Introducing Rigorous Mathematics

Tony Ma

2020-08-25 (Begin)

Preface

It is often important though not easy to learn to think and write rigorously. I would like to write some notes to introduce mainly Mathematical Analysis, which is, loosely speaking, the theory of calculus. The famous book *Principles of Mathematical Analysis* by *Walter Rudin* is very fascinating to me, and I am kind of rewriting the book in e-format and adding **tons of** my own notes, including many motivations (which may possibly be really wordy) to favor my studies and so as yours. Hope you will appreciate.
(As you may already see, I would use color gray when writing some remarks / not important statements.)

1 Introduction

1.1 Very little set theory

Definition 1.1. We shall use $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \overline{\mathbb{R}}, \mathbb{C}$ often, described later.

Definition 1.2. A logical statement $A \Rightarrow B$ is equivalent to its *contrapositive*, i.e. $\neg B \Rightarrow \neg A$ ($\neg$ is negation).

Definition 1.3. $\exists! x : \ldots$ means it exists unique x such that $\ldots$. ('|', ':' both mean 's.t.' (such that).)

Definition 1.4. We say sets $A \subseteq B$ if $x \in B \ \forall x \in A$. We say $A = B$ if $A \subseteq B \wedge B \subseteq A$, else $A \neq B$.

Corollary 1.5. $A \neq B \iff (A \nsubseteq B \vee B \nsubseteq A)$

Definition 1.6. We call A a *proper subset* of B if $A \subseteq B$ and $\exists x \in B | x \notin A$, denoted by $A \subsetneq B$.

1.2 Little about groups

Definition 1.7. For **a pair of** a set G and a binary operation $\diamond : G \times G \to G$, denoted by $(G, \diamond)$, we define:

1. *Closure*: $\forall a, b \in G, a \diamond b \in G$;
2. *Associativity*: $\forall a, b, c \in G, a \diamond (b \diamond c) = (a \diamond b) \diamond c$;
3. *Identity element*: $\exists e \in G : \ \forall a \in G, e \diamond a = a \diamond e = a$;
4. *Inverse element*: $\forall a \in G, \exists b \in G : b \diamond a = a \diamond b = e$;
5. *Commutativity*: $\forall a, b \in G, a \diamond b = b \diamond a$.

With the above conditions (or *axioms*), $(G, \diamond)$ is called

- *Semi-group*, if conditions 1 and 2 are satisfied;
- *Group*, if conditions 1 to 4 are satisfied;
- *Abelian group*, if conditions 1 to 5 are satisfied.

It is worth noticing that in the statement '$\forall a, b \in G$', a, b may be the same, so you shall understand how $(\{0\}, +)$, where $+$ is the ordinary addition of integer, defines a group. Also, we no longer need to add brackets for $a \diamond b \diamond c$ because by associativity it is clear what we are talking about. Associativity is actually more useful in many sense than commutativity, so for a group, commutativity is only a bonus.
Notice that 'e' in condition 4 represents **the identity element in condition 3** (more precisely, '$\exists e$ such that both conditions 3 and 4 are satisfied'), unless after we prove the uniqueness of identity element in a group then it is clear. Also, condition 1 is not necessary to be stated **if you** already defined $\diamond : G \times G \to G$, but it is helpful to mention.

Definition 1.8. For convention, in a group, we denote, for any $a \in G$, $a^n = \underbrace{a \diamond \cdots \diamond a}_{n}$ $\forall n \in \mathbb{N}$ and $a^0 = e$.

Theorem 1.9. *For a group $(G, \diamond)$, the identity element is unique.*

Proof. If e_1, e_2 are two identity elements of $(G, \diamond)$, then $e_1 = e_1 \diamond e_2 = e_2$.
Now there are no two **different** identity elements. ■

Notice that both $e \diamond a = a$ and $a \diamond e = a$ is used in the above proof but we need not to use commutativity.

Theorem 1.10. *For a group $(G, \diamond)$, the inverse element to any element is unique.*

Proof. For $a \in G$, if b_1, b_2 are two inverse elements of a, then $b_1 = (b_2 \diamond a) \diamond b_1 = b_2 \diamond (a \diamond b_1) = b_2$ ■

Notice that both $b \diamond a = e$ and $a \diamond b = e$ is used in the above proof.

Remark 1.11. To prove these elementary results in group theory, simply manipulate the 4 existing 'rules'.

Next we do little investigation on the necessity of to include both left and right identity element / inverse element.

Definition 1.12. In Definition 1.7,

- In Condition 3, $e \in G$ is called an *left identity element* (e is on the left) if $\forall a \in G, e \diamond a = a$;
- In Condition 4, (given, in condition 3, a left / right / both-sided inverse element e), $\forall a \in G, b$ is called a *left inverse element* if $b \diamond a = e$.

  (Likewise for *right identity element* and *right inverse element*)

Definition 1.13. For a semi-group $(G, \diamond)$, $a \in G$ is called *idempotent* if $a \diamond a = a$.

Theorem 1.14. *For a semi-group $(G, \diamond)$, if any of the followings holds:*

1. *Left identity element exists and for such e all elements has left inverse element, or*
2. *Right identity element exists and for such e all elements has right inverse element, or*
3. *$(G, \diamond)$ is a group,*

then $\big(a \diamond a = a \iff a = e\big)$ $\forall a \in G$ and $(G, \diamond)$ is a group.

Proof. WLOG we shall only consider the first case, for all $a \in G$, there exists **at least** one left inverse element b, $\big(a \diamond a = a \Rightarrow a = (b \diamond a) \diamond a = b \diamond a = e\big)$, and btw implying the uniqueness of left identity element e, $(a \diamond b) \diamond (a \diamond b) = a \diamond (b \diamond a) \diamond b = a \diamond b$ and hence b is also right inverse element of a, now $a \diamond e = a \diamond b \diamond a = e \diamond a = a$, satisfying all the axioms required for a group. ■

As you may see, associativity is frequently used in the above proof.

Corollary 1.15. *In a semi-group $(G, \diamond)$ with left identity element and **all**-left inverse element (**corresponding to** the left identity element), then $(G, \diamond)$ is a group (may replace both 'left' by 'right').*

However, 'left identity element' & 'all-right inverse element' is not sufficient, consider the following example.

Example 1.16. Let $G = \{0, 1\}$ (or any other 2 symbols defined to be **distinct**), and define operation $\diamond$ s.t. $a \diamond b = b \ \forall a, b \in G$. We can show that $(G, \diamond)$ is a semi-group with left identity element and all-right inverse element corresponding to that identity element, but it is not a group.

Proof. Closure and associativity (roughly speaking always equal to the last term in expression) is trivial. We can take 0 to be the left-identity element and $a \diamond 0 = 0 \ \forall a \in G$. But you can easily check that it is not a group by the axioms, or simply noticing the non-uniqueness of left identity element (contrapositively, if it is a group, then by Theorem 1.14 the left identity must be unique). ■

Exercise 1.17. Where does the proof of Theorem 1.14 fail with the above example?

Now we can move on from studying group axioms.

1.3 Definition of field

Definition 1.18. A set of F with binary operations $+ : F \times F \to F$ and $\times : F \times F \to F$, denoted by $(F, +, \times)$, is called a *field* if:

1. $(F, +)$ is an Abelian group (let 0 be the identity element of $+$), and
2. $(F \setminus \{0\}, \times)$ is an Abelian group (let 1 be the identity element of $\times$ and denote $ab = a \times b$), and
3. *Distributivitiy*: $\forall a, b, c \in F$, $(a + b)c = ac + bc$ and $a(b + c) = ab + ac$.

Fields are define in such way so that we can study the algebraic structure just like $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ we are used to.

Example 1.19. Consider $F = \{0, 1\}$ with binary operations $+ : F \times F \to F$ and $\times : F \times F \to F$ such that $\forall a, b \in F, (a + b = a \text{ xor } b)$ and $(a \times b = a)$. If only the first part of *distributivity* (above definition) is required, then $\times$ is generally not commutative here. (a xor b is simply $a + b \bmod 2$.)

Proof. You are suggested to draw operation tables for $+$ and $\times$, and notice that all the criteria in Definition 1.18 is satisfied except the second *distributivity*, and $0 \times 1 = 0 \neq 1 = 1 \times 0$. ■

Notice that in the definition of field we consider Abelian group $(F \setminus \{0\}, \times)$, so we are actually missing general commutativity and associativity of $\times$.

Definition 1.20. In Definition 1.18, for $(F, +, \times)$ the following definition of *field* is equivalent:

1. $\forall a, b \in F, a + b, a \times b \in F$, and
2. $\forall a, b \in F, a + b = b + a, a \times b = b \times a$, and

3. $\forall a,b,c \in F, a+(b+c)=(a+b)+c, a\times(b\times c)=(a\times b)\times c$, and
4. $\exists!0 \in F : \big(\forall a \in F, a+0=a\big)$; $\exists!1 \in F \wedge 1 \neq 0 : \big(\forall a \in F, a\times 1=a\big)$, and
5. $\forall a \in F, \big(\exists! b \in F : a+b=0\big)$; $\forall a \in F \wedge a \neq 0, \big(\exists! c \in F : a\times c=1\big)$, and
6. *Distributivity* (either one of the two is enough).

Notice again that it is a convention to use symbols $+, \times$ and $0, 1$ for a field, and every '$\exists!$' above is not necessary as you know how to prove the uniqueness after existence, but the above is just an alternative way to write the axioms which you should **understand** what all these actually mean. Also, we need not require identity or inverse element, nor distributivity to be both-sided because commutativity of $+$ and $\times$ are required already, and by this reason I put commutativity early than most axioms of the above (just I want to).

Proof. To see how the definitions are equivalent, we shall use concept / trick like Definition 1.13. To prove equivalence, we shall prove that the first definition imply the second and vice versa. First of all, if $(F,+,\times)$ is a field defined by Definition 1.18, then we can show $\forall a \in F, 0\times a = a\times 0 = 0$ by $0\times a = (0+0)\times a = 0\times a + 0\times a$ and similarly $a\times 0 = a\times(0+0) = a\times 0 + a\times 0$. We notice that the commutativity and associativity of multiplication holds generally now (originally only commutative and associative in $F\setminus\{0\}$). On the other hand, if it is defined by Definition 1.20, then the other part of distributivity is easily provable by the given part (as multiplication is generally commutative in this definition). ■

Corollary 1.21. *For a field $(F,+,\times)$, the second distributivity can be replaced by '$\forall x \in F,\ x\times 0 = 0$'.*

Theorem 1.22. *If $(F,+,\times)$ is a field, then $ab=0 \iff (a=0 \vee b=0)$ for any $a,b \in F$.*

Proof. Contrapositively, if non-zero $a, b \in F$, assume in contrary, then

$$1 = \Big(\frac{1}{b}\Big)\Big(\frac{1}{a}\Big)(a)(b) = (...)(0) = 0.$$

■

I hope you appreciate the proof of $0\times a = a\times 0 = 0$, which is only true after distributive laws are introduced.

1.4 Totally ordered sets

Definition 1.23. For a set S, a *total strict order* $<$ on S is an operation (giving 'T/F') with the followings satisfied:

1. *Trichotomous*: $\forall x,y \in S$, **one and only one** of the following holds

   $x<y, x=y, y<x$; and
2. *Transitivity*: $\forall x,y,z \in S, \big((x<y)\wedge(y<z)\Rightarrow(x<z)\big)$.

   Also, S is then called a *totally ordered set* when such $<$ is defined.

Notice that we should use 'strict' order here, that is, $x<y \Rightarrow x\neq y$. Also, $x>y$ and $x\leq y$ are defined to mean $y<x$ and $(x<y)\vee(x=y)$ respectively. We can use something like $(S,<)$, but as we always use $<$ to represent the order, one may simply write '(totally) ordered set S'.

Definition 1.24. For a set S, Definition 1.23 of *total strict order* is equivalent to having followings satisfied:

1. *Irreflexivity*: $\forall x \in S, x \not< x$, and

2. The order is *total*: $\forall x, y \in S, (x < y) \vee (y < x) \vee (x = y)$, and

3. *Transitivity*.

Exercise 1.25. Prove that the two definitions above are equivalent (holds if and only if each other).
Hint: Transitivity can be used to achieve a contradiction somehow.

Example 1.26. $\mathbb{Q}$ is a totally ordered set where $a < b$ means $b - a \in \mathbb{Q}^+$ for $a, b \in \mathbb{Q}$.

Definition 1.27. For a totally ordered set S and $E \subseteq S$, if $\exists \alpha \in S : (\forall x \in E, x \leq \alpha)$, we say E is *bounded above* and α is an *upper bound* of E (Likewise for *lower bound*).

Example 1.28. An empty set as a subset of any totally ordered set S is bounded above and below by any $\alpha \in S$.

Notice that the upper bound of a subset of totally ordered set is generally not unique, so we would like to consider the smallest one among all the upper bounds.

Definition 1.29. For a totally ordered set S and $E \subseteq S$, if for some $\alpha \in S$ such that:

1. α is an upper bound of E, **and**

2. $\forall \beta \in S | \beta < \alpha, \beta$ is **not** an upper bound of E.

then α is called the *least-upper-bound* or *supremum* of E, denoted by $\alpha = \sup E$.

Exercise 1.30. In the above definition, show that the supremum of such E is unique **if exists**.

Infimum, or *greatest-lower-bound* is defined in the same manner (with 'upper bound' replaced by 'lower bound').

Example 1.31. Consider the set $A = \left\{ \frac{1}{n} \middle| n \in \mathbb{N} \right\}$, there is no lower bound **larger** than 0, and $\inf A = 0, 0 \notin A$

Definition 1.32. A totally ordered set S has *least-upper-bound property* if any **non-empty** and **bounded above** $E \subseteq S$ has a *supremum* ($\sup E$ exists in S).

Remark 1.33. $\mathbb{R}$ is a totally ordered field which we shall prove it later. The point of considering (total strict) order is that our common $\mathbb{Q}, \mathbb{R}$ are not only fields. The property to have any two elements comparable (order-able) in totally ordered field is not very obviously granted, which, for example, a strict order cannot be defined in $\mathbb{C}$, which we shall explain later. And the importance of $\mathbb{R}$ (instead of only $\mathbb{Q}$) can be explained by the amazing *least-upper-bound property* of $\mathbb{R}$, useful in further *Analysis*. Through construction of number systems (from $\mathbb{N}$ to $\mathbb{C}$) we know more about how numbers work, and we learn how numbers work by developing (basic) algebra and (basic) analysis.

Greatest-lower-bound property is defined likewise and it has a close relation with *least-upper-bound property*.

Theorem 1.34. *A totally ordered set S has least-upper-bound property if and only if it has greatest-lower-bound property.*

Proof. We shall prove the forward direction ($\Rightarrow$) WLOG. For any $E \subseteq S$ non-empty and bounded below, we let $F \subseteq S$ (all the things we discuss are in S as background) be all lower bounds of E, which F is non-empty (as E is bounded below) and bounded above. Any element in F is a lower bound of E by definition and any element in E is an upper bound of F (E is subset the set of upper bound of F but may not be equal to). Or consider $\forall x \in E \wedge y \in F, x \leq y$. Let $\alpha = \sup F$, the **key** of the proof would be $(\beta < \alpha \Rightarrow \beta \notin E)$ and $(\gamma > \alpha \Rightarrow \gamma \notin F)$. The first one is because β is not an upper bound of F with α being the least, and being not an upper bound of F means that $\beta \notin E$; The second one is obvious as α is an upper bound of F. Notice that we use both conditions in Definition 1.29 for $\alpha = \sup F$. Now, $\beta \in E$ implies $\beta \geq \alpha$ (contrapositively) and $\gamma > \alpha$ means γ is not a lower bound of E (by definition of F) respectively, and now we arrive at $\alpha = \inf E$. Notice how $\sup F$ and $\inf E$ are so nicely related to each other and how delicate and symmetric(?) the proof is. Also, btw, notice that for elements in S, smaller than α means not being in E and larger than α means not being in F, where $\alpha \in F$ ($\alpha = \sup F = \max F$) and α may or may not be in E. The above proof are words to help you understand the proof but you should really try to find the key (grasp the essence) of the proof. ■