition 10.2 (Triangle Trigonometric Identities)**.** Let $x + y + z = \pi$

1. $\cos^2 x + \cos^2 y + \cos^2 z + 2\cos x \cos y \cos z = 1$
2. $\tan x \tan y \tan z = \tan x + \tan y + \tan z$
3. $\tan\frac{x}{2}\tan\frac{y}{2} + \tan\frac{y}{2}\tan\frac{z}{2} + \tan\frac{z}{2}\tan\frac{x}{2} = 1$

Theorem 10.3 (S.O.S method)**.** Consider $S = S_a(b-c)^2+S_b(c-a)^2+S_c(a-b)^2$, where S_a, S_b, S_c are functions of a, b, c

1. If $S_a, S_b, S_c \geq 0$ then $S \geq 0$.
2. If $a \geq b \geq c$ and $S_b, S_b + S_c, S_b + S_a \geq 0$ then $S \geq 0$.
3. If $a \geq b \geq c$ and $S_a, S_c, S_a + 2S_b, S_c + 2S_b \geq 0$ then $S \geq 0$.
4. If $a \geq b \geq c$ and $S_b, S_c, a^2S_b + b^2S_a \geq 0$ then $S \geq 0$.
5. If $S_a + S_b + S_c \geq 0$ and $S_aS_b + S_bS_c + S_cS_a \geq 0$ then $S \geq 0$.

Other reference: Supersums of Square-Weights (SOS): A Dumbass's Perspective

Theorem 10.4 (Hilbert 17th problem)**.**

Proposition 10.5.

1. $(a^2 + nb^2)(c^2 + nd^2) = (ac + nbd)^2 + n(ad - bc)^2$
2. $(a^2 - nb^2)(c^2 - nd^2) = (ac + nbd)^2 - n(ad + bc)^2$
3. $(n + 1)(a^2 + nb^2) = (a + nb)^2 + n(a - b)^2$
4. Brahmagupta–Fibonacci two-square identity, Euler four-square identity, Degen eight-square identity, and Pfister theorem

10.1 Functional Equation

Proposition 10.6 (Cauchy functional equation)**.** Consider the functional equation $f(x + y) = f(x) + f(y)$

1. If $f : \mathbb{Q} \to \mathbb{C}$, then $f(x) = ax$ for some complex constant a.
2. If $f : \mathbb{R} \to \mathbb{C}$, then $f(x) = ax$ if any of the following condition is satisfied:
   (a) f is continuous at some point.
   (b) f is monotonic on some interval.
   (c) f is bounded on some interval.
   (d) f is Lebesgue measurable.
3. In general, pick a Hamel basis H of $\mathbb{R}$ over the rationals, and $f : H \to \mathbb{C}$ any function, then the linear extension $\overline{f} : \mathbb{R} \to \mathbb{C}$ satisfies the Cauchy functional equation. In fact, the converse also holds, i.e. any function satisfying the equation must be a linear extension of some (complex-valued) map on a Hamel basis.

Proposition 10.7.

1. Quadratic functional equation: $f(x + y) + f(x - y) = 2[f(x) + f(y)]$
2. Jensen functional equation: $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$
3. d'Alembert functional equation: $f(x + y) + f(x - y) = 2[f(x)f(y)]$
4. Sinusoidal system of functional equations:

$$\begin{aligned}(S(x))^2 + (C(x))^2 &= 1\\ S(x + y) &= S(x)C(y) + S(y)C(x)\\ C(x + y) &= C(x)C(y) - S(x)S(y)\end{aligned}$$

   Reference: Defining sine and cosine from functional equations.

10.1.1 Polynomial functional equation

Theorem 10.8 (Lagrange interpolation)**.** Let $(x_i, y_i)_{i=1}^n$ be points in $\mathbb{R}^2$ such that x_i are pairwise distinct, then there exists a unique polynomial $p(x)$ of degree at most n such that $p(x_i) = y_i$. In fact,

$$p(x) = \sum_{i=1}^{n} y_i \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$$

10.2 Irreducibility Criteria

Lemma 10.9 (Gauss primitivity lemma)**.** If R is a unique factorization domain (UFD), then the set of primitive polynomials in $R[x]$ is multiplicatively closed. Moreover, the content of a product is the product of each individual content (for polynomials).

Corollary 10.10 (Gauss irreducibility lemma)**.** If R is a unique factorization domain (UFD), and F is its field of fractions then a non-constant polynomial in $R[x]$ is irreducible (in $R[x]$) if and only if it is irreducible in $F[x]$ and primitive in $R[x]$.

Theorem 10.11 (Content of a polynomial)**.** Let R be a commutative ring. The content $\operatorname{cont}(f)$ of a polynomial f in $R[x]$ is the ideal (in R) generated by the coefficients of f. Then

1. $\operatorname{cont}(fg) \subseteq \operatorname{cont}(f)\operatorname{cont}(g) \subseteq \sqrt{\operatorname{cont}(fg)}$
2. If R is a GCD domain, then $\gcd(\operatorname{cont}(fg)) = \gcd(\operatorname{cont}(f))\gcd(\operatorname{cont}(g))$, where $\gcd(I)$ is the minimal principal ideal containing I (it awalys exists and is unique for finitely generated ideal).

Theorem 10.12 (Eisenstein criterion)**.** If D is an integral domain, and $f(x) = \sum_{i=0}^{n} a_i x^i$ is a polynomial over D. If there is a prime ideal $\mathfrak{p}$ of D such that

1. $a_i \in \mathfrak{p}$ for $i \neq n$
2. $a_n \notin \mathfrak{p}$
3. $a_0 \notin \mathfrak{p}^2$

then $f(x)$ is not the product of 2 non-constant polynomials in $D[x]$. As a result, if $f(x)$ is also primitive, then it is irreducible in $D[x]$.

Theorem 10.13 (Cohn criterion)**.** If a prime number $p \in \mathbb{N}$ is expressed in base b as $p = \sum_{i=0}^{n} a_i b^i$, then $f(x) = \sum_{i=0}^{n} a_i x^i$ is irreducible in $\mathbb{Z}[x]$.

Theorem 10.14 (Perrion criterion)**.** Let $f(x) = \sum_{i=0}^{n} a_i x^i$ be a polynomial over $\mathbb{Z}$ such that $a_0 \neq 0, a_n = 1$ (i.e. $f(x)$ is monic) and either one of the following conditions hold

1. $|a_{n-1}| > 1 + |a_{n-2}| + \cdots + |a_0|$
2. $|a_{n-1}| \geq 1 + |a_{n-2}| + \cdots + |a_0|$ and $f(\pm 1) \neq 0$

Then $f(x)$ is irreducible over $\mathbb{Z}$.

10.3 Constructible numbers

Reference

1. Rotman, Joseph J. (2006), *A First Course in Abstract Algebra with Applications* (3rd ed.), Prentice Hall, ISBN 978-0-13-186267-8
2. Kazarinoff, Nicholas D. (2003) [1970], *Ruler and the Round /Classic Problems in Geometric Constructions*, Dover, ISBN 0-486-42515-0
3. Gleason, Andrew M. (1988), *Angle trisection, the heptagon, and the triskaidecagon*, American Mathematical Monthly, 95 (3): 185–194, doi:10.2307/2323624, MR 0935432. Footnote 8, p. 191. https://web.archive.org/web/20141105205944/http://apollonius.math.nthu.edu.tw/d1/ne01/jyt/li
4. Roger C. Alperin; Robert J. Lang (2009). *One-, Two-, and Multi-Fold Origami Axioms* (PDF). 4OSME. A K Peters.

10.3.1 Galois construction

We say that an (geometric) operation is Galois if, whenever α is constructible from that operation (along with some basic construction), then any of its conjugates must also be constructible.

Example 10.15.

Straightedge

Compass

Trisector

10.3.2 Straightedge and compass

10.3.3 Ruler and bisector

10.3.4 p-sector

10.3.5 Lang Origami

Chapter 11

Calculus

Reference

1. Tran Phuong. *Diamonds in Mathematical inequalities*. Viet Publishing.

Theorem 11.1 (Bertrand paradox)**.** The probability of choosing a chord on a unit circle so that its length is at least the length of a side of an inscribed equilateral triangle (which is $\sqrt{3}/2$) is $1/3$.

Proof. The probability is the ratio $\lambda(C_g)/\lambda(A_g)$ where $C_g = \{(x,y,z,w) : x^2 + y^2 = z^2 + w^2 = 1, (x-z)^2 + (y-w)^2 \geq 3/4\}$, and $A_g = \{(x,y,z,w) : x^2 + y^2 = z^2 + w^2 = 1\}$.
Let $g(\theta,\varphi) = (\cos\theta, \sin\theta, \cos\varphi, \sin\varphi)$ $(\theta,\varphi \in [0,2\pi])$ be a parametrization of A_g. Then $C_g = \{g(\theta,\varphi) : |\theta - \varphi| \in [2\pi/3, 4\pi/3]\}$ (by presenting $(\cos\theta, \sin\theta)$ and $(\cos\varphi, \sin\varphi)$ as 2 points on a unit circle, and find the distance between them).
As $g_\theta = (-\sin\theta, \cos\theta, 0, 0)$ and $g_\varphi = (0, 0, -\sin\varphi, \cos\varphi)$, we get

$$\lambda(A_g) = \int_0^{2\pi}\int_0^{2\pi} \sqrt{\det \begin{bmatrix} g_\theta \cdot g_\theta & g_\theta \cdot g_\varphi \\ g_\varphi \cdot g_\theta & g_\varphi \cdot g_\varphi \end{bmatrix}} d\theta d\varphi = \int_0^{2\pi}\int_0^{2\pi} d\theta d\varphi = 4\pi^2$$

$$\begin{aligned}\lambda(C_g) &= \int_{\{(\theta,\varphi):|\theta-\varphi|\in[2\pi/3,4\pi/3]\}} d\theta d\varphi = 2\int_{\{(\theta,\varphi):\theta-\varphi\in[2\pi/3,4\pi/3]\}} d\theta d\varphi \\ &= 2\cdot\frac{3}{18}\cdot 4\pi^2 = \frac{4}{3}\pi^2\end{aligned}$$

So the probability is $1/3$. □

Theorem 11.2 (Inequalities)**.**

1. Bernoulli inequality:
   (a) $(1+x)^r \geq 1 + rx$ for every integer $r \geq 0$ and real number $x \geq -1$. It is strict if $r \geq 2$ and $x \neq 0$.

(b) $(1+x)^r \geq 1+rx$ for every even integer $r \geq 0$ and real number x.

(c) $(1+x)^r \geq 1+rx$ for every real number $r \geq 1$ and $x \geq -1$. It is strict if $r \notin \{0,1\}$ and $x \neq 0$.

(d) $(1+x)^r \leq 1+rx$ for every real number $0 \leq r \leq 1$ and $x \geq -1$.

2. Cauchy-Schwarz inequality (also called Cauchy–Bunyakovsky-Schwarz inequality): in an inner product space

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle \text{ or}$$
$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

Special cases:

(a) Sedrakyan lemma / Engel form / T2 lemma / Titu lemma: for $u_i, v_i > 0$

$$\sum_{i=1}^{n} \frac{u_i^2}{v_i} \geq \frac{\left(\sum_{i=1}^{n} u_i\right)^2}{\sum_{i=1}^{n} v_i}$$

(b) In $\mathbb{R}^n$

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

3. Chebyshev inequality (also called the Bienaymé–Chebyshev inequality): let

(a) $t > 0$, and $0 < p < \infty$.

(b) (X, Σ, μ) be a measure space.

(c) $f : X \to [-\infty, \infty]$ is an extended real-valued measurable function on X.

(d) $g : (0, \infty) \to [0, \infty]$ is an extended real-valued measurable function and nondecreasing, with $g(t) \neq 0$

then

$$\mu\{x \in X : f(x) \geq t\} \leq \frac{1}{g(t)} \int_X g \circ f \, d\mu$$

Special cases:

(a) When $g(x) = |x|^p$ for $x \geq t$ and 0 for $0 < x < t$

$$\mu\{x \in X : |f(x)| \geq t\} \leq \frac{1}{t^p} \int_X |f|^p \, d\mu$$

(b) Probabilistic version: let X (integrable) be a random variable with finite expected value μ and finite non-zero variance σ^2

$$\Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

(c) Discrete version: let $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ be monotonic sequences

i. If they are either all increasing, or all decreasing:

$$\frac{1}{n}\sum_{i=1}^{n} a_i b_i \geq \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} b_i\right)$$

ii. If one is increasing, and the other is decreasing:

$$\frac{1}{n}\sum_{i=1}^{n} a_i b_i \leq \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)\left(\frac{1}{n}\sum_{i=1}^{n} b_i\right)$$

There are also the continuous version: let $f, g : [a, b] \to \mathbb{R}$ be integrable functions

i. If they are either all increasing, or all decreasing:

$$\frac{1}{b-a}\int_a^b f(x)g(x)dx \geq \left(\frac{1}{b-a}\int_a^b f(x)dx\right)\left(\frac{1}{b-a}\int_a^b g(x)dx\right)$$

ii. If one is increasing, and the other is decreasing:

$$\frac{1}{b-a}\int_a^b f(x)g(x)dx \leq \left(\frac{1}{b-a}\int_a^b f(x)dx\right)\left(\frac{1}{b-a}\int_a^b g(x)dx\right)$$

4. Hölder inequality: let (S, Σ, μ) be a measure space, and $p, q \in [1, \infty]$ be such that $1/p + 1/q = 1$ (i.e. they are Hölder conjugate). Then for all $f \in L^p(S)$ and $g \in L^q(S)$, we have $fg \in L^1(S)$ and

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q$$

Equality occurs if and only if $|f|^p$ and $|g|^q$ are linearly dependent in $L^1(S)$ (i.e. $\alpha|f|^p = \beta|g|^q$ μ-a.e. for some $\alpha^2 + \beta^2 \neq 0$).

Special cases: in $\mathbb{R}^n$ or $\mathbb{C}^n$

$$\sum_{i=1}^{n} |x_i y_i| \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}\left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}$$

5. Karamata inequality / Majorization inequality: let I be an interval and $f : I \to \mathbb{R}$ be a convex function. Suppose $x_k, y_k \in I$ such that $(x_1, \ldots, x_n)$ majorizes $(y_1, \ldots, y_n)$, i.e.

(a) $x_1 \geq x_2 \geq \cdots \geq x_n$ and $y_1 \geq y_2 \geq \cdots \geq y_n$

(b) $x_1 + \cdots + x_k \geq y_1 + \cdots + y_k$ for $1 \leq k < n$.

(c) $x_1 + \cdots + x_n = y_1 + \cdots + y_n$

then $f(x_1) + \cdots + f(x_n) \geq f(y_1) + \cdots + f(y_n)$. If f is strictly convex, then equality holds iff $x_k = y_k$ for all $1 \leq k \leq n$.

6. Jensen inequality: suppose $(\Omega, \mathcal{A}, \mu)$ is a measure space with finite measure (i.e. $\mu(\Omega) < \infty$). If $g : \Omega \to \mathbb{R}$ is integrable, and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex (or at least on the image of Ω under g), then

$$\varphi\left(\frac{1}{\mu(\Omega)} \int_\Omega g d\mu\right) \leq \frac{1}{\mu(\Omega)} \int_\Omega \varphi \circ g d\mu$$

Special cases: if $\lambda_i \geq 0$ such that $\sum_{i=1}^n \lambda_i = 1$, then

$$\varphi\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i \varphi(x_i)$$

7. Minkowski inequality (or simply, the triangle inequality in L^p space): if $f, g \in L^p(S, \Sigma, \mu)$ (with $1 \leq p \leq \infty$)

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Equality happens exactly when f, g are linearly dependent in L^p.

Special cases: for $x_i, y_i \in \mathbb{C}$

$$\left(\sum_{i=1}^n |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |y_i|^p\right)^{1/p}$$

8. Newton inequality: if $a_1, a_2, \ldots, a_n \in \mathbb{R}$, and $e_k(a_1, a_2, \ldots, a_n)$ denotes the k-th elementary symmetric polynomial (i.e. $e_k(a_1, a_2, \ldots, a_n) = \sum_{1 \leq j_1 < j_2 < \cdots < j_k \leq n} a_{j_1} a_{j_2} \cdots a_{j_k}$), then

$$S_{k-1} S_{k+1} \leq S_k^2$$

where $S_k = e_k / \binom{n}{k}$. If all a_i are non-zero, then equality holds iff all a_i are equal.

9. Maclaurin inequality: with the same S_k as above, with all a_i positive, then

$$S_1 \geq S_2^{1/2} \geq \cdots \geq S_n^{1/n}$$

with equality if and only if all a_i are equal.

10. Young inequality: if $a, b \geq 0$, and $p, q > 1$ are Hölder conjugate, then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

Equality occurs exactly when $a^p = b^q$.

11. Muirhead inequality: given $a = (a_i)_{i=1}^n, x = (x_i)_{i=1}^n \in \mathbb{R}^n$ with $x_i > 0$, define

$$[a] = \frac{1}{n!} \sum_{\sigma} \prod_{i=1}^{n} x_{\sigma_i}^{a_i}$$

where the sum is over all permutation of $\{1, 2, \ldots, n\}$.

Then, $[a] \leq [b]$ for all $x_i > 0$ if and only if there exists some doubly stochastic matrix P such that $a = Pb$. Equality happens iff $a = b$ or all x_i are equal to each other.

The condition $a = Pb$ can be stated alternatively as follows: re-labeling indices in a and b so that $a_1 \geq a_2 \geq \cdots a_n$ and $b_1 \geq b_2 \geq \cdots b_n$, then $a = Pb$ for some doubly stochastic matrix P if and only if b majorizes a.

12. Generalized mean inequality: let $-\infty \leq p \leq \infty$, we define the weighted p-mean of $x_i > 0$ with weight $w_i > 0$ as (for $\sum_{i=1}^n w_i = 1$)

    (a) $M_{-\infty}(\vec{x}; \vec{w}) = \min_{i=1}^n w_i x_i$
    (b) $M_{\infty}(\vec{x}; \vec{w}) = \max_{i=1}^n w_i x_i$
    (c) $M_0(\vec{x}; \vec{w}) = \prod_{i=1}^n x_i^{w_i}$
    (d) $M_p(\vec{x}; \vec{w}) = (\sum_{i=1}^n w_i x_i^p)^{1/p}$ for $p \notin \{0, \pm\infty\}$.

    Then for $-\infty \leq p < q \leq \infty$, we have $M_p(\vec{x}; \vec{w}) \leq M_q(\vec{x}; \vec{w})$, with equality if and only if all x_i where $w_i > 0$ are equal. Furthermore, $M_p(\vec{x}; \vec{w})$ is continuous in $p \in [-\infty, \infty]$.

13. Schur inequality: for all non-negative x, y, z, t

$$x^t(x-y)(x-z) + y^t(y-z)(y-x) + z^t(z-x)(z-y) \geq 0$$

with equality if and only if $x = y = z$, or 2 of them are equal and the other is 0. The inequality is also true if t is an even positive integer, and x, y, z are any real numbers.

14. Rearrangement inequality: suppose $x_1 \leq x_2 \leq \cdots \leq x_n$ and $y_1 \leq y_2 \leq \cdots \leq y_n$, then for any permutation σ of $\{1, 2, \ldots, n\}$

$$x_n y_1 + \cdots + x_1 y_n \leq x_{\sigma(1)} y_1 + \cdots + x_{\sigma(n)} y_n \leq x_1 y_1 + \cdots + x_n y_n$$

Theorem 11.3 (Abstract Concreteness Method)**.**

1. Suppose $f(x, y, z)$ is a function that is either monotonic, convex, or concave in variable x on $\mathbb{R}$, then $f(abc, ab + bc + ca, a + b + c)$ achieves extremum (with $ab + bc + ca$ and $a + b + c$ fixed) when 2 of 3 variables a, b, c are equal.

2. Suppose $f(x, y, z)$ is a function that is either monotonic, convex, or concave in variable x on $[0, \infty)$, then $f(abc, ab+bc+ca, a+b+c)$ achieves extremum (with $ab + bc + ca$ and $a + b + c$ fixed) when 2 of 3 variables a, b, c are equal, or one of them is 0.

Chapter 12

Complex Analysis

Theorem 12.1 (Rouché theorem)**.**

Chapter 13

Abstract Vector Space

Reference

1. Paul Richard Halmos. *Finite-Dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, 1974. ISBN: 0387900934.
2. Walter Rudin. *Real and Complex Analysis*. New York City, NY: McGraw Hill Education India, 2015. ISBN: 0070542341.
3. Barry Simon. *Notes on infinite determinants of Hilbert space operators*. In: Advances in Mathematics 24.3 (1977), pp. 244–273. ISSN: 0001-8708. DOI: https://doi.org/10.1016/0001-8708(77) 90057-3. URL: http://www.sciencedirect.com/science/article/pii/0001870877900573.
4. Sergei Winitzki. *Linear Algebra via Exterior Products*. 2010. ISBN: 140929496X.
5. Is there a definition of determinants that does not rely on how they are calculated?.

13.1 Algebraic Trace and Determinant

13.1.1 Inner product

An *involution*, or *conjugation*, $\sigma : F \to F$ on a field F is a field automorphism of order at most 2 (i.e. $\sigma^2 = 1_F$). Associated with an involution, is the modular valuation $N : F \to F, a \to a\sigma(a)$ which satisfies

1. $N(0) = 0$ and $N(1) = N(-1) = 1$.
2. Multiplicative: $N(ab) = N(a)N(b)$.
3. Invariant under involution: $N(a) = N(\sigma(a)) = \sigma(N(a))$.

A σ-inner product on a vector space V over F is a map $\langle \cdot, \cdot \rangle : V \times V \to F$ such that

1. σ-symmetry: for all $x, y \in V$,
$$\langle x, y \rangle = \sigma(\langle y, x \rangle)$$
2. Linearity in 1st argument: for all $x, y, z \in V$ and $a \in F$
$$\langle x + ay, z \rangle = \langle x, z \rangle + a\langle y, z \rangle$$
3. Modular definitness:
   (a) For each $x \in V$, there exists some $a \in F$ such that $\langle x, x \rangle = N(a)$.
   (b) If $\langle x, x \rangle = 0$, then $x = 0$.

A vector space with such inner product is called a σ-inner product space (or just inner product space given the context)

Proposition 13.1.

1. The orthogonality relation is symmetric: $\langle x, y \rangle = 0$ iff $\langle y, x \rangle = 0$.
2. Non-degeneracy: if $\langle x, y \rangle = 0$ for all $y \in V$ then $x = 0$. Similarly, if $\langle x, y \rangle = 0$ for all $x \in V$ then $y = 0$.
3. σ-linearity in 2nd argument: for all $x, y, z \in V$ and $a \in F$
$$\langle x, y + az \rangle = \langle x, y \rangle + \sigma(a)\langle x, z \rangle$$

Example 13.2.

1. The standard Euclidean inner is the inner product when σ is the identity on $\mathbb{R}$. In fact, it is the only field automorphism on $\mathbb{R}$.
2. The Hermitian inner product is the inner product when $\sigma(z) = \overline{z}$. Along with the identity map, these 2 involutions comprise all field automorphisms on $\mathbb{C}$ such that $\mathbb{R}$ is fixed.

Theorem 13.3 (Riesz representation theorem)**.** Given V a finite-dimensional σ-inner product space, and V^* its (algebraic) dual, then for each functional f in V^*, there exists a unique α in V such that $f(v) = \langle v, \alpha \rangle$ for all $v \in V$.

Proof. Denote $\Phi_\alpha(v) = \langle v, \alpha \rangle$, then by non-degeneracy
$$\Phi_\alpha = \Phi_\beta \leftrightarrow \langle v, \alpha - \beta \rangle = 0, \forall v \in V \leftrightarrow \alpha = \beta$$
Thus, the right-multiplication map $\Phi : V \to V^*$ is injective, which is also bijective since V is finite-dimensional. Note that Φ is also σ-linear, that is
$$\Phi_{x+cy} = \Phi_x + \sigma(c)\Phi_y$$
□

13.1.2 Trace

Let $\mathcal{L}(V,W)$ be the space of all linear maps from V to W, and $V^* = \mathcal{L}(V,F)$ be the dual space of V.

Theorem 13.4. The map $\eta : V^* \otimes W \to \mathcal{L}(V,W)$, where $\eta(\alpha \otimes w)(v) = \alpha(v)w$ (and extends linearly), is a canonical embedding.

The image η, sometimes identified with $V^* \otimes W$, is called *space of trace class maps* (from V to W), and denoted as $\mathrm{Tr}(V,W) = \eta(V^* \otimes W)$. When $V = W$, and $T = \eta\left(\sum_i c_i \alpha_i \otimes v_i\right)$ then the *trace* of T is defined as

$$\mathrm{tr}(T) = \sum_i c_i \alpha_i(v_i)$$

Clearly the trace is a linear map from $\mathrm{Tr}(V) = \mathrm{Tr}(V,V)$ to F.

Proposition 13.5 (Properties of trace)**.**

1. Linearity: if T, S are trace-class and $c \in F$, then $T + cS$ is also trace-class. Additionally, when $V = W$, we have
$$\mathrm{tr}(T + cS) = \mathrm{tr}(T) + c\,\mathrm{tr}(S)$$
2. Closed under composition: if $T : V \to W$ and $S : U \to V$ are trace-class, then $T \circ S : U \to W$ is also trace-class. Hence, $\mathrm{Tr}(V)$ is a rng (a ring without multiplicative identity).
3. Invariant under cyclic composition, if $T : V \to W$ and $S : W \to V$ are trace-class, then
$$\mathrm{tr}(TS) = \mathrm{tr}(ST)$$
4. Identity as trace-class: the identity is a trace-class operator iff V is finite-dimensional. Therefore, the vector spaces (including infinite-dimensional ones) with trace-class maps as morphisms will not form a category.
5. Transpose: the dual T^* is trace-class if and only if T is trace-class. If so, we also get $\mathrm{tr}(T) = \mathrm{tr}(T^*)$.
6. Adjoint [???]
7. Direct sum: the direct sum of trace-class maps is again trace-class, and $\mathrm{tr}(T \oplus S) = \mathrm{tr}(T) + \mathrm{tr}(S)$
8. Tensor product: the tensor product of trace-class maps is also trace-class, and $\mathrm{tr}(T \otimes S) = \mathrm{tr}(T)\,\mathrm{tr}(S)$

Theorem 13.6 (Characteriztion of Trace Class Operators)**.** An operator $T : V \to W$ is trace class if and only if it has finite rank, i.e. its image is finite-dimensional.

Corollary 13.7. All linear maps $V \to W$ are trace class if and only if either V or W has finite dimension.

Corollary 13.8 (Computation of trace)**.** Let $T : V \to V$ be a linear operator with finite rank, and suppose $\beta = \{v_1, v_2, \ldots, v_n\}$ is a basis of $T(V)$ with β^* is an extension of β to a basis of the whole V. If f_i is the coefficient of v_i in $T(v_i)$ then

$$\operatorname{tr}(T) = \sum_{i=1}^{n} f_i$$

13.1.3 Determinant

Let V^I be the vector space of all functions $f : I \to V$. A map $\phi : V^I \to F$ is called

1. Multilinear if for all $k \in I$ and $(\alpha_i)_{i\in I}, (\beta_i)_{i\in I} \in V^I$ and $c \in F$ such that $\alpha_i = \beta_i$ for all $i \neq k$, we have $\phi(\gamma) = \phi(\alpha) + c\phi(\beta)$ where $\gamma_i = \alpha_i$ for $i \neq k$ but $\gamma_k = \alpha_k + c\beta_k$.

2. Alternating if for any $(\alpha_i)_{i\in I} \in V^I$ such that $\alpha_n = \alpha_m$ for some $n \neq m \in I$, then $\phi(\alpha) = 0$.

3. Symmetric if for any transpose map $\tau : I \to J$, $\phi(\alpha) = \phi(\alpha \circ \tau)$

4. Anti-symmetric if for any transpose map $\tau : I \to J$, $\phi(\alpha) = -\phi(\alpha \circ \tau)$

The space of all alternating multilinear maps (of arity I) is denoted as $\Lambda^I V$.

Proposition 13.9. If $(\alpha_i)_{i\in I}$ are dependent vectors in V, then $\phi(\alpha) = 0$ for all alternating multilinear map. As a result, $\Lambda^I V = 0$ if the cardinality of I is strictly greater than the dimension of V.

Proposition 13.10 (Independent of indices)**.** If I is equinumerous to J, then $\Lambda^I V$ is isomorphic to $\Lambda^J V$.

Theorem 13.11. If $\dim V = |I|$, then $\Lambda^I V$ is one-dimensional.

Theorem 13.12.

1. Transpose: T is determinant-class iff its dual is determinant-class, and $\det T = \det T^*$.

2. Adjoint: T is determinant-class iff its adjoint is determinant-class, and $\det T = \det T^\sigma$.

3. Direct sum
4. Tensor product
5. Fredholm property: if $T : V \to V$ is trace-class, then $1_V + T$ is determinant-class.

Proposition 13.13.

1. Trace-class operators are stable under conjugation
   (a) Given $T : V \to W$ trace-class and $U : V \to W$ an isomorphism, UTU^{-1} is also trace-class
   (b) $\det(1_V + T) = \det(1_W + UTU^{-1})$
2. Sylvester determinant theorem: if $T : V \to W$ and $S : W \to V$ are trace-class, then
$$\det(1_W + ST) = \det(1_V + ST)$$

13.2 Volumetric space

13.2.1 Simplicial volume on a vector space

A map $d_n : V^{n+1} \to \mathbb{R}$ is an n-volume of a vector space V if

1. Non-negative: $d_n(x_0, x_1, \ldots, x_n) \geq 0$ for all $x_i \in V$.
2. Symmetry: $d_n(x_{\sigma(0)}, x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = d_n(x_0, x_1, \ldots, x_n)$ for any permutation $\sigma \in S_{n+1}$.
3. Triangle inequality: if Δ denotes the collection of all subsets of $\{0, 1, 2, \ldots, n+1\}$ of size $n+1$, and $P \in \Delta$, then
$$\sum_{I \neq P, I \in \Delta} d_n(x_i)_{i \in I} \geq d_n(x_i)_{i \in P}$$
4. Translational invariance: $d_n(x_0+v, x_1+v, \ldots, x_n+v) = d_n(x_0, x_1, \ldots, x_n)$ for all $x_i, v \in V$.
5. Homogeneity: $d_n(0, \alpha x_1, x_2, \ldots, x_n) = |\alpha| d_n(0, x_1, x_2, \ldots, x_n)$ for all $x_i \in V$ and $\alpha \in F$.

A family of n-volumes (for $m \geq n \geq 1$) on V is geometrically compatible if

1. $d_1(x, y) = 0$ if and only if $x = y$.
2. $d_n(x_0, x_1, \ldots, x_n) = 0$ if and only if for some $P \in \Delta$, $\sum_{I \neq P} d_{n-1}(x_i)_{i \in I} = d_{n-1}(x_i)_{i \in P}$.

Proposition 13.14. If $x_0, x_1, \ldots, x_n$ are linearly dependent, then $d_n(x_0, x_1, \ldots, x_n) = 0$.

13.2.2 Gramian volume

Given V a Hilbert space, the Gramian volume of $x_1, \ldots, x_n$ is

$$\frac{1}{n!}\sqrt{\det\begin{bmatrix}\langle x_1, x_1\rangle & \langle x_1, x_2\rangle & \cdots & \langle x_1, x_n\rangle \\ \langle x_2, x_1\rangle & \langle x_2, x_2\rangle & \cdots & \langle x_2, x_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle x_n, x_1\rangle & \langle x_n, x_2\rangle & \cdots & \langle x_n, x_n\rangle\end{bmatrix}}$$

13.2.3 Cayley-Menger volume

In a metric space M, let $d_{ij} = d(x_i, x_j)$, then the Cayley-Menger volume is

$$\sqrt{\frac{1}{(n!)^2 2^n}\det\begin{bmatrix}2d_{01}^2 & d_{01}^2 + d_{02}^2 - d_{12}^2 & \cdots & d_{01}^2 + d_{0n}^2 - d_{1n}^2 \\ d_{01}^2 + d_{02}^2 - d_{12}^2 & 2d_{02}^2 & \cdots & d_{02}^2 + d_{0n}^2 - d_{2n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ d_{01}^2 + d_{0n}^2 - d_{1n}^2 & d_{02}^2 + d_{0n}^2 - d_{2n}^2 & \cdots & 2d_{0n}^2\end{bmatrix}}$$

$$= \sqrt{\frac{(-1)^{n+1}}{(n!)^2 2^n}\det\begin{bmatrix}0 & d_{01}^2 & d_{02}^2 & \cdots & d_{0n}^2 & 1 \\ d_{01}^2 & 0 & d_{12}^2 & \cdots & d_{1n}^2 & 1 \\ d_{02}^2 & d_{12}^2 & 0 & \cdots & d_{2n}^2 & 1 \\ \vdots & \vdots & \ddots & \vdots & & \\ d_{0n}^2 & d_{1n}^2 & d_{2n}^2 & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0\end{bmatrix}}$$

Chapter 14

Algebraic Number Theory

Theorem 14.1 (Pell equation)**.** Let $x^2 - ny^2 = 1$ be a Pell equation (with n fixed and square-free)

1. If $[a_0; a_1, a_2, \cdots]$ is the continued fraction of $\sqrt{n}$, then some convergent x_1/y_1 will satisfy the equation. The smallest (x_1, y_1) with respect to the first coordinate is called the fundamental solution.

2. All remaning solution can be found by taking the power of the fundamental solution: if $x_k + y_k\sqrt{n} = (x_1 + y_1\sqrt{n})^k$, then (x_k, y_k) is a solution. Vice versa, if (a, b) is a solution, then $a+b\sqrt{n} = (x_1+y_1\sqrt{n})^k$ for some positive integer k. A recurrence relation for x_k and y_k is described below

$$\begin{aligned} x_{k+1} &= x_1x_k + ny_1y_k \\ y_{k+1} &= x_1y_k + y_1x_k \end{aligned}$$

Lemma 14.2 (Bézout identity)**.** If $\gcd(a, b) = d$ then $\{ax+by : x, y \in \mathbb{Z}\} = d\mathbb{Z}$.

Lemma 14.3 (Hensel lemma)**.** Let $\mathfrak{m}$ be a maximal ideal of a commutative ring R, $h(x) = a_0x^n + \cdots + a_{n-1}x + a_n$ be a polynomial over R such that the leading coefficient a_0 is not in $\mathfrak{m}$. Then every factorization of $h(x)$ modulo $\mathfrak{m}$ (note that $(R/\mathfrak{m})[x]$ is a PID) can be lifted uniquely into a factorization modulo $\mathfrak{m}^k$ for each $k \geq 1$.

More precisely, if $h(x) \equiv a_0f(x)g(x) \pmod{\mathfrak{m}}$ where f, g are monic and coprime modulo $\mathfrak{m}$, then there exists unique monic polynomials $f_k, g_k \pmod{\mathfrak{m}^k}$ for each $k \geq 1$ such that

$$\begin{aligned} h(x) &\equiv a_0f_k(x)g_k(x) \pmod{\mathfrak{m}^k} \\ f_k(x) &\equiv f(x) \pmod{\mathfrak{m}} \\ g_k(x) &\equiv g(x) \pmod{\mathfrak{m}} \end{aligned}$$

Lemma 14.4 (Lifting the exponent - LTE)**.** Let x, y be integers, n, p be positive integers, where p is a prime such that it does not divide x and y. Then

1. If p is odd:
   (a) If $p \mid x - y$: $\nu_p(x^n - y^n) = \nu_p(x - y) + \nu_p(n)$
   (b) If n is odd and $p \mid x + y$: $\nu_p(x^n + y^n) = \nu_p(x + y) + \nu_p(n)$
2. If $p = 2$:
   (a) If $4 \mid x - y$: $\nu_2(x^n - y^n) = \nu_2(x - y) + \nu_2(n)$
   (b) If $2 \mid x - y$ and n is even: $\nu_2(x^n - y^n) = \nu_2(x - y) + \nu_2(x + y) + \nu_2(n) - 1$

Theorem 14.5 (Chinese Remainder Theorem)**.** If $I_1, I_2, \ldots, I_n$ are pairwise coprime ideals of a commutative ring R (i.e. the sum of ideals is equal to R), then

$$R/\bigcap_{k=1}^{n} I_k \cong \prod_{k=1}^{n} R/I_k$$

Theorem 14.6 (Sylvester Coinage Theorem)**.** Let a, b be relatively prime positive integers, then

1. $(a - 1)(b - 1) - 1$ is largest number that is not a sum of nonnegative multiplies of a and b.
2. There are $(a - 1)(b - 1)/2$ of such numbers.

Theorem 14.7 (Wilson Theorem)**.** n is prime if and only if $(n - 1)! \equiv -1 \pmod{n}$.

Theorem 14.8 (Fermat 2-square theorem)**.** An odd prime p is a sum of 2 squares if and only if $p \equiv 1 \pmod 4$. In the language of algebra, all prime $p \equiv 1 \pmod 4$ is not prime as Gaussian integers $\mathbb{Z}[i]$.

Proof. If $p = x^2 + y^2$, then we can check the cases of x and y modulo 4 to show that $p \equiv 1 \pmod 4$. As for the converse, we first have $(p - 1)! \equiv -1 \pmod p$ by Wilson theorem, or

$$\left[\left(\frac{p-1}{2}\right)!\right]^2 (-1)^{(p-1)/2} \equiv -1 \pmod p$$

Since $p \equiv 1 \pmod 4$, $\left[\left(\frac{p-1}{2}\right)!\right]^2 \equiv -1 \pmod p$. In other words, there exists some m such that $m^2 + 1$ is divisible by p. Consider the set $\{am + b : 0 \le a, b \le \lceil\sqrt{p}\rceil - 1\}$: it has $\lceil\sqrt{p}\rceil > p$ elements, so by pigeonhole principle, there exists some $a_1, a_2, b_1, b_2 \in [0, \lceil\sqrt{p}\rceil - 1]$ such that

$$a_1 m + b_1 \equiv a_2 m + b_2 \pmod p, \text{ or}$$
$$(a_1 - a_2)^2 m^2 \equiv (b_1 - b_2)^2 \pmod p$$

Since $m^2 \equiv -1 \pmod{p}$, we get $(a_1 - a_2)^2 + (b_1 - b_2)^2 \equiv 0 \pmod{p}$. On the other hand, $(a_1 - a_2)^2 + (b_1 - b_2)^2 \leq 2(\lceil \sqrt{p} \rceil - 1)^2 < 2p$, so the only possible values for the expression $(a_1 - a_2)^2 + (b_1 - b_2)^2$ is p (as it is divisible by p). □

14.1 Euler totient theorem

In March 2015, when the author was in preparation for 2015 Vietnam Team Selection Test (TST) in April, one of his coaches and teachers (Tran Nam Dung) gave this problem as an exercise: let $(f_n)_{n \in \mathbb{N}}$ be a sequence defined recursively as

1. $f_1 = a, f_2 = a^2 - 2b$ where a, b are non-negative integers.
2. $f_{n+2} = af_{n+1} - bf_n$

Show that

1. For any prime p and any non-negative integer m, $f_{p^{m+1}} \equiv f_{p^m} \pmod{p^{m+1}}$
2. Does $f_{n+\varphi(n)} \equiv f_n \pmod{n + \varphi(n)}$ for all n?

2nd part is fairly easy, just choose $n = 4$ with specific a and b as a counterexample. Even if one change the modulus $n + \varphi(n)$ to n, $n = 4$ can still provide a counterexample. Even still, the correct form $f_{n-\varphi(n)} \equiv f_n \pmod{n - \varphi(n)}$ has $n = 6$ as a counterexample for some a and b.

We will show an even stronger statement

Theorem 14.9. Let K be an (algebraic) number field (i.e. finite extension of $\mathbb{Q}$) and $\mathcal{O}_K$ be its ring of integers. If $\alpha_1, \alpha_2, \ldots, \alpha_k$ are roots (multiplicity counted) of the polynomial equation

$$x^k = c_1 x^{k-1} + c_2 x^{k-2} + \cdots + c_{k-1} x + c_k$$

over K (i.e. $c_i \in K$). Let $f_n = \sum_{i=1}^{k} \alpha_i^n$, then

1. $f_n \in K$ for all $n \in \mathbb{N}$.
2. Recall the norm $N(I)$ of an ideal is the cardinality of its residue $\mathcal{O}_K/I$. Show that for a prime ideal $\mathfrak{p}$ and $t \geq 0$ such that $v_{\mathfrak{p}}(c_i) \geq 0$ (for $1 \leq i \leq k$), we get
$$f_{N(\mathfrak{p})^{t+1}} \equiv f_{N(\mathfrak{p})^t} \pmod{\mathfrak{p}^{t+1}}$$

To prove, we introduce the following lemmas

Lemma 14.10 (Newton identities on symmetric polynomial)**.** Define

1. The elementary polynomials $e_k(x_1, x_2, \ldots, x_n)$

$$e_k(x_1, x_2, \ldots, x_n) = \begin{cases} 0 & \text{if } k < 0 \\ 1 & \text{if } k = 0 \\ \sum_{1 \le i_1 < i_2 < \cdots i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k} & \text{if } 1 \le k \le n \\ 0 & \text{if } k > n \end{cases}$$

2. The power sum polynomials $p_m(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i^m$.

Then for integer $m \ge 1$

$$p_m = m \sum_{\substack{r_1+2r_2+\cdots+nr_n=m \\ r_1, r_2, \ldots, r_n \ge 0}} (-1)^m \frac{(r_1 + r_2 + \cdots + r_n - 1)!}{r_1! r_2! \cdots r_n!} \prod_{i=1}^{n} (-e_i)^{r_i}$$

Proof. It could be proven using induction, but we will use the generating function method (Example 11, chapter 8, pp.169-172 in [Titu Andreescu and Gabriel Dospinescu. Problems from the book. Plano, TX: XYZ Press, 2008. ISBN: 978-0-9799269-0-7]).

Fix x_i and write $e_i = e_i(x_1, x_2, \ldots, x_n), p_m = p_m(x_1, x_2, \ldots, x_n)$. By Viète formula

$$P(z) = \frac{1}{\sum_{i=0}^{n} (-1)^i e_i z^i} = \prod_{i=1}^{n} \frac{1}{1 - x_i z}$$

We then compute its logarithm in 2 different ways. Firstly,

$$\begin{aligned} \ln P(z) &= -\sum_{i=1}^{n} \ln(1 - x_i z) \\ &= \sum_{i=1}^{n} \sum_{k=1}^{\infty} \frac{(x_i z)^k}{k} \\ &= \sum_{k=1}^{\infty} \frac{z^k}{k} \sum_{i=1}^{n} x_i^k \\ &= \sum_{k=1}^{\infty} \frac{p_k}{k} z^k \end{aligned}$$

On the other hand, by the multinomial formula

$$\begin{aligned}
\ln P(z) &= -\ln \sum_{i=0}^{n} (-1)^i e_i z^i \\
&= -\ln \left(1 - \sum_{i=1}^{n} (-1)^{i+1} e_i z^i\right) \\
&= \sum_{m=1}^{\infty} \frac{1}{m} \left(\sum_{i=1}^{n} (-1)^{i+1} e_i z^i\right)^m \\
&= \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\substack{r_1+r_2+\cdots+r_n=m \\ r_1,r_2,\ldots,r_n\geq 0}} \binom{m}{r_1, r_2, \ldots, r_n} \prod_{i=1}^{n} \left[(-1)^{i+1} e_i z^i\right]^{r_i} \\
&= \sum_{m=1}^{\infty} \sum_{\substack{r_1+r_2+\cdots+r_n=m \\ r_1,r_2,\ldots,r_n\geq 0}} \frac{(m-1)!}{r_1! r_2! \cdots r_n!} \prod_{i=1}^{n} (-1)^{ir_i} (-e_i)^{r_i} z^{ir_i} \\
&= \sum_{r_1,r_2,\ldots,r_n\geq 0} \frac{(r_1+r_2+\cdots+r_n-1)!}{r_1! r_2! \cdots r_n!} \prod_{i=1}^{n} (-1)^{ir_i} (-e_i)^{r_i} z^{ir_i}
\end{aligned}$$

Comparing the coefficient of z^k, we get

$$\begin{aligned}
\frac{p_k}{k} &= \sum_{\substack{r_1+2r_2+\cdots+nr_n=m \\ r_1,r_2,\ldots,r_n\geq 0}} \frac{(r_1+r_2+\cdots+r_n-1)!}{r_1! r_2! \cdots r_n!} \prod_{i=1}^{n} (-1)^{ir_i} (-e_i)^{r_i} \\
&= \sum_{\substack{r_1+2r_2+\cdots+nr_n=m \\ r_1,r_2,\ldots,r_n\geq 0}} (-1)^m \frac{(r_1+r_2+\cdots+r_n-1)!}{r_1! r_2! \cdots r_n!} \prod_{i=1}^{n} (-e_i)^{r_i}
\end{aligned}$$

□

Lemma 14.11 (Euler totient theorem)**.** For any $\omega \in K, t \in \mathbb{N}$ and prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ such that $v_{\mathfrak{p}}(\omega) \geq 0$, we have

$$\omega^{N(\mathfrak{p})^{t+1}} \equiv \omega^{N(\mathfrak{p})^t} \pmod{\mathfrak{p}^{t+1}}$$

Proof. We first show for $\omega \in \mathcal{O}_K \setminus \mathfrak{p}$, then $\omega \in \mathcal{O}_K$, and finally $\omega \in K$.

*Assuming $\omega \in \mathcal{O}_K \setminus \mathfrak{p}$, we consider the multiplication map by ω on the multiplicative group of $\mathcal{O}_K/\mathfrak{p}^{t+1}$

$$f(x + \mathfrak{p}^{t+1}) = \omega x + \mathfrak{p}^{t+1}$$

1. Well-defined: we prove that $\omega + \mathfrak{p}^{t+1}$ is also a unit of $\mathcal{O}_K/\mathfrak{p}^{t+1}$ (and so for). Since $\omega \notin \mathfrak{p}$, $\omega + \mathfrak{p}$ is non-zero in $\mathcal{O}_K/\mathfrak{p}$. But $\mathfrak{p}$ is maximal (as $\mathcal{O}_K$ is a Dedekind domain), so $\mathcal{O}_K/\mathfrak{p}$ is a field, meaning that $\omega + \mathfrak{p}$ is invertible in $\mathcal{O}_K/\mathfrak{p}$.

   Let $\gamma + \mathfrak{p}$ be its (multiplicative) inverse, i.e. $\omega\gamma + \pi = 1$ for some $\pi \in \mathfrak{p}$. Now define the following sequences

$$\begin{cases} y_1 = \gamma, & p_1 = \pi \\ y_n = \pi y_{n-1} + \gamma, & p_n = \pi p_{n-1} \end{cases}$$

   Using induction, we have $p_n = \pi^n$ and

$$\omega y_n + p_n = \omega(\pi y_{n-1} + \gamma) + \pi p_{n-1} = \omega\gamma + \pi(\omega y_{n-1} + p_{n-1}) = \omega\gamma + \pi = 1$$

   In other words, $\omega y_n + \pi^n = 1$, so $\omega + \mathfrak{p}^n$ is invertible in $\mathcal{O}_K/\mathfrak{p}^n$ for each $n \in \mathbb{N}$.

2. Bijective: since $\omega + \mathfrak{p}^{t+1}$ is invertible in $\mathcal{O}_K/\mathfrak{p}^{t+1}$, there exists some $\gamma \in \mathcal{O}_K$ such that $\omega\gamma \equiv 1 \pmod{\mathfrak{p}^{t+1}}$. Then the inverse of f is given by

$$f^{-1}(x + \mathfrak{p}^{t+1}) = \gamma x + \mathfrak{p}^{t+1}$$

   Indeed, $f \circ f^{-1}(x + \mathfrak{p}^{t+1}) = \omega(\gamma x) + \mathfrak{p}^{t+1} = x + \mathfrak{p}^{t+1}$ and similarly, $f^{-1} \circ f(x + \mathfrak{p}^{t+1}) = \gamma(\omega x) + \mathfrak{p}^{t+1} = x + \mathfrak{p}^{t+1}$.

On the other hand, $\mathcal{O}_K/I$ is finite for any ideal I, so $(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times$ is also finite, and since $f : (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times \to (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times$ is bijective, we get

$$\prod_{x+\mathfrak{p}^{t+1} \in (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times} x + \mathfrak{p}^{t+1} = \prod_{x+\mathfrak{p}^{t+1} \in (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times} \omega x + \mathfrak{p}^{t+1}$$

$$\omega^{|(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times|} r \equiv r \pmod{\mathfrak{p}^{t+1}}$$

$$\omega^{|(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times|} \equiv 1 \pmod{\mathfrak{p}^{t+1}}$$

where r is some number in $\mathcal{O}_K$ such that $r + \mathfrak{p}^{t+1} = \prod_{x+\mathfrak{p}^{t+1} \in (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times} x + \mathfrak{p}^{t+1} \in (\mathcal{O}_K/\mathfrak{p}^{t+1})^\times$ (since the multiplicative group ... is a group).

To conclude the case, we need to compute the cardinality of $(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times$.

1. If $\theta + \mathfrak{p}^{t+1}$ is not an invertible element of $\mathcal{O}_K/\mathfrak{p}^{t+1}$, then θ must be in $\mathfrak{p}$. Otherwise, just as we have shown above for ω, $\theta + \mathfrak{p}^{t+1}$ is also an invertible element, a contradiction. Thus, $\theta + \mathfrak{p}^{t+1}$ must belong to $\mathfrak{p}/\mathfrak{p}^{t+1}$.

2. Vice versa, $\theta + \mathfrak{p}^{t+1} \in \mathfrak{p}/\mathfrak{p}^{t+1}$, then $\theta \in \mathfrak{p}$ by definition. But then $\theta^{t+1} \in \mathfrak{p}^{t+1}$, i.e. $\theta + \mathfrak{p}^{t+1}$ is a nilpotent element of $\mathcal{O}_K/\mathfrak{p}^{t+1}$. In particular, it cannot be in $(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times$.

3. Hence, $\mathcal{O}_K/\mathfrak{p}^{t+1}$ is a disjoint union of its multiplicative group, and $\mathfrak{p}/\mathfrak{p}^{t+1}$.

$$N(\mathfrak{p})^{t+1} = |\mathcal{O}_K/\mathfrak{p}^{t+1}| = |(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times| + |\mathfrak{p}/\mathfrak{p}^{t+1}|$$

On the other hand, by the 3rd isomorphism theorem for ring $(\mathcal{O}_K/\mathfrak{p}^{t+1})/(\mathfrak{p}/\mathfrak{p}^{t+1}) \cong \mathcal{O}_K/\mathfrak{p}$, so

$$|\mathfrak{p}/\mathfrak{p}^{t+1}| = \frac{|(\mathcal{O}_K/\mathfrak{p}^{t+1})|}{|(\mathcal{O}_K/\mathfrak{p})|} = N(\mathfrak{p})^t$$

and

$$|(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times| = N(\mathfrak{p})^{t+1} - N(\mathfrak{p})^t$$

Rewriting the previous congruence, we obtain the following

$$\omega^{N(\mathfrak{p})^{t+1}-N(\mathfrak{p})^t} \equiv 1 \pmod{\mathfrak{p}^{t+1}}$$
$$\omega^{N(\mathfrak{p})^{t+1}} \equiv \omega^{N(\mathfrak{p})^t} \pmod{\mathfrak{p}^{t+1}}$$

*For the next case, $\omega \in \mathcal{O}_K$, since we already show the identity for $\omega \in \mathcal{O}_K \setminus \mathfrak{p}$, we only need to show the same holds for $\omega \in \mathfrak{p}$, but this is straightforward: we already have $\omega^n \equiv 0 \pmod{\mathfrak{p}^{t+1}}$ for any $n \geq t+1$. On the other hand, $N(\mathfrak{p}) \geq 2$ (since it is the cardinality of a finite field), so $N(\mathfrak{p})^t \geq 2^t \geq t+1$ $(t \geq 0)$. In other words,

$$\omega^{N(\mathfrak{p})^{t+1}} \equiv \omega^{N(\mathfrak{p})^t} \equiv 0 \pmod{\mathfrak{p}^{t+1}}$$

*For the final case, $\omega \in K$, since K coincides with $\mathcal{O}_K$'s field of fractions, we can write $\omega = \eta/\delta$ for some number η, δ in $\mathcal{O}_K$ (possibly have some common factor). If $v_\mathfrak{p}(\omega) \geq 1$, then $v_\mathfrak{p}(\omega^{N(\mathfrak{p})^t}) = N(\mathfrak{p})^t \geq 2^t \geq t+1$, so

$$\omega^{N(\mathfrak{p})^{t+1}} \equiv \omega^{N(\mathfrak{p})^t} \equiv 0 \pmod{\mathfrak{p}^{t+1}}$$

Otherwise, suppose $v_\mathfrak{p}(\eta) = v_\mathfrak{p}(\delta) = s$ for some non-negative integer s. Just like the case $\omega \in \mathcal{O}_K \setminus \mathfrak{p}$, we will show that

$$\omega^{N(\mathfrak{p})^{t+1}-N(\mathfrak{p})^t} \equiv 1 \pmod{\mathfrak{p}^{t+1}}$$

which will lead to Q.E.D.

Denote $\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1}) = N(\mathfrak{p})^{t+1} - N(\mathfrak{p})^t = |(\mathcal{O}_K/\mathfrak{p}^{t+1})^\times|$. By assumption, we can factorize the principal ideals $\eta\mathcal{O}_K$ and $\delta\mathcal{O}_K$ into

$$\eta\mathcal{O}_K = \mathfrak{p}^s\mathfrak{n}$$
$$\delta\mathcal{O}_K = \mathfrak{p}^s\mathfrak{d}$$

such that $\mathfrak{n}, \mathfrak{d}$ are not divisible by $\mathfrak{p}$ (in particular, $\mathfrak{n} \cap \mathfrak{p} = \mathfrak{d} \cap \mathfrak{p} = \{0\}$). Since these are also set equalities, if we fix some $\rho \in \mathfrak{p} \setminus \mathfrak{p}^2, \lambda \in \mathfrak{n}, \kappa \in \mathfrak{d}$ (all non-zero), then there exists some $r_\eta, r_\delta \in \mathcal{O}_K$ (also non-zero) such that

$$\eta r_\eta = \rho^s\lambda$$
$$\delta r_\delta = \rho^s\kappa$$

Taking valuation at $\mathfrak{p}$, we notice that $r_\eta, r_\delta \notin \mathfrak{p}$. Now consider

$$\begin{aligned} v_p\left((\eta r_\eta)^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} - (\delta r_\delta)^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})}\right) &= v_p\left(\rho^{s\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})}\left[\lambda^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} - \kappa^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})}\right]\right) \\ &= s\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1}) + v_p\left(\lambda^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} - \kappa^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})(\mathfrak{p}^{t+1})}\right) \\ &\geq s\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1}) + (t+1) \end{aligned}$$

The last inequality is due to $\lambda^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} \equiv \kappa^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} \equiv 1 \pmod{\mathfrak{p}^{t+1}}$ (λ, κ are non-zero, so they do not belong to $\mathfrak{p}$). Rewriting the above inequality as the following congruence

$$\begin{aligned} (\eta r_\eta)^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} &\equiv (\delta r_\delta)^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} && \pmod{\mathfrak{p}^{s\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})+(t+1)}} \\ \omega^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} r_\eta^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} &\equiv r_\delta^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} && \pmod{\mathfrak{p}^{t+1}} \\ \omega^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} &\equiv 1 && \pmod{\mathfrak{p}^{t+1}} \end{aligned}$$

The second congruence is due to $v_\mathfrak{p}(\delta^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})}) = v_\mathfrak{p}(\delta)\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1}) = s\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})$, while the last congruence is due to $r_\eta^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} \equiv r_\delta^{\varphi_{\mathcal{O}_K}(\mathfrak{p}^{t+1})} \equiv 1 \pmod{\mathfrak{p}^{t+1}}$ (since $r_\eta, r_\delta \notin \mathfrak{p}$). □

Corollary 14.12. Let ω be some number in K and I be some (integral) ideal of $\mathcal{O}_K$ such that $v_\mathfrak{p}(\omega) = 0$ whenever $v_\mathfrak{p}(I) > 0$ for a given prime ideal $\mathfrak{p}$. Prove that

$$\omega^{\varphi_{\mathcal{O}_K}(I)} \equiv 1 \pmod{I}$$

Recall that $\varphi_{\mathcal{O}_K}(I)$ is the cardinality of the multiplicative group of $\mathcal{O}_K/I$.

Proof. Since $\mathcal{O}_K$ is a Dedekind domain, we can factorize I into product of power of prime ideals $I = \prod_{k=1}^n \mathfrak{p}_k^{t_k}$. If $\mathfrak{m}$ is the sum of $\mathfrak{p}_i^{t_i}$ and $\mathfrak{p}_j^{t_j}$ (for $i \neq j$), then it contains $\mathfrak{p}_i^{t_i}$ and $\mathfrak{p}_j^{t_j}$. Since $\mathfrak{p}_i, \mathfrak{p}_j$ are prime, $\mathfrak{m}$ contains the prime ideals themselves. But a prime ideal in Dedekind domain is always maximal, so $\mathfrak{m}$ is the whole ring $\mathcal{O}_K$. Thus, we can apply the Chinese remainder theorem to get

$$\begin{aligned} \mathcal{O}_K/I &\cong \prod_{k=1}^n \mathcal{O}_K/\mathfrak{p}_k^{t_k} \\ (\mathcal{O}_K/I)^\times &\cong \left(\prod_{k=1}^n \mathcal{O}_K/\mathfrak{p}_k^{t_k}\right)^\times = \prod_{k=1}^n (\mathcal{O}_K/\mathfrak{p}_k^{t_k})^\times \\ \varphi_{\mathcal{O}_K}(I) &= \prod_{k=1}^n \varphi_{\mathcal{O}_K}(\mathfrak{p}_k^{t_k}) \end{aligned}$$

By the hypothesis, $v_{\mathfrak{p}_k}(\omega) = 0$ ($k = \overline{1,n}$), so from the previous theorem, we have

$$\begin{aligned} \omega^{\varphi_{\mathcal{O}_K}(\mathfrak{p}_k^{t_k})} &\equiv 1 && \pmod{\mathfrak{p}_k^{t_k}} \\ \omega^{\varphi_{\mathcal{O}_K}(I)} &\equiv 1^{\varphi_{\mathcal{O}_K}(I)/\varphi_{\mathcal{O}_K}(\mathfrak{p}_k^{t_k})} \equiv 1 && \pmod{\mathfrak{p}_k^{t_k}} \end{aligned}$$

But $\mathcal{O}_K/I \cong \prod_{k=1}^{n} \mathcal{O}_K/\mathfrak{p}_k^{t_k}$, so we finally get

$$\omega^{\varphi_{\mathcal{O}_K}(I)} \equiv 1 \pmod{I}$$

□

Proposition 14.13 (Legendre formula)**.** For a prime p and a positive integer n

$$v_p(n!) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor = \frac{n - s_p(n)}{p-1}$$

where $s_p(n)$ is the sum of digits of n in base p.

Proof. From 1 to n, by pigeonhole principle, there are exactly $\lfloor n/p^i \rfloor$ numbers divisible by p^i, so

$$\begin{aligned} v_p(n!) &= \sum_{m=1}^{n} v_p(m) = \sum_{i=1}^{\infty} i \left(\left\lfloor \frac{n}{p^i} \right\rfloor - \left\lfloor \frac{n}{p^{i+1}} \right\rfloor \right) \\ &= \sum_{i=1}^{\infty} i \left\lfloor \frac{n}{p^i} \right\rfloor - \sum_{i=2}^{\infty} (i-1) \left\lfloor \frac{n}{p^i} \right\rfloor = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor \end{aligned}$$

For the second equality, if $\overline{n_l n_{l-1} \dots n_1 n_0}_p$ is the representation of n in base p, then for $p^i \le n$

$$\left\lfloor \frac{n}{p^i} \right\rfloor = \left\lfloor \frac{n_l p^l + n_{l-1} p^{l-1} + \cdots + n_1 p + n_0}{p^i} \right\rfloor = n_l p^{l-i} + n_{l-1} p^{l-1-i} + \cdots + n_{i+1} p + n_i$$

Therefore,

$$\begin{aligned} \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor &= \sum_{i=1}^{l} n_l p^{l-i} + n_{l-1} p^{l-1-i} + \cdots + n_{i+1} p + n_i \\ &= n_l p^{l-1} + n_{l-1} p^{l-2} + \cdots + n_2 p + n_1 \\ &\quad + n_l p^{l-2} + n_{l-1} p^{l-3} + \cdots + n_2 \\ &\quad + \cdots \\ &\quad + n_l \\ &= \sum_{i=1}^{l} n_i (p^{i-1} + p^{i-2} + \cdots + p + 1) \\ &= \sum_{i=1}^{l} n_i \frac{p^i - 1}{p-1} = \sum_{i=0}^{l} n_i \frac{p^i - 1}{p-1} \\ &= \frac{\sum_{i=0}^{l} n_i p^i - \sum_{i=0}^{l} n_i}{p-1} = \frac{n - s_p(n)}{p-1} \end{aligned}$$

□

Corollary 14.14. If $r_i \in \mathbb{N}$ and p is a prime

$$v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) = \frac{\sum_{i=1}^k s_p(r_i) - s_p\left(\sum_{i=1}^k r_i\right)}{p-1}$$

Proof. By the Legendre formula

$$\begin{aligned} v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) &= v_p\left(\frac{\left(\sum_{i=1}^k r_i\right)!}{\prod_{i=1}^k r_i!}\right) \\ &= \frac{\left[\sum_{i=1}^k r_i - s_p\left(\sum_{i=1}^k r_i\right)\right] - \sum_{i=1}^k [r_i - s_p(r_i)]}{p-1} \\ &= \frac{\sum_{i=1}^k s_p(r_i) - s_p\left(\sum_{i=1}^k r_i\right)}{p-1} \end{aligned}$$

□

Corollary 14.15.

$$v_p\left(\binom{p(r_1 + r_2 + \cdots + r_k)}{pr_1, pr_2, \ldots, pr_k}\right) = v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right)$$

Proof. Note that s_p is invariant under padding zero, i.e. $s_p(pn) = s_p(n)$, so

$$\begin{aligned} v_p\left(\binom{p(r_1 + r_2 + \cdots + r_k)}{pr_1, pr_2, \ldots, pr_k}\right) &= \frac{\sum_{i=1}^k s_p(pr_i) - s_p\left(p\sum_{i=1}^k r_i\right)}{p-1} \\ &= \frac{\sum_{i=1}^k s_p(r_i) - s_p\left(\sum_{i=1}^k r_i\right)}{p-1} \\ &= v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) \end{aligned}$$

□

Proposition 14.16. Let $r_1, r_2, \ldots, r_k$ be non-negative integers, and p be a prime. Let τ be a number such that $\tau + 1$ is the maximum number of digits of r_i in base p. Then write r_i in base p and add them up as follows

1. $r_a = \overline{d_{a,\tau} d_{a,\tau-1} \ldots d_{a,0}}_p$ $(a = \overline{1,k})$. Note that $d_{a,\tau}$ can be 0.

2. x_0 is the quotient and y_0 is the remainder of $\sum_{i=1}^k d_{i,0}$ after dividing by p, i.e.

$$\sum_{i=1}^k d_{i,0} = x_0 p + y_0$$

3. Similarly, for each $j = \overline{1,\tau}$, let x_j be the quotient and y_j be the remainder of $\sum_{i=1}^{k} d_{i,j} + x_{j-1}$ after dividing by p

$$\sum_{i=1}^{k} d_{i,j} + x_{j-1} = x_j p + y_j$$

Prove that

$$\sum_{i=1}^{k} r_i = \overline{x_\tau y_\tau y_{\tau-1} \cdots y_0}_p$$
$$\sum_{j=1}^{\tau} x_j = v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right)$$

Proof. The first formula is just addition in base p

$$\begin{aligned}
\sum_{i=1}^{k} r_i &= \sum_{i=1}^{k}\sum_{j=0}^{\tau} p^j d_{i,j} = \sum_{j=0}^{\tau} p^j \sum_{i=1}^{k} d_{i,j} \\
&= y_0 + x_0 p + \sum_{j=1}^{\tau} p^j(-x_{j-1} + y_j + x_j p) \\
&= y_0 + x_0 p - x_0 p + y_1 p + x_1 p^2 - \cdots - x_{\tau-1} p^\tau + y_\tau p^\tau + x_\tau p^{\tau+1} \\
&= y_0 + y_1 p + \cdots + y_\tau p^\tau + x_\tau p^{\tau+1} = \overline{x_\tau y_\tau y_{\tau-1} \cdots y_0}_p
\end{aligned}$$

For the second formula, according to the previous corollary

$$\begin{aligned}
(p-1) v_p\left(\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) &= \sum_{i=1}^{k} s_p(r_i) - s_p\left(\sum_{i=1}^{k} r_i\right) \\
&= \sum_{i=1}^{k}\sum_{j=0}^{\tau} d_{i,j} - s_p(\overline{x_\tau y_\tau y_{\tau-1} \cdots y_0}_p) \\
&= \sum_{j=0}^{\tau}\sum_{i=1}^{k} d_{i,j} - x_\tau - \sum_{j=0}^{\tau} y_j \\
&= \sum_{j=0}^{\tau}(x_j p + y_j) - \sum_{j=1}^{\tau} x_{j-1} - x_\tau - \sum_{j=0}^{\tau} y_j \\
&= (p-1)\sum_{j=0}^{\tau} x_j
\end{aligned}$$

□

Lemma 14.17. Let p be a prime and $r_1, r_2, \ldots, r_k \in \mathbb{N}$ such that r_i are not simultaneously divisible by p (in other words, $v_p(\gcd(r_1, r_2, \ldots, r_k)) = 0$), then

$$v_p\left(\frac{1}{r_1 + r_2 + \cdots + r_k}\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) \geq 0$$

Proof. According to the previous proposition

$$v_p\left(\frac{1}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\right)=\sum_{j=0}^{\tau}x_j-v_p(\overline{x_\tau y_\tau y_{\tau-1}\cdots y_0}_p)$$

If $v_p(\overline{x_\tau y_\tau y_{\tau-1}\cdots y_0}_p)=0$, then we get Q.E.D since x_j is a quotient, which always non-negative (as long as the dividend is non-negative). Otherwise, suppose $v_p(\overline{x_\tau y_\tau y_{\tau-1}\cdots y_0}_p)=c+1$ for some $c\geq 0$. Equivalently, $y_c=y_{c-1}=\cdots=y_0=0$ but $y_{c+1}\neq 0$. On the other hand, r_i are not simultaneously divisible by p, so at least one of r_i's last digit (in base p) $d_{i,0}$ must be non-zero, which makes $x_0p=x_0p+y_0=\sum_{i=1}^{k}d_{i,0}\neq 0$. In other words, $x_0\neq 0$ or equivalently, $x_0\geq 1$. For each $1\leq j\leq c$, we deduce by induction that

$$x_jp=x_jp+y_j=x_{j-1}+\sum_{i=1}^{k}d_{i,j}\geq x_{j-1}\geq 1$$

Thus, $x_j\geq 1$ for all $0\leq j\leq c$. We then observe

$$\begin{aligned}v_p\left(\frac{1}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\right)&=\sum_{j=0}^{\tau}x_j-v_p(\overline{x_\tau y_\tau y_{\tau-1}\cdots y_0}_p)\\&=\sum_{j=0}^{\tau}x_j-(c+1)\geq\sum_{j=0}^{c}x_j-(c+1)\geq 0\end{aligned}$$

□

Corollary 14.18.

$$v_p\left(\frac{1}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\right)\geq -v_p(\gcd(r_1,r_2,\ldots,r_k))$$

Proof. Let $s=v_p(\gcd(r_1,r_2,\ldots,r_k))$ and $r_i=p^sz_i$, then $v_p(\gcd(z_1,z_2,\ldots,z_k))=0$. By corollary 14.15 and lemma 14.17

$$\begin{aligned}&v_p\left(\frac{1}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\right)\\&\quad=v_p\left(\frac{1}{p^s(z_1+z_2+\cdots+z_k)}\binom{p^s(z_1+z_2+\cdots+z_k)}{p^sz_1,p^sz_2,\ldots,p^sz_k}\right)\\&\quad=-s+v_p\left(\frac{1}{z_1+z_2+\cdots+z_k}\binom{z_1+z_2+\cdots+z_k}{z_1,z_2,\ldots,z_k}\right)\\&\quad\geq -s=-v_p(\gcd(r_1,r_2,\ldots,r_k))\end{aligned}$$

□

Corollary 14.19. Suppose $m=v_1r_1+v_2r_2+\cdots+v_kr_k$ where v_i are integers (but not all zero), r_i are non-negative integers, and m is a positive integer, then

$$\frac{m}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\in\mathbb{N}$$

Note: the denominator is non-zero since r_i cannot all be zero (which will also make $m = 0$, contradicting the assumption).

Proof. For each prime p, if $s = v_p(\gcd(r_1, r_2, \ldots, r_k))$, then $p^s \mid r_i$ and so $p^s \mid \sum_{i=1}^k v_i r_i = m$. Hence,

$$v_p\left(\frac{m}{r_1 + r_2 + \cdots + r_k}\binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k}\right) \geq v_p(m) - s = s - s = 0$$

But p is arbitrary, so the only possibility is that the expression must be an integer. □

Lemma 14.20. For every $r_1, r_2, \ldots, r_k \in \mathbb{N}$, not all simultaneously 0, $t \in \mathbb{N}$ and prime p

$$\binom{p(r_1 + r_2 + \cdots + r_k)}{pr_1, pr_2, \ldots, pr_k} \equiv \binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k} \pmod{p^{1+v_p\left(\sum_{i=1}^k r_i\right)}}$$

Proof. Let $s = v_p(\gcd(r_1, r_2, \ldots, r_k))$ and $r_i = p^s z_i$ (so that z_i are not simultaneously divisible by p). The following is a series of equivalence

$$\begin{aligned}
\binom{p(r_1 + r_2 + \cdots + r_k)}{pr_1, pr_2, \ldots, pr_k} &\equiv \binom{r_1 + r_2 + \cdots + r_k}{r_1, r_2, \ldots, r_k} \pmod{p^{1+v_p\left(\sum_{i=1}^k r_i\right)}} \\
\frac{\left(p\sum_{i=1}^k r_i\right)!}{\prod_{i=1}^k (pr_i)!} &\equiv \frac{\left(\sum_{i=1}^k r_i\right)!}{\prod_{i=1}^k r_i!} \pmod{p^{1+v_p\left(\sum_{i=1}^k r_i\right)}} \\
\frac{\left(p\sum_{i=1}^k r_i\right)!}{\left(\sum_{i=1}^k r_i\right)!} &\equiv \frac{\prod_{i=1}^k (pr_i)!}{\prod_{i=1}^k r_i!} \equiv \prod_{i=1}^k \frac{(pr_i)!}{r_i!} \pmod{p^{w_p^{[1]}}}
\end{aligned}$$

where $w_p^{[1]} = 1 + v_p\left(\sum_{i=1}^k r_i\right) - v_p\left(\left(\sum_{i=1}^k r_i\right)!\right) + v_p\left(\prod_{i=1}^k (pr_i)!\right)$

$$p^{\sum_{i=1}^k r_i} \prod_{\substack{t=0 \\ p\nmid t}}^{p\sum_{i=1}^k r_i - 1} \left(p\sum_{i=1}^k r_i - t\right) \equiv p^{\sum_{i=1}^k r_i} \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{pr_i - 1} (pr_i - t) \pmod{p^{w_p^{[1]}}}$$

since for each factor γ in the denominator, there is always $p\gamma$ appeared in the numerator. Vice versa, if $1 \leq \lambda \leq p\sum_{i=1}^k r_i$ is divisible by p, then $1 \leq \lambda/p \leq \sum_{i=1}^k r_i$, so λ/p should appear in the denominator. After each pair of corresponding factors cancels, we are left with $\sum_{i=1}^k r_i$ factors of p, and other factors which are not divisible by p. If we write each factor that is not divisible by p as $p\sum_{i=1}^k r_i - t$ for $0 \leq t \leq p\sum_{i=1}^k r_i - 1$ (or $pr_i - t$ for RHS), then p cannot divide t. Vice versa, if p does not divide t, then p neither divide $p\sum_{i=1}^k r_i - t$.

$$\prod_{\substack{t=0 \\ p\nmid t}}^{p\sum_{i=1}^k r_i - 1} \left(p\sum_{i=1}^k r_i - t\right) \equiv \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{pr_i - 1} (pr_i - t) \pmod{p^{w_p^{[2]}}}$$

where $w_p^{[2]} = w_p^{[1]} - \sum_{i=1}^k r_i$

$$\prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1}\sum_{i=1}^k z_i - 1} \left(p^{s+1} \sum_{i=1}^k z_i - t \right) \equiv \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1} z_i - 1} (p^{s+1} z_i - t) \pmod{p^{w_p^{[2]}}}$$

We then try to prove this congruence instead. But first, let's simplify $w_p^{[2]}$

$$\begin{aligned}
w_p^{[2]} &= 1 + v_p\left(\sum_{i=1}^k r_i\right) - v_p\left(\left(\sum_{i=1}^k r_i\right)!\right) + v_p\left(\prod_{i=1}^k (pr_i)!\right) - \sum_{i=1}^k r_i \\
&= 1 + v_p\left(\sum_{i=1}^k r_i\right) - v_p\left(\left(\sum_{i=1}^k r_i\right)!\right) + \sum_{i=1}^k v_p\left((pr_i)!\right) - \sum_{i=1}^k r_i \\
&= 1 + v_p\left(\sum_{i=1}^k r_i\right) + \frac{-\left[\sum_{i=1}^k r_i - s_p\left(\sum_{i=1}^k r_i\right)\right] + \left[\sum_{i=1}^k pr_i - s_p(pr_i)\right]}{p-1} - \sum_{i=1}^k r_i \\
&= 1 + v_p\left(\sum_{i=1}^k r_i\right) + \sum_{i=1}^k r_i - \frac{\sum_{i=1}^k s_p(r_i) - s_p\left(\sum_{i=1}^k r_i\right)}{p-1} - \sum_{i=1}^k r_i \\
&= 1 + v_p\left(\sum_{i=1}^k r_i\right) - v_p\left(\binom{r_1 + r_2 + \cdots r_k}{r_1, r_2, \ldots, r_k}\right) \\
&= 1 - v_p\left(\frac{1}{r_1 + r_2 + \cdots + r_k}\binom{r_1 + r_2 + \cdots r_k}{r_1, r_2, \ldots, r_k}\right) \\
&\leq 1 + v_p(\gcd(r_1, r_2, \ldots, r_k)) = 1 + s
\end{aligned}$$

From here, if we can show the following stronger congruence (same one we need to prove, but with larger exponent in the modulus), then we have Q.E.D.

$$\prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1}\sum_{i=1}^k z_i - 1} \left(p^{s+1} \sum_{i=1}^k z_i - t \right) \equiv \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1} z_i - 1} (p^{s+1} z_i - t) \pmod{p^{s+1}}$$

Simplifying each side, we get

$$\text{LHS} = \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1}\sum_{i=1}^k z_i - 1} \left(p^{s+1} \sum_{i=1}^k z_i - t \right) \equiv \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1}\sum_{i=1}^k z_i - 1} (-t) \pmod{p^{s+1}}$$

$$\text{RHS} = \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1} z_i - 1} (p^{s+1} z_i - t) \equiv \prod_{i=1}^k \prod_{\substack{t=0 \\ p\nmid t}}^{p^{s+1} z_i - 1} (-t) \pmod{p^{s+1}}$$

However, on the other hand

$$\begin{aligned}
\text{LHS} &\equiv \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_1-1} (-t) \prod_{\substack{t=p^{s+1}z_1\\ p\nmid t}}^{p^{s+1}\sum_{i=1}^k z_i-1} (-t) \\
&= \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_1-1} (-t) \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}\sum_{i=2}^k z_i-1} [-(t+p^{s+1}z_1)] \\
&\equiv \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_1-1} (-t) \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}\sum_{i=2}^k z_i-1} (-t) \equiv \cdots \\
&\equiv \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_1-1} (-t) \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_2-1} (-t) \cdots \prod_{\substack{t=0\\ p\nmid t}}^{p^{s+1}z_k-1} (-t) \equiv \text{RHS} \pmod{p^{s+1}}
\end{aligned}$$

□

Proof of theorem 14.9.
Part 1: by lemma 14.10, we have $f_n \in K$ for all $n \in \mathbb{N}$. In fact, $f_n \in \mathcal{O}_K$ if $c_i \in \mathcal{O}_K$, but this is irrelevant to the current discussion.
Part 2: since $\mathcal{O}_K \setminus \mathfrak{p}$ is a finite field, let $N(\mathfrak{p}) = p^\tau$ where p is the characteristic of $\mathcal{O}_K \setminus \mathfrak{p}$ (and so $p \in \mathfrak{p}$) and $\tau \geq 1$. By Newton identities on symmetric polynomials

$$f_{N(\mathfrak{p})^{t+1}} = f_{p^{\tau(t+1)}} = \sum_{\substack{s_1+2s_2+\cdots+ks_k=p^{\tau(t+1)}\\ s_1,s_2,\ldots,s_k\geq 0}} \frac{p^{\tau(t+1)}}{s_1+s_2+\cdots+s_k} \binom{s_1+s_2+\cdots+s_k}{s_1,s_2,\ldots,s_k} \prod_{i=1}^k c_i^{s_i}$$

For tuples $(s_1, s_2, \ldots, s_k)$ such that $v_p(\gcd(s_1, s_2, \ldots, s_k)) \leq \tau - 1$ (i.e. s_i are not simultaneously divisible by p^τ), then by corollary 14.18

$$v_p\left(\frac{1}{s_1+s_2+\cdots+s_k}\binom{s_1+s_2+\cdots+s_k}{s_1,s_2,\ldots,s_k}\right) \geq -v_p(\gcd(s_1,s_2,\ldots,s_k)) \geq -\tau+1$$

So

$$\frac{p^{\tau(t+1)}}{s_1+s_2+\cdots+s_k}\binom{s_1+s_2+\cdots+s_k}{s_1,s_2,\ldots,s_k} \in p^{\tau(t+1)-\tau+1}\mathcal{O}_K = p^{\tau t+1}\mathcal{O}_K \subseteq \mathfrak{p}^{\tau t+1} \subseteq \mathfrak{p}^{t+1}$$

since $p \in \mathfrak{p}$ (as $p = 0$ in $\mathcal{O}_K/\mathfrak{p}$). Under modulo $\mathfrak{p}^{t+1}$, such term (i.e. $v_p(\gcd(s_1, s_2, \ldots, s_k)) \leq \tau-1$) in $f_{N(\mathfrak{p})^{t+1}}$ will be zero, leaving us other terms such that s_i are simultaneously divisible by p^τ.

For each remaining such term, let $s_i = p^\tau r_i$

$$\begin{aligned} f_{N(\mathfrak{p})^{t+1}} &\equiv \sum_{\substack{s_1+2s_2+\cdots+ks_k=p^{\tau(t+1)} \\ p^\tau \mid s_1,s_2,\ldots,s_k}} \frac{p^{\tau(t+1)}}{s_1+s_2+\cdots+s_k}\binom{s_1+s_2+\cdots+s_k}{s_1,s_2,\ldots,s_k}\prod_{i=1}^{k} c_i^{s_i} \\ &= \sum_{\substack{r_1+2r_2+\cdots+kr_k=p^{\tau t} \\ r_1,r_2,\ldots,r_k\geq 0}} \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{p^\tau(r_1+r_2+\cdots+r_k)}{p^\tau r_1,p^\tau r_2,\ldots,p^\tau r_k}\prod_{i=1}^{k} c_i^{p^\tau r_i} \pmod{\mathfrak{p}^{t+1}} \end{aligned}$$

This is similar to $f_{N(\mathfrak{p})^t}$, using the same Newton identity

$$f_{N(\mathfrak{p})^t} = \sum_{\substack{r_1+2r_2+\cdots+kr_k=p^{\tau t} \\ r_1,r_2,\ldots,r_k\geq 0}} \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\prod_{i=1}^{k} c_i^{r_i}$$

Comparing the corresponding terms, we conjecture that the following congruence holds (for $r_1+2r_2+\cdots+kr_k = p^{\tau t}$ and $r_1, r_2, \ldots, r_n \geq 0$)

$$\begin{aligned} &\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{p^\tau(r_1+r_2+\cdots+r_k)}{p^\tau r_1,p^\tau r_2,\ldots,p^\tau r_k}\prod_{i=1}^{k} c_i^{p^\tau r_i} \\ &\quad\equiv \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\prod_{i=1}^{k} c_i^{r_i} \pmod{\mathfrak{p}^{t+1}} \end{aligned}$$

If so, then it is straightforward to see that $f_{N(\mathfrak{p})^{t+1}} \equiv f_{N(\mathfrak{p})^t} \pmod{\mathfrak{p}^{t+1}}$.

Let ω be the quotient and γ be the remainder of $v_p(\gcd(r_1, r_2, \ldots, r_k))$ after dividing τ, and set $r_i = (p^\tau)^\omega z_i, z_i = p^\gamma x_i$. From corollary 14.18 and corollary 14.19

$$\begin{aligned} &\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \in \mathbb{N} \\ v_p&\left(\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\right) \geq \tau t - (\tau\omega+\gamma) = \tau(t-\omega)-\gamma \end{aligned}$$

Note that $r_1 + 2r_2 + \cdots + kr_k = p^{\tau t}$, so $x_1 + 2x_2 + \cdots + kx_k = p^{\tau t-(\tau\omega+\gamma)} = p^{\tau(t-\omega)-\gamma}$. In particular, $\tau(t-\omega)-\gamma$ must be non-negative.

1. If $\gamma = 0$, then $\tau(t-\omega)-\gamma = \tau(t-\omega) \geq (t-\omega)$.
2. Otherwise, $1 \leq \gamma \leq \tau - 1$ (since γ is the remainder). $\tau(t-\omega)-\gamma$ is non-negative, so $\tau(t-\omega) \geq \gamma \geq 1$. In particular, $t-\omega \geq 1$. We can then show that

$$\tau(t-\omega)-\gamma \geq \tau(t-\omega)-(\tau-1) = (\tau-1)(t-\omega-1)+(t-\omega) \geq t-\omega$$

In any case, we always have $\tau(t-\omega)-\gamma \geq t-\omega$, so

$$\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \in p^{\tau(t-\omega)-\gamma} \subseteq p^{t-\omega} \subseteq \mathfrak{p}^{t-\omega}$$

On the other hand, note that $v_{\mathfrak{p}}(c_i) \geq 0$, so by the Euler totient theorem

$$c_i^{p^\tau r_i} = (c_i^{z_i})^{p^{\tau(\omega+1)}} = (c_i^{z_i})^{N(\mathfrak{p})^{\omega+1}} \equiv (c_i^{z_i})^{N(\mathfrak{p})^{\omega}} = c_i^{r_i} \pmod{\mathfrak{p}^{\omega+1}}$$

$$\prod_{i=1}^{k} c_i^{p^\tau r_i} \equiv \prod_{i=1}^{k} c_i^{r_i} \pmod{\mathfrak{p}^{\omega+1}}$$

Multiplying both side by $\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \in \mathfrak{p}^{t-\omega}$, we get

$$\begin{aligned}&\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\prod_{i=1}^{k} c_i^{p^\tau r_i} \\ &\quad\equiv \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k}\prod_{i=1}^{k} c_i^{r_i} \pmod{\mathfrak{p}^{t+1}} \qquad (14.1)\end{aligned}$$

On the other hand, by lemma 14.20

$$\binom{p^{\tau+1}(r_1+r_2+\cdots+r_k)}{p^{\tau+1}r_1,p^{\tau+1}r_2,\ldots,p^{\tau+1}r_k} \equiv \binom{p^{\tau}(r_1+r_2+\cdots+r_k)}{p^{\tau}r_1,p^{\tau}r_2,\ldots,p^{\tau}r_k} \pmod{p^{1+v_p\left(p^\tau \sum_{i=1}^k r_i\right)}}$$

Note that $1+v_p\left(p^\tau \sum_{i=1}^k r_i\right) = 1+\tau+v_p\left(\sum_{i=1}^k r_i\right) \geq 1+v_p\left(\sum_{i=1}^k r_i\right)$, so

$$\begin{aligned}\binom{p^{\tau+1}(r_1+r_2+\cdots+r_k)}{p^{\tau+1}r_1,p^{\tau+1}r_2,\ldots,p^{\tau+1}r_k} &\equiv \binom{p^{\tau}(r_1+r_2+\cdots+r_k)}{p^{\tau}r_1,p^{\tau}r_2,\ldots,p^{\tau}r_k} \\ &\equiv \binom{p^{\tau-1}(r_1+r_2+\cdots+r_k)}{p^{\tau-1}r_1,p^{\tau-1}r_2,\ldots,p^{\tau-1}r_k} \\ &\equiv \cdots \equiv \binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \pmod{p^{1+v_p\left(\sum_{i=1}^k r_i\right)}}\end{aligned}$$

Multiplying $p^{\tau t}/(r_1+r_2+\cdots+r_k)$ (which increase the exponent of p in the modulus by $\tau t - v_p(r_1+r_2+\cdots+r_k)$) on both sides, we get

$$\begin{aligned}&\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{p^{\tau+1}(r_1+r_2+\cdots+r_k)}{p^{\tau+1}r_1,p^{\tau+1}r_2,\ldots,p^{\tau+1}r_k} \\ &\quad\equiv \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \pmod{p^{\tau t+1}}\end{aligned}$$

Yet $p^{\tau t+1}\mathcal{O}_K \subseteq p^{t+1}\mathcal{O}_K \subseteq \mathfrak{p}^{t+1}$,

$$\begin{aligned}&\frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{p^{\tau+1}(r_1+r_2+\cdots+r_k)}{p^{\tau+1}r_1,p^{\tau+1}r_2,\ldots,p^{\tau+1}r_k} \\ &\quad\equiv \frac{p^{\tau t}}{r_1+r_2+\cdots+r_k}\binom{r_1+r_2+\cdots+r_k}{r_1,r_2,\ldots,r_k} \pmod{\mathfrak{p}^{t+1}} \qquad (14.2)\end{aligned}$$

Finally, by combining (14.1) and (14.2), we get Q.E.D. □

14.2 Structure of multiplicative group of a residue ring for a number field

We call a prime p *ramified* in some prime ideal $\mathfrak{p}$ if $v_{\mathfrak{p}}(p\mathcal{O}_K)$ is strictly greater than 1. If the exponent is exactly 1, then we call such prime *unramified* in $\mathfrak{p}$. To uncover the structure of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$, we introduce the following lemmas.

Lemma 14.21. If $N(\mathfrak{p}) = p^m$, then

1. $p \in \mathfrak{p}$, and $q \notin \mathfrak{p}$ for all prime $q \neq p$.
2. If $\mu = v_{\mathfrak{p}}(p)$ (which is at least 1), then

$$v_{\mathfrak{p}}(x) = \mu v_p(x) \text{ for all } x \in \mathbb{Z}$$

Proof. *Since $N(\mathfrak{p}) = p^m = |\mathcal{O}_K/\mathfrak{p}|$, $\mathcal{O}_K/\mathfrak{p}$ must be a finite field of characteristic p, so $p(1+\mathfrak{p}) = p + \mathfrak{p} = \mathfrak{p}$. In other words, $p \in \mathfrak{p}$. If q is another prime in $\mathfrak{p}$, then recall the Bézout lemma, $px + qy = 1$ for some integers x, y. But $\mathbb{Z}$ is a subring of $\mathcal{O}_K$, so $1 \in \mathfrak{p}$, contradicting the primeness of $\mathfrak{p}$.

*As a result, $v_{\mathfrak{p}}(q) = 0$ for prime $q \neq p$, so for any integer x, only the prime power of p matters

$$v_{\mathfrak{p}}(x) = v_{\mathfrak{p}}(p^{v_p(x)}) = v_{\mathfrak{p}}(p)v_p(x) = \mu v_p(x)$$

□

Lemma 14.22 (Odd prime)**.** For any odd prime p and integer $k \geq 2$, we have $k \geq v_p(k) + 2$. Equality happens exactly when either $k = 2$, or $k = p = 3$.

Proof. If $v_p(k) = 0$, then $k \geq 2 = 2 + v_p(k)$. Otherwise, suppose $v_p(k) \geq 1$. Consider the function $f(x) = p^x - x - 2$ for $x \geq 0$: its derivative $f'(x) = \ln p \cdot p^x - 1 \geq \ln 3 \cdot p^0 - 1 > 0$ is always positive, so f is *strictly* increasing on $[0, \infty)$. In particular, $f(v_p(k)) \geq f(1) = p - 3 \geq 0$, so $v_p(k) + 2 \leq p^{v_p(k)} \leq k$.

In the case $v_p(k) = 0$, if equality occurs then $k = 2$, which indeed makes $v_p(k) = 0$ regardless of (odd prime) p. In the other case, $v_p(k) \geq 1$, since f is strictly increasing, if equality occurs then $v_p(k) = 0$ (since $f(v_p(k)) = f(1)$) and $p = 3$. But that also means, $k = 3^{v_p(k)} = v_p(k) + 2 = 3$. □

Lemma 14.23 (Even prime)**.**

1. For all integer $k \geq 3$, we have $k \geq v_2(k) + 2$. Equality happens exactly when $k = 4$.
2. For all integer $k \geq 5$, $k \geq v_2(k) + 5$. Equality happens exactly when $k = 5, 6$ or 8.

Proof. *For the first inequality: we divide into 2 cases base on 2-adic value

1. If $v_2(k) \leq 1$: then $2 + v_2(k) \leq 3 \leq k$.
2. If $v_2(k) \geq 2$: let $f(x) = 2^x - x - 2$, where its derivative $f'(x) = \ln 2 \cdot 2^x - 1 \geq \ln 2 \cdot 2^2 - 1 > 0$ is positive for all $x \geq 2$, so f is strictly increasing on $[2, \infty)$. We then have $f(v_2(k)) \geq f(2) = 0$, so $v_2(k) + 2 \leq 2^{v_2(k)} \leq k$.

If equality occurs, then either $v_2(k) = 1$ and $k = 3 = v_2(k) + 2$ (which is impossible), or $v_2(k) = 2$ and $k = 2^{v_2(k)} = v_2(k) + 2 = 4$ (which is indeed true).

*For the second inequality: we divide into the following cases (still base on 2-adic value)

1. If $v_2(k) = 0$: then $5 + v_2(k) = 5 \leq k$.
2. If $v_2(k) = 1$: then $k \geq 6 = v_2(k) + 1$ (the least even number over 5) .
3. If $v_2(k) = 2$: then $k \geq 8 > v_2(k) + 5$ (the least number over 5 that is divisble by 4).
4. If $v_2(k) \geq 3$: by induction (or calculus like the above), one can show that $2^n \geq n + 5$ for all $n \geq 3$ (with equality exactly at $n = 3$). Hence, $k \geq 2^{v_2(k)} \geq v_2(k) + 5$.

If equality occurs, then either $v_2(k) = 0$ and $k = 5 = 5 + v_2(k)$ (which is true), or $v_2(k) = 1$ and $k = 6 = v_2(k) + 5$ (which is also true), or $v_2(k) = 3$ and $k = 2^{v_2(k)} = v_2(k) + 5 = 8$ (which is indeed true). Note the inequality is strict in the case $v_2(k) = 2$. □

Lemma 14.24. For all integers $n, w \geq 1$, prime p, and $1 \leq k \leq p^n$

$$v_p\left(\binom{p^n w}{k}\right) \geq n - v_p(k)$$

Proof. Based on proposition 14.16, we write k and $p^n - k$ in base p as

$$k = \overline{\cdots a_{n-1} \ldots a_0}$$
$$p^n w - k = \overline{\cdots b_{n-1} \ldots b_0}$$

with x_j be the quotient, $0 \leq y_j \leq p-1$ be the remainder of $x_{j-1} + a_j + b_j$ after dividing by p

$$a_0 + b_0 = p x_0 + y_0$$
$$a_j + b_j + x_{j-1} = p x_j + y_j \text{ for } j \geq 1$$

We only pay attention to the last n digits of k and $p^n w - k$

1. Since $p^n w = k + (p^n w - k) = \overline{\cdots y_{n-1} \cdots y_0}$, we know that $y_0 = y_1 = \cdots = y_{n-1} = 0$. Also, since $0 \leq a_j, b_j \leq p-1$, we also have $0 \leq x_j \leq 1$. Indeed, under induction

(a) Base case $j = 0$: since $px_0 \leq px_0 + y_0 = a_0 + b_0 \leq 2(p-1)$, so $x_0 \leq 2 - 2/p$, or $x_0 \leq 1$ since x_0 is an integer.

(b) Induction case: suppose $x_{j-1} \leq 1$, then $px_j \leq px_j + y_j = a_0 + b_0 + x_{j-1} \leq 2p - 1$. In other words, $x_j \leq 2 - 1/p$, or $x_j \leq 1$.

2. To show $v_p\left(\binom{p^n w}{k}\right) \geq n - v_p(k)$ for $k \geq 1$, consider two separate cases

(a) If $x_j = 0$ for all $0 \leq j < n$, then $a_0 + b_0 = px_0 + y_0 = 0$ and $a_j + b_j = a_j + b_j + x_{j-1} = px_j + y_j = 0$ for each $1 \leq j < n$. Therefore, $a_j = b_j = 0$ for all $0 \leq j < n$. Since $k \geq 1$ (i.e. k is non-zero), $v_p(k) \geq n$, so

$$v_p\left(\binom{p^n w}{k}\right) = \sum_j x_j \geq \sum_{j=0}^{n-1} x_j = 0 \geq n - v_p(k)$$

(b) There exists some $0 \leq j < n$ such that $x_j \neq 0$, or $x_j = 1$. Pick j_0 to be the smallest such. We divide into 2 cases

i. If $j_0 = 0$, then $a_0 + b_0 = px_0 + y_0 = p$. But $a_0, b_0 \leq p - 1$, so a_0 and b_0 cannot be zero (otherwise, the other is equal to p). Hence, $v_p(k) = 0 = j_0$.

ii. If $1 \leq j_0 < n$, then by definition, $x_j = 0$ for all $j < j_0$. However, $a_0 + b_0 = px_0 + y_0 = 0$ and $a_j + b_j + x_{j-1} = px_j + y_j = 0$ for all $1 \leq j < j_0 < n$, so $a_j = b_j = 0$ for all $0 \leq j < j_0$. This shows that p^{j_0} must divide k, or simply $j_0 \leq v_p(k)$. On the other hand, $a_{j_0} + b_{j_0} = a_{j_0} + b_{j_0} + x_{j_0 - 1} = px_{j_0} + y_{j_0} = p$, while $a_{j_0}, b_{j_0} \leq p-1$, so both a_{j_0} and b_{j_0} cannot be 0. Therefore, $v_p(k) \leq j_0$, or just $v_p(k) = j_0$ since we have already shown $j_0 \leq v_p(k)$.

In any case, we have $v_p(k) = j_0$. Additionally, for $n > j > j_0 \geq 0$ (so $j \geq 1$), we have $a_j + b_j + x_{j-1} = px_j + y_j = px_j$. By induction, if $x_{j-1} = 1$, then $px_j \geq 1$, and so $x_j = 1$ necessarily (as $0 \leq x_j \leq 1$). Therefore, $x_j = 1$ for all $j_0 \leq j < n$. In a sense, x_j represents the carries of the addition between $\overline{1 \ldots 1} \times p^{j_0}$ and $\overline{1} \times p^{j_0}$. Finally, by proposition 14.16, we get

$$v_p\left(\binom{p^n w}{k}\right) = \sum_j x_j \geq \sum_{j=0}^{n-1} x_j = \sum_{j=j_0}^{n-1} x_j = n - 1 - j_0 + 1 = n - v_p(k)$$

□

As a result from the proof above, equality happens when $w = 1$.

Proposition 14.25. For all integer $n \geq 1$, prime p, and $1 \leq k \leq p^n$

$$v_p\left(\binom{p^n}{k}\right) = n - v_p(k)$$

Lemma 14.26.

1. If $k \geq 1$, then $v_2\left(3^{2^k} - 1\right) = k + 2$.
2. $v_2(3^k + 1) \leq 2$ for all nonnegative integer k. Equality happens exactly when k is odd (i.e. $v_2(3^k + 1) = 1$ whenever k is even).

Proof. By the binomial theorem

$$3^{2^k} = (1+2)^{2^k} = \sum_{i=0}^{2^k} \binom{2^k}{i} 2^i$$

Furthermore, by lemma 14.23 and proposition 14.25

$$v_2\left(\binom{2^k}{i}2^i\right) = k - v_2(i) + i \geq k + 2 \text{ for } i \geq 3$$

$$\begin{aligned} 3^{2^k} &= 1 + 2^k \cdot 2 + \frac{2^k(2^k-1)}{2} \cdot 2^2 + \sum_{i=3}^{2^k} \binom{2^k}{i} 2^i \\ &= 1 + 2^{2k+1} + \sum_{i=3}^{2^k} \binom{2^k}{i} 2^i \equiv 1 \pmod{2^{k+2}} \end{aligned}$$

Note that $k \geq 1$, which means that $2^k \geq 2$ (that's why we can extract the first 3 terms of the sum) and $2k + 1 \geq k + 2$ (for the last congruence). On the other hand,

1. If $k = 1$, then $v_2(3^{2^k} - 1) = 3 = k + 2$
2. If $k \geq 2$, then $2^k \geq 4$ and $2k + 1 \geq k + 3$, so

$$v_2\left(\binom{2^k}{i}2^i\right) = k - v_2(i) + i \geq k + 5 \text{ for } i \geq 5$$

$$\begin{aligned} 3^{2^k} &= \sum_{i=0}^{4} \binom{2^k}{i} 2^i + \sum_{i=5}^{2^k} \binom{2^k}{i} 2^i \\ &\equiv 1 + 2^{2k+1} + 2^{k+3}\frac{(2^k-1)(2^{k-1}-1)}{3} \\ &\quad + 2^{k+2}\frac{(2^k-1)(2^{k-1}-1)(2^k-3)}{3} \\ &\equiv 1 + 2^{k+2}\frac{(2^k-1)(2^{k-1}-1)(2^k-3)}{3} \not\equiv 1 \pmod{2^{k+3}} \end{aligned}$$

In other words, $v_2(3^{2^k}-1) < k+3$, which, when combines with the previous lower bound (i.e. $3^{2^k} \equiv 1 \pmod{2^{k+2}}$), gives us $v_2(3^{2^k} - 1) = k + 2$.

*For the second part

1. If n is even (i.e. $n = 2k$), then $3^n + 1 = 9^k + 1 \equiv 2 \pmod 8$, thus $v_2(3^n + 1) = 1$.
2. If n is odd (i.e. $n = 2k + 1$), then $3^n + 1 = 9^k \cdot 3 + 1 \equiv 4 \pmod 8$, thus $v_2(3^n + 1) = 2$.

□

14.2.1 p is odd and unramified in $\mathfrak{p}$

Theorem 14.27. If $N(\mathfrak{p}) = p^m$ and p is unframified in $\mathfrak{p}$, then

$$(\mathcal{O}_K/\mathfrak{p}^t)^\times \cong (C_{p^{t-1}})^m \times C_{N(\mathfrak{p})-1}$$

where C_n is the cyclic group of order n.

Proof. When $t = 1$, the isomorphism becomes $(\mathcal{O}_K/\mathfrak{p})^\times \cong C_{N(\mathfrak{p})-1}$, which is true since the multiplicative group of a field is cyclic (and $N(\mathfrak{p}) = p^m$ is the cardinality of $\mathcal{O}_K/\mathfrak{p}$). For $t \geq 2$, let

1. $\mathfrak{g} = g_0 + \mathfrak{p}$ be a generator of $(\mathcal{O}_K/\mathfrak{p})^\times$.
2. $\beta = \{\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_m\}$ be a basis of $\mathcal{O}_K/\mathfrak{p}$ over $\mathbb{F}_p$ (note that $\mathcal{O}_K/\mathfrak{p} \cong \mathbb{F}_{p^m}$). Also, let h_i be some number in $\mathcal{O}_K$ such that $\mathfrak{h}_i = h_i + \mathfrak{p}$. If $h_i \in \mathfrak{p}$, then $h_i + \mathfrak{p} = \mathfrak{p}$, contradicting the fact that β is a basis of $\mathcal{O}_K/\mathfrak{p}$. Thus, $h_i \notin \mathfrak{p}$, meaning $v_\mathfrak{p}(h_i) = 0$.

*We prove that $1 + ph_i$ is of order p^{t-1} modulo $\mathfrak{p}^t$. By lemma 14.21, lemma 14.22, and lemma 14.24

$$v_\mathfrak{p}\left(\binom{p^{t-1}}{k} p^k h_i^k\right) = v_\mathfrak{p}(p) v_p\left(\binom{p^{t-1}}{k} p^k\right) + v_\mathfrak{p}(h_i^k) \geq t - 1 - v_p(k) + k \geq t + 1$$

whenever $k \geq 2$

$$\begin{aligned}(1 + ph_i)^{p^{t-1}} &= \sum_{k=0}^{p^{t-1}} \binom{p^{t-1}}{k} p^k h_i^k = 1 + p^{t-1} \cdot ph_i + \sum_{k=2}^{p^{t-1}} \binom{p^{t-1}}{k} p^k h_i^k \\ &\equiv 1 + p^t h_i \equiv 1 \pmod{\mathfrak{p}^t}\end{aligned}$$

We can split the sum since $p^{t-1} \geq 3^{2-1} = 3$. On the other hand, since $p^{t-2} \geq 3^{2-2} = 1$, we have

$$v_\mathfrak{p}\left(\binom{p^{t-2}}{k} p^k h_i^k\right) = v_\mathfrak{p}(p) v_p\left(\binom{p^{t-2}}{k} p^k\right) + v_\mathfrak{p}(h_i^k) \geq t - 2 - v_p(k) + k \geq t$$

whenever $k \geq 2$

$$\begin{aligned}(1 + ph_i)^{p^{t-2}} &= \sum_{k=0}^{p^{t-2}} \binom{p^{t-2}}{k} p^k h_i^k = 1 + p^{t-2} \cdot ph_i + \sum_{k=2}^{p^{t-2}} \binom{p^{t-2}}{k} p^k h_i^k \\ &\equiv 1 + p^{t-1} h_i \not\equiv 1 \pmod{\mathfrak{p}^t}\end{aligned}$$

(note that $h_i \notin \mathfrak{p}$ so $v_{\mathfrak{p}}(p^{t-1}h_i) = t-1 < t$). Let H_i be the subgroup of $\mathcal{O}_K/\mathfrak{p}^t$ generated by $1 + ph_i + \mathfrak{p}^t$. By the result above, we know that H_i is a cyclic group of order p^{t-1}.

*Next, let s be the order of g_0 modulo $\mathfrak{p}^t$ (it can be different based on the choice of g_0, but it always have order $N(\mathfrak{p})-1$ modulo $\mathfrak{p}$). Since $\mathfrak{p}^t \subseteq \mathfrak{p}$, we also have $g^s \equiv 1 \pmod{\mathfrak{p}}$, so s must be a multiple of $N(\mathfrak{p})-1$ by Lagrange theorem. Let $s = r[N(\mathfrak{p})-1]$, then g_0^r will have order $N(\mathfrak{p})-1$ modulo $\mathfrak{p}^t$. If G is the subgroup generated by $g_0^r + \mathfrak{p}^t$, then it should be a cyclic group of order $N(\mathfrak{p})-1$.

*Since $|G|\prod_{i=1}^m |H_i| = [N(\mathfrak{p})-1]\cdot\left(p^{t-1}\right)^m = N(\mathfrak{p})^{t-1}[N(\mathfrak{p})-1]$, and $|(\mathcal{O}_K/\mathfrak{p}^t)^\times| = \varphi_{\mathcal{O}_K}(\mathfrak{p}^t) = N(\mathfrak{p})^t - N(\mathfrak{p})^{t-1}$ (according to the Euler totient theorem), the order of $G \times \prod_{i=1}^m H_i$ and $(\mathcal{O}_K/\mathfrak{p}^t)^\times$ are the same. We conjecture that they are actually isomorphic.

Since H_i are cyclic subgroups of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$ (of order p^{t-1}), the kernel the multiplication map $\sigma : \prod_{i=1}^m H_i \to (\mathcal{O}_K/\mathfrak{p}^t)^\times$ (sending $(b_1, b_2, \ldots, b_m) \mapsto b_1 b_2 \cdots b_m \pmod{\mathfrak{p}^t}$) consists of elements of the form

$$\prod_{i=1}^m (1+ph_i)^{x_i} \equiv 1 \pmod{\mathfrak{p}^t}$$

for some $0 \le x_1, x_2, \ldots, x_m < p^{t-1}$. Suppose at least one x_i is not zero, then $w = v_p(\gcd(x_1, x_2, \ldots, x_m))$ is non-negative (not $-\infty$), and $x_i = p^w y_i$ such that y_i are not simultaneously divisible by p. We also have $w < t-1$ (since $x_i < p^{t-1}$), so the congruence above works for smaller modulus $\mathfrak{p}^{w+2}$.

$$\prod_{i=1}^m (1+ph_i)^{x_i} \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

$$\prod_{i=1}^m (1+ph_i)^{p^w y_i} \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

$$\prod_{i=1}^m \sum_{k=1}^{p^w y_i} \binom{p^w y_i}{k} p^k h_i^k \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

By lemma 14.22 and lemma 14.24, for $k \ge 2$, we have

$$v_{\mathfrak{p}}\left(\binom{p^w y_i}{k} p^k h_i^k\right) = v_p\left(\binom{p^w y_i}{k}\right) + k \ge w - v_p(k) + k \ge w+2$$

Consider

1. If $x_i = p^w y_i \ge 1$, then we can split the sum as

$$\begin{aligned}\sum_{k=1}^{p^w y_i} \binom{p^w y_i}{k} p^k h_i^k &= 1 + p^w y_i \cdot ph_i + \sum_{k=2}^{p^w y_i} \binom{p^w y_i}{k} p^k h_i^k \\ &\equiv 1 + p^{w+1} y_i h_i \pmod{\mathfrak{p}^{w+2}}\end{aligned}$$

2. If $x_i = p^w y_i = 0$, then $y_i = 0$, and the sum turns into

$$\sum_{k=1}^{p^w y_i} \binom{p^w y_i}{k} p^k h_i^k = 1 = 1 + p^{w+1} y_i h_i$$

In either case, the previous congruence turns into

$$\prod_{i=1}^{m} \sum_{k=1}^{p^w y_i} \binom{p^w y_i}{k} p^k h_i^k \equiv \prod_{i=1}^{m} \left(1 + p^{w+1} y_i h_i\right) \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

If we distribute the above product (in the middle), every term that multiplies at least 2 different factors of the form $p^{w+1} y_i h_i$, then such term has to be divisible by p^{2w+2}. But $v_{\mathfrak{p}}(p^{2w+2}) = 2w + 2 \geq w + 2$, so those terms can be cancelled modulo $\mathfrak{p}^{w+2}$. That leaves us with 1 and $p^{w+1} y_i h_i$

$$1 + \sum_{i=1}^{m} p^{w+1} y_i h_i \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

$$p^{w+1} \sum_{i=1}^{m} y_i h_i \equiv 0 \pmod{\mathfrak{p}^{w+2}}$$

Since $v_{\mathfrak{p}}(p^{w+1}) = w + 1$,

$$v_{\mathfrak{p}}\left(\sum_{i=1}^{m} y_i h_i\right) = v_{\mathfrak{p}}\left(p^{w+1} \sum_{i=1}^{m} y_i h_i\right) - (w+1) \geq (w+2) - (w+1) = 1$$

i.e.

$$\sum_{i=1}^{m} y_i h_i \in \mathfrak{p}$$

$$\sum_{i=1}^{m} y_i (h_i + \mathfrak{p}) = \mathfrak{p}$$

$$\sum_{i=1}^{m} y_i \mathfrak{h}_i = \mathfrak{p}$$

Yet $\beta = \{\mathfrak{h}_i : 1 \leq i \leq m\}$ is a basis of $\mathcal{O}_K/\mathfrak{p}$ over $\mathbb{F}_p$, so the coefficients y_i must be zero modulo p, contradicting the fact that they are not simultaneously divisible by p. Thus, the intial assumption is false, which means that all x_i has to be zero. The kernel is then trivial, or equivalently, the multiplication map $\sigma : \prod_{i=1}^{m} H_i \to (\mathcal{O}_K/\mathfrak{p}^t)^\times$ is an embedding, with the image be just $H_1 H_2 \cdots H_m$.

*Finally, consider another multiplication map $\sigma : G \times \prod_{i=1}^{m} H_i \to (\mathcal{O}_K/\mathfrak{p}^t)^\times$: since the order of G (which is $N(\mathfrak{p}) = p^m - 1$) and $\prod_{i=1}^{m} H_i$ (which is $p^{m(t-1)}$) is relatively prime, the kernel is trivial. Indeed,

1. Given $\mathfrak{a} \in G$ and $\mathfrak{b}_i \in H_i$ such that $\mathfrak{a}\prod_{i=1}^m \mathfrak{b}_i = 1 + \mathfrak{p}$ (the identity in $(\mathcal{O}_K/\mathfrak{p}^t)^\times$), then $\mathfrak{a}^{p^{m(t-1)}} = \left(\prod_{i=1}^m \mathfrak{b}_i\right)^{-p^{m(t-1)}} = 1 + \mathfrak{p}$.

2. But $p^m - 1$ and $p^{m(t-1)}$ are relative prime, so by Bézout lemma, there exists some integer x, y such that $(p^m - 1)x + p^{m(t-1)}y = 1$. Hence, $\mathfrak{a} = \mathfrak{a}^{(p^m-1)x+p^{m(t-1)}y} = \left(\mathfrak{a}^{p^m-1}\right)^x \cdot \left(\mathfrak{a}^{p^{m(t-1)}}\right)^y = 1 + \mathfrak{p}$.

3. It also means that $\prod_{i=1}^m \mathfrak{b}_i = \mathfrak{a}^{-1} = 1 + \mathfrak{p}$, but from the previous result, we must have $\mathfrak{b}_i = 1 + \mathfrak{p}$.

Thus, $G \times \prod_{i=1}^m H_i$ embeds into $(\mathcal{O}_K/\mathfrak{p}^t)^\times$. But we already observe that these 2 groups have the same (finite) order, so the embedding is actually an isomorphism. We conclude that

$$(\mathcal{O}_K/\mathfrak{p}^t)^\times \cong G \times \prod_{i=1}^m H_i \cong C_{N(\mathfrak{p})-1} \times \left(C_{p^{t-1}}\right)^m$$

□

Corollary 14.28. For each $t \geq 1$

$$(\mathbb{Z}/p^t\mathbb{Z})^\times \cong C_{p-1} \times C_{p^{t-1}}$$

odd and ramified in $\mathfrak{p}$

9. If $N(\mathfrak{p}) = p^m$ and $v_\mathfrak{p}(p) = \mu$, then for $t \geq \mu$, we get

$$_K/\mathfrak{p}^t)^\times \cong \left(\prod_{j=0}^{\mu-1} C_{p^{\lceil (t-j)/\mu \rceil - 1}} \times C_{\mu-1}\right)^m \times C_{N(\mathfrak{p})-1}$$

generator modulo $\mathfrak{p}$ (i.e. a primitive root) of $(\mathcal{O}_K/\mathfrak{p})^\times$. In , $g_0^{N(\mathfrak{p})-1} \equiv 1 \pmod{\mathfrak{p}}$.

$\ldots, h_m\}$ be a basis modulo $\mathfrak{p}$ of $\mathcal{O}_K/\mathfrak{p}$ over $\mathbb{F}_p$. In particular, h_i $\mathfrak{p}$, and whenever $z_1h_1 + z_2h_2 + \cdots + z_mh_m \equiv 0 \pmod{\mathfrak{p}}$ (with $_i$ are simultaneously divisible by p.

^{2} (i.e. $v_\mathfrak{p}(\omega) = 1$) so that $p \mid \omega^\mu$ (note that, by definition of μ, $\mathfrak{q}$ where $v_\mathfrak{p}(\mathfrak{q}) = 0$)

e show that the order of $1+p\omega^j h_i$ modulo $\mathfrak{p}^t$ is p^{s_j} where $s_j = \lceil (t-j)/\mu \rceil - 1$ $0 \le j < \mu$). Indeed, by lemma 14.21, lemma 14.22, and lemma 14.24

$$\begin{aligned}
v_{\mathfrak{p}}\left(\binom{p^{s_j}}{k} p^k \omega^{jk} h_i^k\right) &= \mu(s_j - v_p(k)) + (\mu + j)k \ge \mu\left(\frac{t-j}{\mu} - 1 - v_p(k) + k\right) + jk \\
&\ge t + \mu + j(k-1) \ge t + \mu + j > t \\
&\text{whenever } k \ge 2, j \ge 0 \\
(1 + p\omega^j h_i)^{p^{s_j}} &= \sum_{k=0}^{p^{s_j}} \binom{p^{s_j}}{k} p^k \omega^{jk} h_i^k = 1 + p^{s_j+1}\omega^j h_i + \sum_{k=2}^{p^{s_j}} \binom{p^{s_j}}{k} p^k \omega^{jk} h_i^k \\
&\equiv 1 + p^{\lceil (t-j)/\mu \rceil} \omega^j h_i \equiv 1 \pmod{\mathfrak{p}^t}
\end{aligned}$$

e last congruence holds because $v_{\mathfrak{p}}\left(p^{\lceil (t-j)/\mu \rceil}\omega^j h_i\right) = \mu\lceil (t-j)/\mu \rceil + j \ge t$. the other hand,

$$\begin{aligned}
v_{\mathfrak{p}}\left(\binom{p^{s_j-1}}{k} p^k \omega^{jk} h_i^k\right) &= \mu(s_j - 1 - v_p(k)) + (\mu + j)k \ge \mu\left(\frac{t-j}{\mu} - 2 - v_p(k) + k\right) + jk \\
&\ge t + j(k-1) \ge t \\
&\text{whenever } k \ge 2, j \ge 0 \\
(1 + p\omega^j h_i)^{p^{s_j-1}} &= \sum_{k=0}^{p^{s_j-1}} \binom{p^{s_j-1}}{k} p^k \omega^{jk} h_i^k = 1 + p^{s_j}\omega^j h_i + \sum_{k=2}^{p^{s_j-1}} \binom{p^{s_j-1}}{k} p^k \omega^{jk} h_i^k \\
&\equiv 1 + p^{\lceil (t-j)/\mu \rceil - 1} \omega^j h_i \not\equiv 1 \pmod{\mathfrak{p}^t}
\end{aligned}$$

ce $v_{\mathfrak{p}}\left(p^{\lceil (t-j)/\mu \rceil - 1}\omega^j h_i\right) < \mu \cdot (t-j)/\mu + j = t$.

ote that the order of $1 + \omega^j h_i$ (modulo $\mathfrak{p}^t$) is a *non-trivial* power of p $1 \le j \le \mu - 1$, so let r_{ij} be such that $(1 + \omega^j h_i)^{r_{ij}}$ has order p. Let W_{ij} be subgroup (of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$) generated by $1 + p\omega^j h_i$, and M_{ij} be the subgroup erated by $(1 + \omega^j h_i)^{p^{r_{ij}}}$. Note that $|W_{ij}| = p^{s_j}$, $|M_{ij}| = p$, and

$$\prod_{i=1}^{m}\prod_{j=0}^{\mu-1} p^{s_j} \cdot \prod_{i=1}^{m}\prod_{j=1}^{\mu-1} p = p^{m(t-1)}$$

ch the order of the p-torsion subgroup of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$, so we will prove that product of these groups embedded into $(\mathcal{O}_K/\mathfrak{p}^t)^\times$ (through multiplication p). Let $0 \le \lambda_{ij} < s_j$ and $0 \le \kappa_{ij} < p$ be such that

$$\prod_{j=0}^{\mu-1}\prod_{i=1}^{m}(1 + p\omega^j h_i)^{\lambda_{ij}} \cdot \prod_{j=1}^{\mu-1}\prod_{i=1}^{m}(1 + \omega^j h_i)^{p^{r_{ij}}\kappa_{ij}} \equiv 1 \pmod{\mathfrak{p}^t}$$

$\mathfrak{p}^\mu$, this congruence turns into

$$\prod_{j=1}^{\mu-1}\prod_{i=1}^{m}(1+\omega^j h_i)^{p^{r_{ij}}\kappa_{ij}} \equiv 1 \pmod{\mathfrak{p}^\mu}$$

□

unramified in $\mathfrak{p}$

0. Suppose $N(\mathfrak{p}) = 2^m$ $(m \geq 1)$ and 2 is unramified in $\mathfrak{p}$. Then

$(\mathcal{O}_K/\mathfrak{p})^\times \cong C_{N(\mathfrak{p})-1}$

$(\mathcal{O}_K/\mathfrak{p}^2)^\times \cong (C_2)^m \times C_{N(\mathfrak{p})-1}$

$(\mathcal{O}_K/\mathfrak{p}^t)^\times \cong (C_2 \times C_{2^{t-2}}) \times (C_{2^{t-1}})^{m-1} \times C_{N(\mathfrak{p})-1}$

$t = 1$ proceeds just the same as the previous theorem (by noting licative group of a finite field is cyclic). For the case $t = 2$, let

be a generator of the (cyclic) group $(\mathcal{O}_K/\mathfrak{p})^\times$.

$_2, \ldots, \mathfrak{h}_m\} = \{h_i + \mathfrak{p} : 1 \leq i \leq m\}$ be a basis of $\mathcal{O}_K/\mathfrak{p}$ over $\mathbb{F}_2$.

proof for odd prime, we first show the order of $1 + 2h_i$ is 2 eed, note that $h_i \notin \mathfrak{p}$ (otherwise β is not a basis), so $2h_i \not\equiv 0$ ence, $1 + 2h_i \not\equiv 1 \pmod{\mathfrak{p}^2}$. On the other hand, since $v_\mathfrak{p}(2) = 1$ in $\mathfrak{p}$), we get $(1 + 2h_i)^2 = 1 + 4h_i + 4h_i^2 \equiv 1 \pmod{\mathfrak{p}^2}$.

yclic) subgroup of $(\mathcal{O}_K/\mathfrak{p}^2)^\times$ generated by $1+2h_i+\mathfrak{p}^2$. Consider he kernel of the multiplication map $\sigma : \prod_{i=1}^m H_i \to (\mathcal{O}_K/\mathfrak{p}^2)^\times$ $\ldots, b_m) \mapsto b_1 b_2 \cdots b_m$)

$$\prod_{i=1}^{m}(1+2h_i)^{u_i} \equiv 1 \pmod{\mathfrak{p}^2}$$

≤ 1. In simple term, it is just a product of $1 + 2h_i$ (but not from 1 to m). Therefore, since $v_\mathfrak{p}(4) = 2$, we have

$$1 + 2\sum_{i=1}^{m} u_i h_i \equiv 1 \pmod{\mathfrak{p}^2}$$

$$2\sum_{i=1}^{m} u_i h_i \equiv 0 \pmod{\mathfrak{p}^2}$$

s then necessary that $\sum_{i=1}^{m} u_i h_i \in \mathfrak{p}$. But since $\mathfrak{h}_i = h_i + \mathfrak{p}$ are linearly
ependent over $\mathbb{F}_2$, $u_i = 0 \pmod 2$, or just simply $u_i = 0$ since $0 \leq u_i \leq 1$.
nce, the kernel of the multiplication map is trivial, i.e. $\prod_{i=1}^{m} H_i$ embeds into
$_K/\mathfrak{p}^2)^\times$.

xt, let s be the order of g_0 modulo $\mathfrak{p}^2$, so in particular, we have $g_0^s \equiv 1$
od $\mathfrak{p}$). Since g_0 is of order $N(\mathfrak{p}) - 1$ modulo $\mathfrak{p}$, s must be a multiple of
$\mathfrak{p}) - 1$, i.e. $s = [N(\mathfrak{p}) - 1]r$ for some $r \geq 1$. It is then true that g_0^r is of order
$\mathfrak{p}) - 1$ modulo $\mathfrak{p}^2$, so let G be the subgroup of $(\mathcal{O}_K/\mathfrak{p}^2)^\times$ generated by $g_0^r + \mathfrak{p}^2$.
ce the order of G is $N(\mathfrak{p}) - 1 = 2^m - 1$, which is relatively prime to the order
H_i ($|H_i| = 2$), we can argue similarly (like the above proof for odd prime)
t $G \times \prod_{i=1}^{m} H_i$ embeds into $(\mathcal{O}_K/\mathfrak{p}^2)^\times$ through the multiplication map. But
$\times \prod_{i=1}^{m} H_i| = |G| \prod_{i=1}^{m} |H_i| = 2^m[N(\mathfrak{p}) - 1] = N(\mathfrak{p})^2 - N(\mathfrak{p}) = |(\mathcal{O}_K/\mathfrak{p}^2)^\times|$,
the 2 are isomorphic

$$(\mathcal{O}_K/\mathfrak{p}^2)^\times \cong G \times \prod_{i=1}^{m} H_i \cong C_{N(\mathfrak{p})-1} \times (C_2)^m$$

or $t \geq 3$

1. $N(\mathfrak{p}) = 2$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^t) = 2^t - 2^{t-1} = 2^{t-1}$
   (b) $(\mathcal{O}_K/\mathfrak{p}^t)^\times \cong \langle -1 + \mathfrak{p}^t \rangle \times \langle 3 + \mathfrak{p}^t \rangle \cong C_2 \times C_{2^{t-2}}$
2. $N(\mathfrak{p}) = 4, t = 3$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^3) = 2^4 \cdot 3$
   (b) $\mathcal{O}_K/\mathfrak{p} = \mathrm{span}_{\mathbb{F}_2}\{1 + \mathfrak{p}, h + \mathfrak{p}\}$
   (c) $h^2 \equiv h + 1 \pmod{\mathfrak{p}}$
   (d)

$$\begin{aligned}
\langle -1 \rangle &= \{-1, 1\} \\
\langle 1 + 4h \rangle &= \{1, 1 + 4h\} \\
\langle 1 + 2h \rangle &= \{1, 1 + 2h, 5, -3 + 2h\} \\
(\mathcal{O}_K/\mathfrak{p}^3)^\times &= \langle -1 \rangle \cdot \langle 1 + 4h \rangle \cdot \langle 1 + 2h \rangle \cdot \langle h \rangle \cong (C_2)^2 \times C_4 \times C_3
\end{aligned}$$

$$\begin{aligned}
\langle -1 \rangle &= \{-1, 1\} \\
\langle 3 \rangle &= \{1, 3\} \\
\langle 1 + 2h \rangle &= \{1, 1 + 2h, 5, -3 + 2h\} \\
(\mathcal{O}_K/\mathfrak{p}^3)^\times &\neq \langle -1 \rangle \cdot \langle 3 \rangle \cdot \langle 1 + 2h \rangle \cdot \langle h \rangle
\end{aligned}$$

3. $N(\mathfrak{p}) = 4, t = 4$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^3) = 2^6 \cdot 3$
   (b) $\mathcal{O}_K/\mathfrak{p} = \mathrm{span}_{\mathbb{F}_2}\{1+\mathfrak{p}, h+\mathfrak{p}\}$
   (c) $h^2 \equiv h+1 \pmod{\mathfrak{p}}$
   (d)

$$\begin{aligned}\langle -1\rangle &= \{1,-1\}\\ \langle 1+4h\rangle &= \{1, 1+4h, 1+8h, 1-4h\}\\ \langle 1+2h\rangle &= \{1, 1+2h, 1+4h+4h^2, 1-2h+4h^2, 9, 9+2h, 9+4h+4h^2, 9-2h+4h^2\}\\ (\mathcal{O}_K/\mathfrak{p}^4)^\times &= \langle -1\rangle \cdot \langle 1+4h\rangle \cdot \langle 1+2h\rangle \cdot \langle h\rangle \cong C_2 \times C_4 \times C_8 \times C_3\end{aligned}$$

4. $N(\mathfrak{p}) = 4$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^t) = 2^{2(t-1)} \cdot 3$
   (b) $\mathcal{O}_K/\mathfrak{p} = \mathrm{span}_{\mathbb{F}_2}\{1+\mathfrak{p}, h+\mathfrak{p}\} \cong \mathbb{F}_4$
   (c) $h^3 \equiv 1, h^2 \equiv h+1 \pmod{\mathfrak{p}}$
   (d)

$$(\mathcal{O}_K/\mathfrak{p}^t)^\times = \langle -1\rangle \cdot \langle 1+4h\rangle \cdot \langle 1+2h\rangle \cdot \langle h\rangle \cong C_2 \times C_{2^{t-2}} \times C_{2^{t-1}} \times C_3 (-1)^u (1+2^{r+1}w(2^r w-1)+2^{r+1}w$$

5. $N(\mathfrak{p}) = 8, t = 3$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^3) = 2^6 \cdot 7$
   (b) $\mathcal{O}_K/\mathfrak{p} = \mathrm{span}_{\mathbb{F}_2}\{1+\mathfrak{p}, h+\mathfrak{p}, h^2+\mathfrak{p}\} \cong \mathbb{F}_8 \cong \mathbb{F}_2[h]/\langle h^3+h+1\rangle$
   (c) $(\mathcal{O}_K/\mathfrak{p})^\times = \langle h\rangle$
   (d)

$$\begin{aligned}\langle -1\rangle &= \{1,-1\}\\ \langle 3\rangle &= \{1,3\}\\ \langle 1+4h\rangle &= \{1, 1+4h\}\\ \langle 1+2h\rangle &= \{1, 1+2h, 1+4h+4h^2, 1-2h+4h^2\}\\ \langle 1+4h^2\rangle &= \{1, 1+4h^2\}\\ \langle 1+2h^2\rangle &= \{1, 1+2h^2, 1+4h, 1+4h+2h^2\}\\ (\mathcal{O}_K/\mathfrak{p}^3)^\times &= \langle -1\rangle \cdot \langle 3\rangle \cdot \langle 1+2h\rangle \cdot \langle 1+2h^2\rangle\\ &\cong (C_2)^2 \times (C_4)^2 \times C_7\end{aligned}$$

6. $N(\mathfrak{p}) = 2^m$:
   (a) $\varphi_{\mathcal{O}_K}(\mathfrak{p}^t) = 2^{m(t-1)} \cdot (2^m - 1)$

(b) $\mathcal{O}_K/\mathfrak{p} = \text{span}_{\mathbb{F}_2}\{h^i : 0 \leq i \leq m-1\}$

(c) $(\mathcal{O}_K/\mathfrak{p})^\times = \langle h \rangle$

(d)

$$\prod_{i=1}^{m-1}(1+2h^i)^{2^w y_i} \equiv 1 \pmod{\mathfrak{p}^{w+2}}$$

$$\sum_{i=1}^{m-1} y_i h^i + y_i(2^w y_i - 1)h^{2i}] \equiv 0 \pmod{\mathfrak{p}}$$

□

Temp Idea. *For $t \geq 3$, we can still obtain g_0 and h_i from the case $t = 2$. However, instead of using the basis $\beta = \{h_i + \mathfrak{p} : 1 \leq i \leq m\}$, we normalize it by multiply the inverse of h_m. Denote $e_i = h_i h_m^{-1}$, then $h_m^{-1}\beta = \{e_i + \mathfrak{p} : 1 \leq i \leq m\} = \{\mathfrak{e}_i : 1 \leq i \leq m\}$ is also a basis for $(\mathcal{O}_K/\mathfrak{p})^\times$ (over $\mathbb{F}_2$). Note that $e_i \not\equiv e_m = 1 \pmod{\mathfrak{p}}$ for all $1 \leq i \leq m-1$ since $\mathfrak{e}_i$ must be linearly independent over $\mathbb{F}_2$.

Let H_i be the subgroup of $1 + 2e_i + \mathfrak{p}^t$ (for $1 \leq i \leq m-1$). We prove that its order is 2^{t-1}. Indeed, since $2^{t-1} \geq 2^{t-2} \geq 2$, we can split the following sums

$$v_\mathfrak{p}\left(\binom{2^{t-1}}{k}2^k e_i^k\right) \geq t - 1 - v_2(k) + k \geq t+1 \text{ for } k \geq 3$$

$$\begin{aligned}(1+2e_i)^{2^{t-1}} &= \sum_{k=0}^{2^{t-2}}\binom{2^{t-2}}{k}2^k e_i^k \\ &= 1 + 2^t e_i + 2^t(2^{t-1}-1)e_i^2 + \sum_{k=3}^{2^{t-2}}\binom{2^{t-2}}{k}2^k e_i^k \\ &\equiv 1 \pmod{\mathfrak{p}^t}\end{aligned}$$

$$v_\mathfrak{p}\left(\binom{2^{t-2}}{k}2^k e_i^k\right) \geq t - 2 - v_2(k) + k \geq t \text{ for } k \geq 3$$

$$\begin{aligned}(1+2e_i)^{2^{t-2}} &= 1 + 2^{t-1}e_i + 2^{t-1}(2^{t-2}-1)e_i^2 + \sum_{k=3}^{2^{t-2}}\binom{2^{t-2}}{k}2^k e_i^k \\ &\equiv 1 + 2^{t-1}e_i[1 + (2^{t-2}-1)e_i] \pmod{\mathfrak{p}^t}\end{aligned}$$

In the last congruence, we have $e_i \not\equiv e_m = 1 \pmod{\mathfrak{p}}$, so $1 + (2^{t-2}-1)e_i \not\equiv$

$1 + (2^{t-2} - 1) \equiv 0 \pmod{\mathfrak{p}}$. In other words, $v_{\mathfrak{p}}(1 + (2^{t-2} - 1)e_i) = 0$, and thus

$$\begin{aligned} v_{\mathfrak{p}}\left(2^{t-1}e_i[1 + (2^{t-2} - 1)e_i]\right) &= v_{\mathfrak{p}}(2)v_2(2^{t-1}) + v_{\mathfrak{p}}(e_i) + v_{\mathfrak{p}}(1 + (2^{t-2} - 1)e_i) \\ &= t - 1 < t \end{aligned}$$

$$(1 + 2e_i)^{2^{t-2}} \equiv 1 + 2^{t-1}e_i[1 + (2^{t-2} - 1)e_i] \not\equiv 1 \pmod{\mathfrak{p}^t}$$

Let E be the (exceptional) subgroup of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$ generated by $2e_m + 1 + \mathfrak{p}^t = 3 + \mathfrak{p}^t$. By lemma 14.26, $v_2\left(3^{2^k} - 1\right) = k + 2$, so $v_{\mathfrak{p}}\left(3^{2^{t-2}} - 1\right) = v_{\mathfrak{p}}(2)v_2\left(3^{2^{t-2}} - 1\right) = t$ and $v_{\mathfrak{p}}\left(3^{2^{t-3}} - 1\right) = t - 1 < t$. This should be enough to show that the order of 3 is 2^{t-2} modulo $\mathfrak{p}^t$. Denote $S = \{\pm 1 + \mathfrak{p}^t\}$ to be the (cylic) subgroup of $(\mathcal{O}_K/\mathfrak{p}^t)^\times$.

We show that the multiplication map $S \times E \times \prod_{i=1}^{m-1} H_i \to (\mathcal{O}_K/\mathfrak{p}^t)^\times$ is an embedding. Consider an element of the kernel: there exists some $0 \le z \le 1, 0 \le y_m < 2^{t-2}$ and $0 \le x_i < 2^{t-1}$ such that

$$(-1)^{-z}3^{-y_m}\prod_{i=1}^{m-1}(1 + 2e_i)^{x_i} \equiv 1 \pmod{\mathfrak{p}^t}$$

$$\prod_{i=1}^{m-1}(1 + 2e_i)^{x_i} \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^t}$$

Suppose x_i are not simultaneously zero, then $w = v_2(\gcd(x_1, x_2, \ldots, x_{m-1})) \ge 0$ and y_i are not simultaneously divisible by 2, where $x_i = 2^w y_i$. Since $x_i < 2^{t-1}$, $w < t - 1$, so we can reduce the power of the modulus from t to $w + 2$.

$$\prod_{i=1}^{m-1}(1 + 2e_i)^{x_i} \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^{w+2}}$$

$$\prod_{i=1}^{m-1}(1 + 2e_i)^{2^w y_i} \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^{w+2}}$$

$$\prod_{i=1}^{m-1}\sum_{k=0}^{2^w y_i}\binom{2^w y_i}{k}2^k e_i^k \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^{w+2}}$$

But $v_{\mathfrak{p}}\left(\binom{2^w y_i}{k}2^k e_i^k\right) \ge w - v_2(k) + k \ge w + 2$ for $k \ge 3$, so consider each factor on the left

1. If $x_i = 2^w y_i = 0$: then $y_i = 0$, and so the sum turns into

$$\sum_{k=0}^{2^w y_i}\binom{2^w y_i}{k}2^k e_i^k = 1 = 1 + 2^{w+1}y_i e_i\left[1 + (2^w y_i - 1)e_i\right]$$

2. If $x_i = 2^w y_i = 1$: then $w = 0$ and $y_i = 1$. The sum turns into

$$\sum_{k=0}^{2^w y_i} \binom{2^w y_i}{k} 2^k e_i^k = 1 + 2e_i = 1 + 2^{w+1} y_i e_i \left[1 + (2^w y_i - 1)e_i\right]$$

3. If $x_i = 2^w y_i \geq 2$: then we split the sum into

$$\begin{aligned}\sum_{k=0}^{2^w y_i} \binom{2^w y_i}{k} 2^k e_i^k &= \sum_{k=0}^{2} \binom{2^w y_i}{k} 2^k e_i^k + \sum_{k=3}^{2^w y_i} \binom{2^w y_i}{k} 2^k e_i^k \\ &\equiv 1 + 2^{w+1} y_i e_i + 2^{w+1} y_i (2^w y_i - 1) e_i^2 \\ &\equiv 1 + 2^{w+1} y_i e_i \left[1 + (2^w y_i - 1)e_i\right] \pmod{\mathfrak{p}^{w+2}}\end{aligned}$$

In any case, the previous congruence turns into

$$\prod_{i=1}^{m-1} \left[1 + 2^{w+1} y_i e_i \left(1 + (2^w y_i - 1)e_i\right)\right] \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^{w+2}}$$

Distributing the factors in the product on LHS, then each term with a factor of the form $(2^{w+1} y_i e_i)(2^{w+1} y_j e_j)$ (for $i \neq j$) will be cancelled modulo $\mathfrak{p}^{w+2}$, since the $\mathfrak{p}$-adic valuation is at least $2(w+1) \geq w+2$ (note that $w \geq 0$). Therefore,

$$1 + 2^{w+1} \sum_{i=1}^{m-1} y_i e_i \left[1 + (2^w y_i - 1)e_i\right] \equiv (-1)^z 3^{y_m} \pmod{\mathfrak{p}^{w+2}}$$

□

Chapter 15

Analytic Number Theory

Theorem 15.1 (Möbius inversion formula)**.** The Möbius function

$$\mu(n) = \begin{cases} 1 & \text{if } n \text{ is square-free with an even number of prime factors} \\ -1 & \text{if } n \text{ is square-free with an odd number of prime factors} \\ 0 & \text{otherwise} \end{cases}$$

is the multiplicative inverse of the identity function $1(n) = 1$, with respect to the Dirichlet convolution

$$f * g(n) = \sum_{d|n} f(d)g(n/d)$$

Note that $\varepsilon(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$ is the multiplicative identity with respect to the convolution.

Chapter 16

Graph Theory

Theorem 16.1 (Kuratowski theorem)**.** A finite graph is planar if and only if it does not contain a subgraph that is a subdivision of K_5 or $K_{3,3}$.

Theorem 16.2 (Hall marriage theorem)**.** Let G be a finite bipartite graph with bipartite sets X and Y.

1. An X-perfect matching (or saturated matching) is an injective map $f : X \to Y$ such that x is adjacent to $f(x)$.
2. For every subset W of X, denote $N_G(W)$ to be the neighborhood of W in G (i.e. set of all vertices in Y adjacent to some vertex in W).

Then G has an X-perfect matching if $|W| \leq |N_G(W)|$ for every subset W of X.

Theorem 16.3 (Turán theorem)**.**

Theorem 16.4 (Tutte-Berge formula)**.**

Theorem 16.5 (Ramsey theorem)**.**

Chapter 17

Combinatorics

Theorem 17.1 (Helly theorem)**.**

Theorem 17.2 (Pigeonhole principle)**.** If $f : A \to B$ is a function between 2 finite set A, B such that $|A| = n > k = |B|$, then there exists some element b of B such that $f^{-1}(b)$ has at least $\lceil n/k \rceil$ elements.