baby-rudin reading notes

weiyuan

2022 年 11 月 20 日

Chapter 1

the real and complex number system

1.1 Introduction

First we use $\sqrt{2}$ to construct real number system from integer and rational numbers.

Example 1.1.

$$p^2 = 2 \tag{1.1}$$

p is not a rational number.

证明. (反证法) 假设 p 是有理数, $\exists m, n \in \mathbf{N}$, s.t. $p = m/n$. $\gcd(m, n) = 1$. Then 1.1

$$m^2 = 2n^2. \tag{1.2}$$

m is even, $m = 2k$. 那么有 $(2k)^2 = 2n^2$, $2k^2 = n^2$, k is even, $\gcd(m, n) = 2 \neq 1$, contrary to our choice of m and n. Hence p can't be a rational number. $\square$

After proving $\sqrt{2}$ isn't a rational number, rudin use $\sqrt{2}$ to divide the rationals 在证明 $\sqrt{2}$ 不是有理数后, 使用 $\sqrt{2}$ 将有理数集分成两部分. 引出了分划的概念?

$$A = \{p | p^2 < 2\}$$
$$B = \{p | p^2 > 2\}$$

A contains no largest number,
B contains no smallest number.
$\forall p \in A$, $\exists q \in A$, s.t. $p < q$,
$\forall p \in B$, $\exists q \in B$, s.t. $p > q$,
$\forall p > 0$

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2} \tag{1.3}$$

Then

$$q^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2} \tag{1.4}$$

If $p \in A$, $p^2 < 2$. 1.3 shows that $q > p$, 1.4 shows that $q^2 < 2$, $q \in A$. If $p \in B$, $p^2 > 2$. 1.3 shows that $q < p$, 1.4 shows that $q^2 > 2$, $q \in B$.

Remark 1.2. The purpose of the above discussion has been to show that the rational number system has certain gaps, in spite of the fact that between any two rationals there is another: If $r < s$ then $r < (r + s)/2 < s$. The real number system fills these gaps. This is the principal reason for the fundamental role which it plays in analysis.

> mynotes:
> 有理数的稠密性与实数的连续性. 在分析中, 考察极限等需要的是数系的连续性, 因此需要先建立实数系. 事实上, 我们是先有微积分, 后有实数理论的.
> 三次数学危机：无理数, 微积分基础, 集合论实数理论是极限的基础.

In order to elucidate its structure, as well as that of the complex numbers, we start with a brief discussion of the genral concepts of *ordered set* and *field*.

> mynotes:
> rudin 引入复数的方法非常怪, 对初学者非常不友好, 过于抽象了. 想起一个法国笑话, 问小学生 $2 + 3$ 等于几, 回答 $2 + 3 = 3 + 2$ 加法是一个交换群（Abel 群）...

Here is some of the standard set-theoretic terminology taht will be used throughout this book.

> mynotes:
> 接下来引入一些集合论的定义

Definition 1.3. If A is any set (whose elements[1] may be numbers or any other objects[2]), we write $x \in A$ to indicate that x is a member (or an element) of A.

If x is not a member of A, we write: $x \notin A$.

empty set $\varnothing$ contains no element, If a set has at least one element, it is called *nonempty*.

A, B are sets, $\forall x \in A$, $x \in B$, we say that A is a *subset* of B, $A \subset B$ or $B \supset A$. If $\exists x \in B$, $x \notin A$, A is a *proper subset* of B, $A \subsetneqq B$. Note that $A \subset A$ for every set A.

(Bernstein) If $A \subset B$ and $B \subset A$, we write $A = B$. Otherwise $A \neq B$.

> mynotes:
> 这条性质在证明集合相等时很常用

Definition 1.4. Throughout Chap. 1, the set of all rational numbers will be denoted by $\mathbb{Q}$.

有理数集 $\mathbb{Q}$

1.2 Ordered sets

有序集

[1] 这里 elements 还没定义, 笑啦

[2] object 指代什么？我个人认为集合理解的难点在于集合的集合. 这一点可以引出罗素悖论

Definition 1.5. Let S be a set. An *order* on S is a relation, denoted by $<$, with the following two properties:

(i) If $x \in S$ and $y \in S$ then one and only one of the statements

$$x < y, \quad x = y, \quad y < x$$

The statement $x < y$ may be read as x is less than y, or x is smaller than y, or x precedes y. (It's often convenient to write $y > x$ in place of $x < y$) (less-great, smaller-bigger, precedes-succeeds)

(ii) If $x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$.

$x \leq y$ indicates taht $x < y$ or $x = y$, without specofying which of these two is to hold. In other words, $x \leq y$ is the negation of $x > y$.

> mynotes:
> 偏序关系: 1. 三歧性, 2. 传递性.
> 建立偏序关系后, 可以使用不等式进行分析. 在后续根据极限定义计算时, 需要大量使用不等式分析数列和函数的极限计算结果.

Definition 1.6. An *ordered set* is a set S in which an order is defined.

For Example, $\mathbb{Q}$ is an ordered set if $r < s$ is defined to mean that $s - r$ is a positive rational number.

> mynotes:
> 存在偏序关系的集合称为有序集 $\mathbb{Q}, \mathbb{R}$ 均是有序集, 但 $\mathbb{C}$ 不是有序集.

Definition 1.7. Suppose S is an ordered set, and $E \subset S$. If there exists a $\beta \in S$ such that $x \leq \beta$ for every $x \in E$, we say that E is *bounded above*, and call β an *upper bound* of E.

Lower bounds are defined in the same way (with $\geq$ in place of $\leq$).

Definition 1.8. Suppose S is an ordered set, $E \subset S$, and E is bounded above. Suppose there exists an $a\alpha \in S$ with the following properties:

(i) α is an upper bound of E. (ii) If $\gamma < \alpha$ then γ is not an upper bound of E.

Then α is called the *least upper bound* of E [that there is at most one such α is clear from (ii)] or the *supremum* of E, and we write

$$\alpha = \sup E.$$

The *greatest lower bound*, or *infimum*, of a set E which is bounded below is defined in the same manner: The statement

$$\alpha = \inf E$$

means that α is a lower bound of E and that no β with $\beta > \alpha$ is a lower bound of E.

> mynotes:
> 从上界引出最小上界, 没有直接定义最大下界, 而是使用对称定义引出. 从最小上界引出的最小上界性质更为常用. Dedekind 分划

Example 1.9. (a) Consider the set A, B

$$A = \{p | p^2 < 2\}, \quad B = \{p | p^2 > 2\}.$$

A has no least upper bound in $\mathbb{Q}$. B has no great lower bound in $\mathbb{Q}$.

(b) If $\alpha = \sup E$ exists, α may be or may not be a member of E.

$$E_1 = \{r | r \in Q, r < 0\}$$
$$E_2 = \{r | r \in Q, r \leq 0\}$$

$$\sup E_1 = \sup E_2 = 0,$$

and $0 \notin E_1$, $0 \in E_2$.

(c) $E = \{1/n | n = 1, 2, 3, ...\}$. Then $\sup E = 1$, which is in E, and $\inf E = 0$, which is not in E.

Definition 1.10. least-upper-bound property

An ordered set S is said to have the *least-upper-bound property* if the following is true:

If $E \subset S$, E is not empty, and E is bounded above, then $\sup E$ exists in S.

Example 1.9(a) shows that $\mathbb{Q}$ does not have the least-upper-bound property.

We shall now show that there is a close relation between greatest lower bounds and least upper bounds, and that every ordered set with the least-upper-bound property also has the greatest-lower-bound property.

Theorem 1.11. *Suppose S is an ordered set with the least-upper-bound property, $B \subset S$, B is not empty, and B is bounded below. Let L be the set of all lower bounds of B. Then*

$$\alpha = \sup L$$

exists in S, and $\alpha = \inf B$.

In particular, $\inf B$ exists in S.

证明. Since B is bounded below, L is not empty. Since L consists of exactly those $y \in S$ which satisfy the inequality $y \leq x$ for every $x \in B$, we see that *every $x \in B$ is an upper bound of L*. Thus L is bounded above. Our hypothesis about S implies therefore that L has a supremum in S; call it α.

If $\gamma < \alpha$ then (see Definition 1.8) γ is not an upper bound of L, hence $\gamma \notin B$. It follows that $\alpha \leq x$ for every $x \in B$. Thus $\alpha \in L$.

If $\alpha < \beta$ then $\beta \notin L$, since α is an upper bound of L.

We have shown that $\alpha \in L$ but $\beta \notin L$ if $\beta > \alpha$. In other words, α is a lower bound of B, but α is not if $\beta > \alpha$. This means that $\alpha = \inf B$. □

mynotes:
这个证明第一次看比较难理清我试着用自己的话重写梳理一下：已知条件 S, ordered set + least-upper-bound property. $B \in S$, $B \neq \varnothing$, B is bounded below. L is the set of all lower bounds of B. $\exists \alpha \in S$, $\alpha = \sup L$, and $\alpha = \inf B$.

证明. 思路由最小上界 $\rightarrow$ 最大下界 $L = \{y | y \in S; \forall x \in B, y \leq x\}$ 关于 L 中有没有不在 S 中的元素这一点我还没想明白. 定理中只是说 L 是 B 的下界组成的. B 是 S 的子集, 但 B 的下界不一定全在 S 中.

L 由 B 在 S 中的全部下界组成

$\forall x \in B, x$ 为 L 的上界. $L \subset S$. S 有最小上界性质, $\therefore \exists \alpha \in S, \alpha = \sup L$.

$\forall \gamma < \alpha$ 由 $\alpha = \sup L$ 的定义 (1.8) γ 不是 L 的上界.

$\forall x \in B, x$ 为 L 的上界, $x \geq \alpha$. $\therefore \alpha \in L$.

$\alpha < \beta, \alpha = \sup L$. $\therefore \beta \notin L$. L 由 B 在 S 中的全部下界组成, $\beta \notin L$. β 不是 B 的下界.

$\therefore \alpha = \inf B, \inf B \in S$. □

1.3 fields

mynotes:
域, 交换除环 $<\mathbb{R}, +, \times>$ $<\mathbb{R}, +>$, $<\mathbb{R}\backslash\{0\}, \times>$ 都是交换群, 且满足分配律. 则 $<\mathbb{R}, +, \times>$ 是域.

Definition 1.12. (A) Axioms for addition

(Al) If $x \in F$ and $y \in F$, then their sum $x + y$ is in F.

(A2) Addition is commutative: $x + y = y + x$ for all $x, y \in F$.

(A3) Addition is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$.

(A4) F contains an element 0 such that $0 + x = x$ for every $x \in F$.

(A5) To every $x \in F$ corresponds an element $-x \in F$ such that

$$x + (-x) = 0.$$

(M) Axioms for multiplication

(M1) If $x \in F$ and $x \in F$, then their product xy is in F.

(M2) Multiplication is commutative: $xy = yx$ for all $x, y \in F$.

(M3) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in F$.

(M4) F contains an element $1 \neq 0$ such that $1x = x$ for every $x \in F$.

(M5) If $x \in F$ and $x \neq 0$ then there exists an element $1/x \in F$ such that

$$x \cdot (1/x) = 1.$$

(D) The distributive law

$$x(y + z) = xy + xz$$

holds for all $x, y, z \in F$.

Remark 1.13. (a) Our usual writes (in any filed)

只定义了加法和乘法, 使用逆元分别表示减法和除法. $x-y=x+(-y), x/y=x\cdot(1/y)$.

(b) The field axioms clearly hold in $\mathbb{Q}$, the set of all rational numbers, if addition and multiplication have their customary meaning. Thus $\mathbb{Q}$ is a field.

全体有理数的集合是一个域.

(c) Although it is not our purpose to study fields (or any other algebraic structures) in detail, it is worthwhile to prove that some familiar properties of $\mathbb{Q}$ are consequences of the field axioms; once we do this, we will not need to do it again for the real numbers and for the complex numbers.

Proposition 1.14. The axioms for addition imply the following statements.

(a) If $x+y=x+z$ then $y=z$.

(b) If $x+y=x$ then $y=0$.

(c) If $x+y=0$ then $y=-x$.

(d) $-(-x)=x$.

Statement (a) is a cancellation law. Note that (b) asserts the uniqueness of the element whose existence is assumed in (A4), and that (c) does the same for (A5).

mynotes:
what is the difference between axiom and proposition?
An axiom is a proposition regarded as self-evidently true without proof. The word "axiom" is a slightly archaic synonym for postulate. Compare conjecture or hypothesis, both of which connote apparently true but not self-evident statements. A proposition is a mathematical statement such as "3 is greater than 4," "an infinite set exists," or "7 is prime." An axiom is a proposition that is assumed to be true. With sufficient information, mathematical logic can often categorize a proposition as true or false, although there are various exceptions (e.g., "This statement is false"). `https://www.nutritionmodels.com/terminology.html`

证明. Proof(rudin)

If $x+y=x+z$, the axioms (A) give

$$\begin{aligned} y = 0 + y &= (-x+x)+y = -x+(x+y) \\ &= -x+(x+z) = (-x+x)+z = 0+z = z \end{aligned}$$

This proves (a). Take $z=0$ in (a) to obtain (b). Take $z=-x$ in (a) to obtain (c). Since $-x+x=0$, (c) (with $-x$ in place of x) gives (d). □

mynotes:
mynotes 我自己证明上述四条性质时都是从定义开始的, 而 rudin 这里在后一步的证明中都利用了刚推导出的结论, 这一点需要借鉴.

Proposition 1.15. The axioms for multiplication imply the following statements.

(a) If $x \neq 0$ and $xy = xz$ then $y = z$.

(b) If $x \neq 0$ and $xy = x$ then $y = 1$.

(c) If $x \neq 0$ and $xy = 1$ then $y = 1/x$.

(d) If $x \neq 0$ then $1/(1/x) = x$.

The proof is so similar to that of Proposition 1.14 that we omit it.

证明. mynotes (a),

$$\begin{aligned} y = 1 \cdot y = \left(\frac{1}{x} \cdot x\right) y = \frac{1}{x}(xy) \\ = \frac{1}{x}(xz) = \left(\frac{1}{x}x\right) z = z \end{aligned}$$

(b), (a) 取 $z = 1$. $y = z = 1$.

(c), (a) 取 $z = \frac{1}{x}$. $y = z = \frac{1}{x}$.

(d), (c) 取 $x = \frac{1}{x'}$. $y = 1/(1/x')$. □

Proposition 1.16. The field axioms imply the following statements, for any $x, y, z \in F$.

(a) $0x = 0$.

(b) If $x \neq 0$ and $y \neq 0$ then $xy \neq 0$.

(c) $(-x)y = -(xy) = x(-y)$.

(d) $(-x)(-y) = xy$.

证明. $0x + 0x = (0+0)x = 0x$. Hence 1.14(b) implies that $0x = 0$, and (a) holds.

Next, assume $x \neq 0$, $y \neq 0$, but $xy = 0$. Then (a) gives

$$1 = \left(\frac{1}{y}\right)\left(\frac{1}{x}\right) xy = \left(\frac{1}{y}\right)\left(\frac{1}{x}\right) 0 = 0.$$

a contradiction. Thus (b) holds.

The first equality in (c) comes from

$$(-x)y + xy = (-x + x)y = 0y = 0,$$

combined with 1.14(c); the other half of (c) is proved in the same way.

Finally,

$$(-x)(-y) = -[x(-y)] = -[-(xy)] = xy$$

by (c) and 1.14(d). □

Definition 1.17. An ordered field is a field F which is also an ordered set, such that

(i) $x + y < x + z$ if $x, y, z \in F$ and $y < z$,

(ii) $xy > 0$ if $x \in F$, $y \in F$, $x > 0$, and $y > 0$.

If $x > 0$, we call x positive; if $x < 0$, x is negative.

For example, $\mathbb{Q}$ is an ordered field.

All the familiar rules for working with inequalities apply in every ordered field: Multiplication by positive [negative] quantities preserves [reverses] inequalities, no square is negative, etc. The following proposition lists some of these.

mynotes: 有序域 F 也是有序集, 由于有理数域 $\mathbb{Q}$, 实数域 $\mathbb{R}$ 都是有序域, 这里使用有理数域 $\mathbb{Q}$ 证明的有序集的性质也可以直接用于实数域. $\mathbb{R}$

Proposition 1.18. The following statements are true in every ordered field.

(a) If $x > 0$ then $-x < 0$, and vice versa.

(b) If $x > 0$ and $y < z$ then $xy < xz$.

(c) If $x < 0$ and $y < z$ then $xy > xz$.

(d) If $x \neq 0$ then $x^2 > 0$. In particular, $1 > 0$.

(e) If $0 < x < y$ then $0 < l/y < l/x$.

证明. (a) $x > 0$, $-x < 0$.

$$\begin{aligned} x > 0 &= (x + -x) \\ x + 0 &> x + (-x) \\ (-x) &< 0 \end{aligned}$$

(b) $x > 0$, $y < z$, $xy < xz$.

$$\begin{aligned} y < z, z - y &> y - y = 0 \\ x(z - y) &> 0 \\ x(z - y) + xy &> 0 + xy \\ xz &> xy \end{aligned}$$

(c)

$$\begin{aligned} & (z - y) > y - y = 0 \\ x < 0, (-x) > 0. \quad & (-x)(z - y) > 0 \\ & x(z - y) < 0 \\ & xz < xy \end{aligned}$$

(d)

$$\begin{aligned} & x > 0 && x^2 > 0 \\ & x < 0 && (-x)^2 > 0, (-x)^2 = -[x(-x)] = -(-(x \cdot x)) = x^2, x^2 > 0 \end{aligned}$$

$\because 1^2 = 1, 1 > 0.$

(e) If $y > 0$ and $v \leq 0$, then $yv \leq 0$. But $y \cdot (1/y) = 1 > 0$. Hence $1/y > 0$. Likewise, $1/x > 0$. If we multiply both sides of the inequality $x < y$ by the positive quantity $(1/x)(1/y)$, we obtain $1/y < 1/x$. □

1.4 THE REAL FIELD

We now state the *existence theorem* which is the core of this chapter.

Theorem 1.19. *There exists an* ordered field $\mathbb{R}$ *which has the least-upper-bound property.*

Moreover, $\mathbb{R}$ *contains* $\mathbb{Q}$ *as a* subfield.

The second statement means that $\mathbb{Q} \subset \mathbb{R}$ and that the operations of addition and multiplication in $\mathbb{R}$, when applied to members of $\mathbb{Q}$, coincide with the usual operations on rational numbers; also, the positive rational numbers are positive elements of $\mathbb{R}$.

The members of $\mathbb{R}$ are called real numbers.

The proof of Theorem 1.19 is rather long and a bit tedious and is therefore presented in an Appendix to Chap. 1. The proof actually constructs $\mathbb{R}$ from $\mathbb{Q}$.

The next theorem could be extracted from this construction with very little extra effort. However, we prefer to derive it from Theorem 1.19 since this provides a good illustration of what one can do with the least-upper-bound property.

mynotes:
$\mathbb{R}$ 具有最小上界性质的有序域 least-upper-bound $\rightarrow$ upper bound in the sets.
ordered field (ordered set, field).
$\mathbb{Q} \in \mathbb{R}$ subfield
$x \in \mathbb{R}$, x is a real number

mynotes:
proof of theorem 1.19 is tedious. construct $\mathbb{R}$ from $\mathbb{Q}$
tedious 乏味的, 冗长的
derive 取得, 得到

Theorem 1.20. *(archimedean property of* $\mathbb{R}$*) (a) If* $x \in \mathbb{R}$*,* $y \in \mathbb{R}$*, and* $x > 0$*, then there is a positive integer* n *such that*

$$nx > y$$

(b) If $x \in \mathbb{R}$*,* $y \in \mathbb{R}$*, and* $x < y$*, then there exists a* $p \in Q$ *such that* $x < p < y$*.*

Theorem 1.21. *For every real* $x > 0$ *and every integer* $n > 0$ *there is one and only one positive real* y *such that* $y^n = x$*.*

This number y is written $\sqrt[n]{x}$ or $x^{1/n}$.

Corollary If a and b are positive real numbers and n is a positive integer, then

$$(ab)^{1/n} = a^{1/n}b^{1/n}.$$

Definition 1.22. (Decimals) We conclude this section by pointing out the relation between real numbers and decimals.

1.5 THE EXTENDED REAL NUMBER SYSTEM

Definition 1.23. The extended real number system consists of the real field $\mathbb{R}$ and two symbols, $+\infty$ and $-\infty$. We preserve the original order in $\mathbb{R}$, and define

$$-\infty < x < +\infty$$

for every $x \in \mathbb{R}$

(a) If x is real then

$$x + \infty = +\infty, \quad x - \infty = -\infty, \quad \frac{x}{+\infty} = \frac{x}{-\infty} = 0.$$

(b) If $x > 0$ then $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$.

(c) If $x < \infty$ then $x \cdot (+\infty) = -\infty$, $x \cdot (-\infty) = +\infty$.

1.6 THE COMPLEX FIELD

mynotes:
rudin 引入复数定义的方法很奇怪, 使用复数的代数定义直接引入（天上掉下来的定义）。理解起来比较困难, 我觉得使用几何方法引入复数更为合理且直观, rudin 这里对初学者不太友好

Definition 1.24. A complex number is an ordered pair (a, b) of real numbers. "Ordered" means that (a, b) and (b, a) are regarded as distinct if $a \neq b$.

Let$x = (a, b)$, $y = (c, d)$ be two complex numbers. We write $x = y$ if and only if $a = c$ and $b = d$. (Note that this definition is not entirely superfluous; think of equality of rational numbers, represented as quotients of integers.) We define

$$x + y = (a + c, b + d),$$
$$xy = (ac - bd, ad + bc).$$

Theorem 1.25. *These definitions of addition and multiplication turn the* set *of all complex numbers into a field, with* $(0, 0)$ *and* $(1, 0)$ *in the role* 0 *and* 1.

proof (A1)–(A5), (M1)–(M5) and (D), then we can prove that $\mathbb{C}$ is a field.

Theorem 1.26. *For any real numbers a and b we have*

$$(a,0)+(b,0)=(a+b,0), \quad (a,0)(b,0)=(ab,0).$$

The proof is trivial.

show that the notation (a,b) is equivalent to the more customary $a+bi$.

Definition 1.27. $i=(0,1)$

Theorem 1.28. $i^2=-1$

证明.

$$i^2=(0,1)(0,1)=(-1,0)=-1.$$

□

Theorem 1.29. *If a and b are real, then $(a,b)=a+bi$.*

证明.

$$\begin{aligned} a+bi &= (a,0)+(b,0)(0,1) \\ &= (a,0)+(0,b)=(a,b) \end{aligned}$$

□

Definition 1.30. $a,b\in\mathbb{R}$, $z=a+bi$,the complex number $\bar{z}=a-bi$ is called the conjugate of z. the numbers a and b are the real part and imaginary part of z. respectively.

$$a=\Re(z), \quad b=\Im(z)$$

Theorem 1.31. *If z and w are complex, then*

(a) $\bar{z+w}=\bar{z}+\bar{w}$,

(b) $\bar{zw}=\bar{z}\cdot\bar{w}$,

(c) $z+\bar{z}=2\Re(z)$, $z-\bar{z}=2\Im(z)$,

(d) $z\bar{z}$ *is real and positive (except when* $z=0$*).*

Proof (a), (b),and (c)are quite trivial. To prove (d), write $z=a+bi$, and note that $z\bar{z}=a^2+b^2$.

Definition 1.32. If z is a complex number, its absolute value $|z|$ is the nonnegative square root of $z\bar{z}$; that is, $|z|=(z\bar{z})^{1/2}$.

The existence (and uniqueness) of $|z|$ follows from Theorem 1.21 and part (d) of Theorem 1.31.

Note that when x is real, then $\bar{x}=x$, hence $|x|=\sqrt{x^2}$. Thus $|x|=x$ if $x>0$, $|x|=-x$ if $x<0$.

Theorem 1.33. *Let z and w be complex numbers. Then*

(a) $|z| > 0$ *unless* $z = 0$, $|0| = 0$,

(b) $\bar{\bar{z}} = z$,

(c) $|zw| = |z||w|$,

(d) $|\Re(z)| \leq |z|$,

(e) $|z + w| \leq |z| + |w|$.

Notation 1.34. (sum) $x_1, x_2, \ldots, x_n \in \mathbb{C}$,

$$x_1 + x_2 + \cdots + x_n = \sum_{j=1}^{n} x_j.$$

Theorem 1.35. *(Schwarz Inequality)*

If $a_1, \ldots, a_n$, $b_1, \ldots, b_n$, *are complex numbers, then*

$$\left|\sum_{j=1}^{n} a_j \bar{b_j}\right|^2 \leq \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

在正式证明之前, 先回忆 $\mathbb{R}$ 中的施瓦茨不等式是怎么证明的. let $A = \sum a_j^2$, $B = \sum b_j^2$, $C = \sum a_j b_j$.

$$\sum (a_j + \lambda b_j)^2 = \sum a_j^2 + 2\sum a_j b_j \lambda + \sum b_j^2 \lambda^2$$

由韦达定理, $\Delta \leq 0$, $\Delta = (2\sum a_j b_j)^2 - 4\sum a_j^2 \sum b_j^2$. 因此 $(\sum a_j b_j)^2 \leq \sum a_j^2 \sum b_j^2$

证明. Put $A = \sum |a_j|^2$, $B = \sum |b_j|^2$, $C = \sum a_j \bar{b_j}$, $j = 1, 2, \ldots, n$.

If $B = 0$, $b_1 = \cdots = b_n = 0$, this conclusion is trivial.

If $B > 0$,

$$\begin{aligned}
\sum |Ba_j - Cb_j|^2 &= \sum (Ba_j - Cb_j)(B\bar{a_j} - \bar{C}\bar{b_j}) \\
&= B^2 \sum |a_j|^2 - B\bar{C} \sum a_j \bar{b_j} - BC \sum \bar{a_j} b_j + |C|^2 \sum |b_j|^2 \\
&= B^2 A - B|C|^2 \\
&= B(AB - |C|^2).
\end{aligned}$$

Since each term in the first sum is nonnegative, we see that

$$B(AB - |C|^2) \geq 0.$$

Since $B > 0$, it follows that $AB - |C|^2 \geq 0$. This is the desired inequality. □

我的想法

$$\begin{aligned}
\sum (a_j + \lambda \bar{b_j})(\bar{a_j} + \lambda b_j) &= \sum (a_j \bar{a_j} + \lambda(\bar{a_j} b_j + a_j \bar{b_j}) + \lambda^2 b_j \bar{b_j}) \\
&= \sum (a_j \bar{a_j} + \lambda 2\Re(a_j \bar{b_j}) + \lambda^2 b_j \bar{b_j})
\end{aligned}$$

由韦达定理, $\Delta \leq 0$, $\Delta = (2\sum \Re(a_j \bar{b_j}))^2 - 4\sum a_j \bar{a_j} \sum b_j \bar{b_j}$, (这里推出的结论比原始结论弱？为什么？)

1.7 Euclidean space

欧式空间

Definition 1.36. For each positive integer k, let $\mathbb{R}^k$ be the set of all ordered k-tuples

$$\mathbf{x} = (x_1, x_2, \ldots, x_k),$$

where $x_1, x_2, \ldots, x_k$ are real numbers, called the **coordinates** of $\mathbf{x}$. The elements of $\mathbb{R}^k$ are called points, or vectors, especially when $k > 1$. We shall denote vectors by boldfaced letters. If $\mathbf{y} = (y_1, y_2, \ldots, y_k)$ and if α is a real number, put

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k),$$
$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \ldots, \alpha x_k)$$

so that $\mathbf{x} + \mathbf{y} \in \mathbb{R}^k$ and $\alpha\mathbf{x} \in \mathbb{R}^k$. This defines addition of vectors, as well as multiplication of a vector by a real number (a scalar). These two operations satisfy the commutative, associative, and distributive laws (the proof is trivial, in view of the analogous laws for the real numbers) and make $\mathbb{R}^k$ into a vector space over the *real field*. The zero element of $\mathbb{R}^k$ (sometimes called the origin or the null vector) is the point **0**, all of whose coordinates are 0.

We also define the so-called "inner product" (or scalar product) of **x** and **y** by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{j=1}^{k} x_j y_j$$

and the norm of **x** by

$$|x| = (x \cdot x)^{1/2} = \left(\sum_{1}^{k} x_j^2 \right)^{1/2}.$$

The structure now defined (the vector space $\mathbb{R}^k$ with the above inner product and norm) is called euclidean k-space.

Theorem 1.37. *Suppose* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^k$*, and* α *is real. Then*

(a) $|\mathbf{x}| \geq 0$*;*

(b) $|\mathbf{x}| = 0$ *if and only if* $\mathbf{x} = 0$*;*

(c) $|\alpha\mathbf{x}| = |\alpha||x|$

(d) $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$*;*

(e) $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$*;*

(f) $|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|$.

(f) 为欧式空间中的三角不等式。

证明. Proof (a), (b), and (c) are obvious, and (d) is an immediate consequence of the Schwarz inequality**??**. By (d) we have

$$\begin{aligned}|\mathbf{x}+\mathbf{y}|^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y})\\ &= \mathbf{x}\cdot\mathbf{x}+2\mathbf{x}\cdot\mathbf{y}+\mathbf{y}\cdot\mathbf{y}\\ &\le |\mathbf{x}|^2+2|\mathbf{x}||\mathbf{y}|+|\mathbf{y}|^2\\ &= (|\mathbf{x}|+|\mathbf{y}|)^2.\end{aligned}$$

so that (e) is proved. Finally, (f) follows from (e) if we replace $\mathbf{x}$ by $\mathbf{x}-\mathbf{y}$ and $\mathbf{y}$ by $\mathbf{y}-\mathbf{z}$. □

Remark 1.38. Theorem 1.37 (a), (b), and (f) will allow us (see Chap. 2) to regard $\mathbb{R}^k$ as a metric space.

$\mathbb{R}^1$ (the set of all real numbers) is usually called the line, or the real line. Likewise, $\mathbb{R}^2$ is called the plane, or the complex plane (compare Definitions **??** and 1.36). In these two cases the norm is just the absolute value of the corresponding real or complex number.

appendix

Theorem 1.19 will be proved in this appendix by constructing $\mathbb{R}$ from $\mathbb{Q}$. We shall divide the construction into several steps.

Step 1 The members of $\mathbb{R}$ will be certain subsets of $\mathbb{Q}$, called *cuts*. A cut is, by definition, any set[3] $\alpha\subset\mathbb{Q}$ with the following three properties.

(I) α is not empty, and $\alpha\neq\mathbb{Q}$.

(II) If $p\in\alpha$,$q\in\mathbb{Q}$, and $q<p$, then $q\in\alpha$.

(III) If $p\in\alpha$, then $p<r$ for some $r\in\alpha$.

The letters $p, q, r, ...$ will always denote rational numbers, and $\alpha, \beta, \gamma, ...$ will denote cuts.

mynotes:
建立分划定义的这三条性质说明了有理数集是稠密而不是连续的

Note that (III) simply says that a has no largest member: (II) implies two facts which will be used freely:

If $p\in\alpha$ and $q\notin\alpha$ then $p<q$.

If $r\notin\alpha$ and $r<s$ then $s\notin\alpha$.

Step 2 Define "$\alpha<\beta$" to mean: α is a proper subset of β.[4]

Let us check that this meets the requirements of Definition 1.5.

If $\alpha<\beta$ and $\beta<\gamma$ it is clear that $\alpha<\gamma$. (A proper subset of a proper subset is a proper subset.) It is also clear that at most one of the three relations

$$\alpha<\beta,\qquad \alpha=\beta,\qquad \beta<\alpha.$$

[3] 分划 α 是一个集合

[4] 这里使用真子集关系定义了分划 (集合) 间的序

can hold for any pair α, β. To show that at least one holds, assume that the first two fail. Then α is not a subset of β. Hence there is a $p \in \alpha$ with $p \notin \beta$. If $q \in \beta$, it follows that $q < p$ (since $p \notin \beta$), hence $q \in \alpha$, by (II). Thus $\beta \subset \alpha$. Since $\beta \neq \alpha$, we conclude: $\beta < \alpha$.

Thus $\mathbb{R}$ is now an ordered set.

mynotes:
利用集合关系定义的偏序关系具有传递性和三歧性

Step 3 The ordered set $\mathbb{R}$ has the least-upper-bound property.

To prove this, let A be a nonempty subset of $\mathbb{R}$, and assume that $\beta \in \mathbb{R}$ is an upper bound of A. Define γ to be the union of all $\alpha \in A$. In other words, $p \in \gamma$ if and only if $p \in \alpha$ for some $\alpha \in A$. We shall prove that $\gamma \in \mathbb{R}$ and that $\gamma = \sup A$.

Since A is not empty, there exists an $a_0 \in A$. This a is not empty. Since $\alpha_0 \in \gamma$, γ is not empty. Next, $\gamma \subset \beta$ (since $\alpha \subset \beta$ for every $\alpha \in A$), and therefore $\gamma \neq \mathbb{Q}$. Thus γ satisfies property (I). To prove (II) and (III), pick $p \in \gamma$. Then $p \in \alpha_1$ for some $\alpha_1 \in A$. If $q < p$, then $q \in \alpha_1$, hence $q \in \gamma$; this proves (II). If $r \in \alpha_1$ is so chosen that $r > p$, we see that $r \in \gamma$ (since $\alpha_1 \subset \gamma$), and therefore γ satisfies (III).

Thus $\gamma \in \mathbb{R}$.

It is clear that $\alpha \leq \gamma$ for every $\alpha \in A$.

Suppose $\delta < \gamma$. Then there is an $s \in \gamma$ and that $s \notin \delta$. Since $s \in \gamma, s \in \alpha$ for some $\alpha \in A$. Hence $\delta < a$, and δ is not an upper bound of A.

This gives the desired result: $\gamma = \sup A$.

mynotes:
这里分划之间所用的 $\in$ 让我很费解, 上文对分划的定义是集合, 那么应该用子集形式而不是元素形式来描述偏序关系. 分划是不是集合的一个元素? 查询原始 pdf 文件发现确实是子集形式描述的!

Step 4 If $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$ we define $\alpha + \beta$ to be the set of all sums $r + s$, where $r \in \alpha$ and $s \in \beta$.

We define 0^* to be the set of all negative rational numbers. It is clear that 0^* is a cut. We verify that the axioms for addition (see Definition 1.12) hold in $\mathbb{R}$, with 0^* playing the role of 0.

Step 5 Having proved that the addition defined in Step 4 satisfies Axioms (A) of Definition 1.12, it follows that Proposition 1.14 is valid in $\mathbb{R}$, and we can prove one of the requirements of Definition **??**:

If $\alpha, \beta, \gamma \in \mathbb{R}$ and $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$.

Indeed, it is obvious from the definition of $+$ in $\mathbb{R}$ that $\alpha+\beta \subset \alpha+\gamma$; if we had $\alpha+\beta = \alpha+\gamma$, the cancellation law (Proposition 1.14) would imply $\beta = \gamma$.

It also follows that $\alpha > 0^*$ if and only if $-\alpha < 0^*$.

Step 6 Multiplication is a little more bothersome than addition in the present context, since products of negative rationals are positive. For this reason we confine ourselves first to $\mathbb{R}^+$, the set of all $\alpha \in \mathbb{R}$ with $\alpha > 0^*$.

If $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$, we define $\alpha\beta$ to be the set of all p such that $p \leq rs$ for some choice of $r \in \alpha$, $s \in \beta$, $r > 0$, $s > 0$.

We define 1^* to be the set of all $q < 1$.

Then the axioms (M) and (D) of Definition 1.12 hold, with $\mathbb{R}^+$ in place of F, and with 1^* in the role of 1.

The proofs are so similar to the ones given in detail in Step 4 that we omit them.

Note, in particular, that the second requirement of Definition 1.17 holds: If $\alpha > 0^*$ and $\beta > 0^*$ then $\alpha\beta > 0^*$.

Step 7 We complete the definition of multiplication by setting $\alpha 0^* = 0^*\alpha = 0^*$, and by setting

$$\alpha\beta = \begin{cases} (-\alpha)(-\beta) & \text{if } \alpha < 0^*, \beta < 0^*, \\ -[(-\alpha)\beta] & \text{if } \alpha < 0^*, \beta > 0^*, \\ -[\alpha \cdot (-\beta)] & \text{if } \alpha > 0^*, \beta < 0^*, \end{cases}$$

The products on the right were defined in Step 6.

Having proved (in Step 6) that the axioms (M) hold in $\mathbb{R}^+$, it is now perfectly simple to prove them in$\mathbb{R}^+$, by repeated application of the identity $\gamma = -(-\gamma)$ which is part of Proposition 1.14. (See Step 5.)

The proof of the distributive law

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

breaks into cases. For instance, suppose $\alpha > 0^*$, $\beta < 0^*$, $\beta + \gamma > 0^*$ Then $\gamma = (\beta + \gamma) + (-\beta)$, and (since we already know that the distributive law holds in $\mathbb{R}^+$)

$$\alpha\gamma = \alpha(\beta + \gamma) + \alpha \cdot (-\beta).$$

But $\alpha \cdot (-\beta) = -(\alpha\beta)$. Thus

$$\alpha\beta + \alpha\gamma = \alpha(\beta + \gamma).$$

The other cases are handled in the same way.

We have now completed the proof that $\mathbb{R}$ is an orderedfield with the least-upper-bound property.

Step 8 We associate with each $r \in \mathbb{Q}$ the set r^* which consists of all $p \in \mathbb{Q}$ such that $p < r$. It is clear that each r^* is a cut; that is, $r^* \in \mathbb{R}$. These cuts satisfy the following relations:

(a) $r^* + s^* = (r + s)^*$,

(b) $r^* s^* = (rs)^*$,

(c) $r^* < s^*$ if and only if $r < s$.

To prove (a), choose $p \in r^* + s^*$. Then $p = u + v$, where $u < r$, $v < s$. Hence $p < r + s$, which says that $p \in (r + s)^*$.

Conversely, suppose $p \in (r+s)^*$. Then $p < r+s$. Choose t so that $2t = r+s-p$, put

$$r' = r - t, s' = s - t.$$

Then $r' \in r^*$, $s' \in s^*$, and $p = r' + s'$, so that $p \in r^* + s^*$

This proves (a). The proof of (b) is similar.

If $r < s$ then $r \in s*$, but $r \notin r^*$; hence $r^* < s^*$.

If $r^* < s^*$ then there is a $p \in s^*$ such that $p \notin r^*$ Hence $r < p < s$, so that $r < s$.

This proves (c).

Step 9 We saw in Step 8 that the replacement of the rational numbers r by the corresponding "rational cuts" $r^* \in \mathbb{R}$ preserves sums, products, and order. This fact may be expressed by saying that the ordered field $\mathbb{Q}$ is isomorphic to the ordered field $\mathbb{Q}^*$ whose elements are the rational cuts. Of course, r^* is by no means the same as r, but the properties we are concerned with (arithmetic and order) are the same in the two fields.

It is this identification of $\mathbb{Q}$ with $\mathbb{Q}^*$ which allows us to regard $\mathbb{Q}$ as a subfield of $\mathbb{R}$.

The second part of Theorem 1.19 is to be understood in terms of this identification. Note that the same phenomenon occurs when the real numbers are regarded as a subfield of the complex field, and it also occurs at a much more elementary level, when the integers are identified with a certain subset of $\mathbb{Q}$.

It is a fact, which we will not prove here, that any two orderedfields with the least-upper-bound property are isomorphic. The first part of Theorem 1.19 therefore characterizes the real field $\mathbb{R}$ completely.

The books by Landau and Thurston cited in the Bibliography are entirely devoted to number systems. Chapter 1 of Knopp's book contains a more leisurely description of how $\mathbb{R}$ can be obtained from $\mathbb{Q}$. Another construction, in which each real number is defined to be an equivalence class of Cauchy sequences of rational numbers (see Chap. 3), is carried out in Sec. 5 of the book by Hewitt and Stromberg.

The cuts in $\mathbb{Q}$ which we used here were invented by Dedekind. The construction of $\mathbb{R}$ from $\mathbb{Q}$ by means of Cauchy sequences is due to Cantor. Both Cantor and Dedekind published their constructions in 1872.

EXERCISES

Unless the contrary is explicitly stated, all numbers that are mentioned in these exercises are understood to be real.

Exercise 1.1. $r \in \mathbb{Q}, r \neq 0, x \notin \mathbb{Q}, x \in \mathbb{R}\; r + x, rx \notin \mathbb{Q}, \in \mathbb{R}$

Solve 1. if $r + x \in \mathbb{Q}$, there exists $m, n \in \mathbb{N}, n \neq 0$, s.t. $r + x = \frac{m}{n}$. $\because r \in \mathbb{Q}, r = \frac{p}{q}, p, q \in$

$\mathbb{N}, q \neq 0.$

$$r + x = \frac{m}{n}$$
$$\frac{p}{q} + x = \frac{m}{n}$$

$$x = \frac{m}{n} - \frac{p}{q} = \frac{mq - np}{nq}$$

then $x \in \mathbb{Q}$ contradict to the supposion that $x \notin \mathbb{Q}$

If $rx \in \mathbb{Q}$, then $rx = \frac{m}{n}, m, n \in \mathbb{N}$, $x = \frac{qm}{pn} \in \mathbb{Q}$, contradictory!

Exercise 1.2. prove that there is no rational number whose square is 12.

Solve 2. If $(p/q)^2 = 12$, $p^2/q^2 = 12$. p must be even, $p = 2m$. $(2m)^2/q^2 = 12$, $m^2/q^2 = 3$. 3 is a prime number, $m = 3n$, $(3n)^2/q^2 = 3$, $3n^2 = q^2$, q have a factor 3, $\gcd(p, q) = \gcd(m, q) = \gcd(n, q) = 3 \neq 1$, contradict to the fact that p, q are coprime.

Exercise 1.3. Prove Proposition 1.15.

Solve 3. (a) $x \neq 0$, $xy \neq xz$. $x \neq 0$, $\exists 1/x$, $1/x \cdot x = 1$.

$$\begin{aligned} y &= \left(\frac{1}{x} x \cdot x\right) y = \frac{1}{x}(xy) \\ &= \frac{1}{x}(xz) = \left(\frac{1}{x} x \cdot x\right) z = z. \end{aligned}$$

(b) $x \neq 0$, $xy = x$ then $y = 1$.

Let $z = 1$ in (a).

(c) $x \neq 0$, $xy = 1$ then $y = 1/x$.

Let $z = 1/x$ in (a).

(d) $x \neq 0$, $1/(1/x) = x$.

$x \cdot \frac{1}{x} = 1$, $\frac{1}{x} \cdot \frac{1}{\frac{1}{x}} = 1$. then $x \cdot \frac{1}{x} = \frac{1}{x} \cdot \frac{1}{\frac{1}{x}}$. so $1/(1/x) = x$.

Exercise 1.4. $E = \varnothing$, E 为有序的非空子集. α 是 E 的下界 β 是 E 的上界 Prove that $\alpha \leq \beta$.

Solve 4. $\forall x \in E, \alpha \leq x, x \leq \beta$. $\alpha \leq x \leq \beta$, $\alpha \leq \beta$.

Exercise 1.5. A 为 $\mathbb{R}$ 的非空子集, A 有下界

$$-A = \{-x | x \in A\}$$

Prove that $\inf A = -\sup(-A)$

Solve 5. (rudin) $\beta = \inf A$, $\alpha = \sup(-A)$.

(1) $\beta < -\alpha$, $\exists x \in A$, $\beta \leq x < \alpha$, $-x > \alpha$. 矛盾.

(2) $\beta > -\alpha$, $\exists x \in A$, $\alpha \geq -x > -\beta$, $x < \beta$. 矛盾.

$\therefore \beta = -\alpha$.

Exercise 1.6. Fix $b > 1$,
(a) If m, n, p, q are integers, $n > 0, q > 0$, and $r = m/n = p/q$, prove that

$$(b^m)^{1/n} = (b^p)^{1/q}$$

Hence it makes sense to define $b^r = (b^m)^{1/n}$.
(b) Prove that $b^{r+s} = b^r b^s$ if r and s are rational.
(c) If x is real, define $B(x)$ to be the set of all numbers b^t, where t is rational and $t \leq x$. Prove that

$$b^r = \sup B(r)$$

when r is rational. Hence it makes sense to define

$$b^x = \sup B(x)$$

for every real x. (d) Prove that $b^{x+y} = b^x b^y$ for all real x and y.

Exercise 1.7. Fix $b > 1, y > 0$, and prove that there is a unique real x such that $b^x = y$, by completing the following outline. (This x is called the logarithm of y to the base b.)
(a) For any positive integer n, $b^n - 1 \geq n(b - 1)$.
(b) Hence $b - 1 \geq n(b^{1/n} - 1)$.
(c) If $t > 1$ and $n > (b - 1)/(t - 1)$, then $b^{1/n} < t$.
(d) If w is such that $b^w < y$, then $b^{w+(1/n)} < y$ for sufficiently large n; to see this, apply part (c) with $t = y \cdot b^{-w}$.
(e) If $b^w > y$, then $b^{w-(1/n)} > y$ for sufficiently large n.
(f) Let A be the set of all w such that $b^w < y$, and show that $x = \sup A$ satisfies $b^x = y$.
(g) Prove that this x is unique.

Exercise 1.8. Prove that no order can be defined in the complex field that turns it into an ordered field.
Hint: -1 is a square.

Exercise 1.9. Suppose $z = a + bi, w = c + di$. Define $z < w$ if $a < c$, and also if $a = c$ but $b < d$. Prove that this turns the set of all complex numbers into an ordered set. (This type of order relation is called a *dictionary order*, or *lexicographic order*, for obvious reasons.) Does this ordered set have the least-upper-bound property?

Exercise 1.10. Suppose $z = a + bi$, $w = u + iv$, and

$$a = \left(\frac{|w| + u}{2}\right)^{1/2}, \qquad b = \left(\frac{|w| - u}{2}\right)^{1/2}.$$

Prove that $z^2 = w$ if $v \geq 0$ and that $(\bar{z})^2 = w$ if $v \leq 0$. Conclude that every complex number (with one exception!) has two complex square roots.

Exercise 1.11. If z is a complex number, prove that there exists an $r \geq 0$ and a complex number w with $|w| = 1$ such that $z = rw$. Are w and r always uniquely determined by z?

Exercise 1.12. If $z_1, ..., z_n$ are complex, prove that

$$|z_1 + z_2 ... + z_n| \leq |z_1| + |z_2| ... + |z_n|.$$

Exercise 1.13. If x, y are complex, prove that

$$||x| - |y|| \leq |x - y|.$$

Exercise 1.14. If z is a complex number such that $|z| = 1$, that is, such that $z\bar{z} = 1$, compute

$$|1 + z|^2 + |1 - z|^2$$

Exercise 1.15. Under what conditions does equality hold in the Schwarz inequality?

Exercise 1.16. Suppose $k \geq 3$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $|\mathbf{x} - \mathbf{y}| = d > 0$, and $r > 0$. Prove:
(a) If $2r > d$, there are infinitely many $\mathbf{z} \in \mathbb{R}^k$ such that

$$|z - x| = |z - y| = r.$$

(b) If $2r = d$, there is exactly one such $\mathbf{z}$,
(c) If $2r < d$, there is no such $\mathbf{z}$.
How must these statements be modified if k is 2 or 1?

Exercise 1.17. Prove that

$$|\mathbf{x} + \mathbf{y}|^2 + |\mathbf{x} - \mathbf{y}|^2 = 2|\mathbf{x}|^2 + 2|\mathbf{y}|^2$$

if $\mathbf{x} \in \mathbb{R}^k$ and $\mathbf{y} \in \mathbb{R}^k$. Interpret this geometrically, as a statement about parallelograms.

Exercise 1.18. If $k > 2$ and $\mathbf{x} \in \mathbb{R}^k$, prove that there exists $\mathbf{y} \in \mathbb{R}^k$ such that $\mathbf{y} \neq 0$ but $\mathbf{x} \cdot \mathbf{y} = 0$. Is this also true if $k = 1$?

Exercise 1.19. Suppose $\mathbf{a} \in \mathbb{R}^k$, $\mathbf{b} \in \mathbb{R}^k$. Find $\mathbf{c} \in \mathbb{R}^k$ and $r > 0$ such that

$$|\mathbf{x} - \mathbf{a}| = 2|\mathbf{x} - \mathbf{b}|$$

if and only if $|\mathbf{x} - \mathbf{a}| = r$. (Solution: $3\mathbf{c} = 4\mathbf{b} - \mathbf{a}$, $3r = 2|\mathbf{b} - \mathbf{a}|$.)

Exercise 1.20. With reference to the Appendix, suppose that property (III) were omitted from the definition of a cut. Keep the same definitions of order and addition. Show that the resulting ordered set has the least-upper-bound property, that addition satisfies axioms (A1) to (A4) (with a slightly different zero-element!) but that (A5) fails.

Chapter 2

Basic topology

2.1 Finite, countable, and uncountable sets

We begin this section with a definition of the **function** concept.

Definition 2.1. Consider two sets A and B, whose elements may be any objects whatsoever, and suppose that with each element x of A there is associated, in some manner, an element of B, which we denote by $f(x)$. Then f is said to be a *function* from A to B (or a *mapping* of A into B). The set A is called the *domain* of f (we also say fis defined on A), and the elements $f(x)$ are called the *values* of f. The set of all values of f is called the *range* of f.

Definition 2.2. Let A and B be two sets and let f be a mapping of A into B. If $E \subset A$, $f(E)$ is defined to be the set of all elements $f(x)$, for $x \in E$. We call $f(E)$ the image of E under f. In this notation, $f(A)$ is the range of f. It is clear that $f(A) \subset B$. If $f(A) = B$, we say that f maps A *onto* B. (Note that, according to this usage, *onto* is more specific than *into*.)[1]

If $E \subset B$, $f^{-1}(E)$ denotes the set of all $x \in A$ such that $f(x) \in E$. We call $f^{-1}(E)$ the *inverse image* of E under f. If $y \in B$, $f^{-1}(y)$ is the set of all $x \in A$ such that $f(x) = y$. If, for each $y \in B$, $f^{-1}(y)$ consists of at most one element of A, then f is said to be a 1-1 (*one-to-one*) mapping of A into B. This may also be expressed as follows: f is a 1-1 mapping of A into B provided that $f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$, $x_1 \in A$, $x_2 \in A$.

(The notation $x_1 \neq x_2$, means that x_1 and x_2 are distinct elements; otherwise we write $x_1 = x_2$.)

Definition 2.3. If there exists a 1-1 mapping of A *onto* B, we say that A and B can be putin 1-1 correspondence, or that A and B have the same cardinal number, or, briefly, that A and B are equivalent, and we write $A \sim B$. This relation clearly has the following properties :

It is reflexive: $A \sim A$.

[1] onto 满射? into 映射?

It is symmetric: If $A \sim B$, then $B \sim A$.

It is transitive: If $A \sim B$ and $B \sim C$, then $A \sim C$.

Any relation with these three properties is called an equivalence relation.[2]

> mynotes:
> 集合等势是一种等价关系, 其满足自反性, 对称性, 传递性.

Definition 2.4. $\forall n \in \mathbb{N}^+$, $J_n = \{1, 2, ..., n\}$, $J = \{1, 2, ..., n, ...\}$, (set consisting of all positive integers).

A is finite, $A \sim J_n$ for some n,

$A = \varnothing$. empty set is also considered to be finite.

A is infinite, A is not finite.

A is countable, $A \sim J$

A is uncountable. A is neither finite nor countable.

countable set and finite set are called at most countable.

> mynotes:
>
> $$\begin{cases} finite & A \sim J_n \\ infinite & \begin{cases} countable & A \sim J \\ uncountable & \end{cases} \end{cases}$$

countable sets, enumerable, denumerable.

$A, B \in$ finite set

$A \sim B \Longleftrightarrow A, B$ contains sasme number of elements

$A, B \in$ infinite set

same number or elements? vague

1-1 correspondence. retains its clarity.

Example 2.5. $f : J \to A$

$$f(n) = \begin{cases} \frac{n}{2} & (n\text{even}) \\ -\frac{n-1}{2} & (n\text{odd}) \end{cases}$$

> mynotes:
> $f(n) = (-1)^n \left\lfloor \frac{n}{2} \right\rfloor$

Remark 2.6. a finite set cannot be equivalent to one of its proper subsets, but it's possible for infinite sets.

$J = 1, 2, 3, 4, ..., A = 0, 1, -1, 2, -2, ...,J, A$are infinite sets, $J \subset A$.

but there exist a function $f : J \to A$, $J \sim A$

[2]等价关系：自反性，对称性，传递性

Definition 2.7. $f(x), x \in J = \mathbb{N}^{+}$.
$\{x_n\}, x_1, x_2, x_3, ...$
x_n, terms of the sequence.
$\forall n \in J$, $x_n \in A$, $\{x_n\}$ is a sequence in A, or a sequence of elements of A.

every countable set is range of a sequence of distinct terms. the elements of any countable set can be "arranged in a sequence". replace $J(\mathbb{N}^{+})$ by $\mathbb{N} = \{x|, x \in Z, x \geq 0\}$, start with 0 rather than 1.

Theorem 2.8. *Every infinite subset of a countable set A is countable*

$E \subset A$. E is infinite. To prove E is countable, we need a 1-1 correspondation of J to E, $f : J \to E$.

mynotes:
my first guess is A is a countable set, $A \sim J$ (by def). $\exists$ 1-1 mapping $g : J$ onto A. $x \in J$, $g(x) \in A$. $E \subset A, \exists g(x) \in E$. $g(x_i) \in E, x_i \in J, g : J \to E$.
再证 x_i 不是有限的. E is infinite, there exist infinite $g(x_i) \in E$. $\because$ g is a 1-1 mapping, $\{x_i\}$ is infinite. $\therefore$ $J \sim E$.

证明. Suppose $E \subset A$, E is infinite. arrange the elements x of A in a sequence $\{x_n\}$ of a distinct elements. Construct a sequence n_k as follows.
Let n_1 be the smallest positive int, s.t. $x_{n_1} \in E$. Having chosen $n_1, ...n_{k-1}$,$(k = 2, 3, ...)$, let n_k be the smallest integer greater than n_{k-1}, s.t. $x_{n_k} \in E$.
Putting $f(k) = x_{n_k}$, $f : J \to E$ is a 1-1 mapping. □

Countable sets represent the "smallest" infinity.
No uncountable set ca be a subset of a countable set.

mynotes:
rudin 这里尝试区分实无穷与浅无穷，使用集合的势来说明更为具体，全体整数组成的集合为 "最小" 的无穷大，其势为 $\aleph_0$，康托尔使用一一对应关系作为无穷集合之间的等价关系

Definition 2.9. $\forall \alpha \in A$, $E_\alpha \subset \Omega$, $\{E_\alpha\}$ debites elements of E_α. collection of sets (or family of sets)[3] union

$$S = \bigcup_{\alpha \in A} E_\alpha \tag{2.1}$$

if A consists of the integers $1, 2, ..., n$.

$$S = \bigcup_{m=1}^{n} E_m \tag{2.2}$$

$$S = E_1 \bigcup E_2 \bigcup \cdots \bigcup E_n. \tag{2.3}$$

[3]sets of sets sounds strange

if A is the set of all positive integers.

$$S = \bigcup_{m=1}^{\infty} E_m. \tag{2.4}$$

intersection

$$P = \bigcap_{\alpha \in A} E_\alpha \tag{2.5}$$

$$S = \bigcap_{m=1}^{n} E_m = E_1 \bigcap E_2 \bigcap \cdots \bigcap E_n. \tag{2.6}$$

$$S = \bigcap_{m=1}^{\infty} E_m. \tag{2.7}$$

A and B intersect if $A \bigcap B$ is not empty, otherwise they are disjoint.

Example 2.10. some example of set relation

Remark 2.11. Many properties of unions and intersections are quite similar to those of sums and products; in fact, the words sum and product were sometimes used in this connection, and the symbols $\sum$ and $\prod$ were written in place of $\bigcup$ and $\bigcap$.

The commutative and associative laws are trivial:

$$A \bigcup B = B \bigcup A; \qquad A \bigcap B = B \bigcap A \tag{2.8}$$

$$\left(A \bigcup B\right) \bigcup C = A \bigcup \left(B \bigcup C\right); \qquad \left(A \bigcap B\right) \bigcap C = A \bigcap \left(B \bigcap C\right); \tag{2.9}$$

Thus the omission of parentheses in 2.3 and 2.6 is justified.

The distributive law also holds:

$$A \bigcap \left(B \bigcup C\right) = \left(A \bigcap B\right) \bigcup \left(A \bigcap C\right). \tag{2.10}$$

To prove this, let the left and right members of 2.10 be denoted by E and F, respectively.

Suppose $x \in E$. Then $x \in A$ and $x \in B \bigcup C$, that is, $x \in B$ or$x \in C$ (possibly both). Hence $x \in A \bigcap B$ or $x \in A \bigcap C$, so that $x \in F$. Thus $E \subset F$.

Next, suppose $x \in F$. $Then x \in A \bigcap B$ or $x \in A \bigcap C$. That is, $x \in A$, and $x \in B \bigcup C$. Hence $x \in A \bigcap (B \bigcup C)$, so that $F \subset E$.

It follows that $E = F$.

We list a few more relations which are easily verified:

$$A \subset A \bigcup B, \tag{2.11}$$

$$A \bigcap B \subset B, \tag{2.12}$$

If 0 denotes the empty set, then[4]

$$A \bigcup 0 = A, \quad A \bigcap 0 = 0. \tag{2.13}$$

If $A \subset B$, then

$$A \bigcup B = B, \quad A \bigcap B = A. \tag{2.14}$$

Theorem 2.12. *Let $\{E_n\}, n = 1, 2, 3, ...,$ be a sequence of countable sets, and put*

$$S = \bigcup_{n=1}^{\infty} E_n. \tag{2.15}$$

Then S is countable.

将 E_n 按顺序排成一张表格，按反对角线重新排列成新的序列，得到 T, $S \sim T$. S is at most countable. 同时存在无限集合 (infinite set) E_1, $E_1 \subset S$, S is countable.

Corollary. Suppose A is at most countable, and, for every $\alpha \in A$, B, is at most countable. Put

$$T = \bigcup_{\alpha \in A} B_\alpha.$$

Then T is at most countable.

For T is equivalent to a subset of 2.15.

Theorem 2.13. *Theorem Let A be a countable set, and let B_n be the set of all n-tuples $(a_1, ..., a_n)$, where $a_k \in A (k = 1, ..., n)$, and the elements $a_1, ..., a_n$ need not be distinct. Then B_n is countable.*

Corollary. The set of all rational numbers is countable.

Theorem 2.14. *Theorem Let A be the set of all sequences whose elements are the digits 0 and 1. This set A is uncountable.*

The elements of A are sequences like $1, 0, 0, 1, 0, 1, 1, 1, ...$

2.2 Metric space

Definition 2.15. set X metric space
$p \in X$, p point.

$\forall p, q \in X$ associate a real number $d(p, q)$ (distance)
a. $d(p, q) > 0$ if $p \neq q$; $d(p, p) = 0$,
b. $d(p, q) = d(q, p)$.
c. $d(p, q) \leq d(p, r) + d(r, q)$, $\forall r \in X$

对称性，正定性，三角不等式。

[4]现在一般使用 $\varnothing$ 指代空集

Example 2.16. the distance of the euclidean space $\mathbb{R}^k$ is defined by

$$d(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}| \qquad (\mathbf{x},\mathbf{y} \in \mathbb{R}^k) \tag{2.16}$$

It's important to observe that every subset Y of metric space X is a metric space in its own right, with the same distance function. For it is clear that if conditions (a) to (c) of Definition 2.15 hold for $p, q, r \in X$, they also hold if we restrict p, q, r to lie in Y.

Thus every subset of a euclidean space is a metric space. Other examples are the spaces $l(K)$ and $L^2(\mu)$, , which are discussed in Chaps. 7 and 11, respectively.

Definition 2.17. By the *segment* (a, b) we mean the set of all real numbers x such that $a < x < b$.

By the *interval* $[a.b]$ we mean the set of all real numbers x such that $a \leq x \leq b$

Occasionally we shall also encounter "half-open intervals" $[a, b)$ and $(a, b]$; the first consists of all x such that $a \leq x < b$, the second of all x such that $a < x \leq b$

If $a_i < b_i$ for $i = 1, ..., k$, the set of all points $\mathbf{x} = (x_1, ..., x_k)$ in $\mathbb{R}^k$ whose coordinates satisfy the inequalities $a_i \leq x_i \leq_i (1 \leq i \leq k)$ is called a *k-cell.*
Thus a 1-cell is an interval, a 2-cell is a rectangle, etc.

If $\mathbf{x} \in \mathbb{R}^k$ and $r > 0$. the *open (or closed) ball* B with center at $\mathbf{x}$ and radius r is defined to be the set of all $\mathbf{y} \in \mathbb{R}^k$ such that $|\mathbf{y}-\mathbf{x}| < r$ (or $|\mathbf{y}-\mathbf{x}| \leq r$).

We call a set $E \subset \mathbb{R}^k$ *convex* if

$$\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$$

whenever $\mathbf{x} \in E$, $\mathbf{y} \in E$, and $0 < \lambda < 1$.

For example, *balls are convex.* For if |y -x| <r, |z-x| <r, and $0 < \lambda < 1$, we have

$$\begin{aligned}|\lambda\mathbf{y} + (1-\lambda)\mathbf{z} - \mathbf{x}| &= |\lambda(\mathbf{y}-\mathbf{x}) + (1-\lambda)(\mathbf{z}-\mathbf{x}) \\ &\leq \lambda|\mathbf{y}-\mathbf{x}| + (1-\lambda)|\mathbf{z}-\mathbf{x}| < \lambda r + (1-\lambda)r \\ &= r.\end{aligned}$$

The same proof applies to closed balls. It is also easy to see that k-cells are convex.

mynotes:
这里给出了开区间, 闭区间, 半开区间以及凸集 convex 的定义

Definition 2.18. Definition Let X be a metric space. All points and sets mentioned below are understood to be elements and subsets of X.

(a) A *neighborhood* of p is a set $N_r(p)$ consisting of all q such that $d(p, q) < r$,for some $r > 0$. The number r is called the *radius* of $N_r(p)$.

(b) A point p is a *limit point* of the set E if *every* neighborhood of p contains a point $q \neq p$ such that $q \in E$.

(c) If $p \in E$ and p is not a limit point of E, then p is called an *isolated point* of E.

(d) E is *closed* if every limit point of E is a point of E.

(e) A point p is an *interior* point of E if there is a neighborhood N of p such that $N \subset E$.

(f) E is *open* if every point of E is an interior point of E.

(g) The *complement* of E (denoted by E^c)is the set of all points $p \in X$ such that $p \notin E$.

(h) E is *perfect* if E is closed and if every point of E is a limit point of E.

(i) E is *bounded* if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$.

(j) E is *dense* in X if every point of X is a limit point of E, or a point of E (or both).

Let us note that in $\mathbb{R}^1$ neighborhoods are segments, whereas in $\mathbb{R}^2$ neighborhoods are interiors of circles.

Theorem 2.19. *Every neighborhood is un open set.*

Theorem 2.20. *If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.*

Corollary. A finite point set has no limit points.

Example 2.21. Let us consider the following subsets of $\mathbb{R}^2$:

(a) The set of all complex z such that $|z| < 1$.

(b) The set of all complex z such that $|z| \leq 1$.

(c) A nonempty finite set.

(d) The set of all integers.

(e) The set consisting of the numbers $1/n(n = 1, 2, 3, ...)$. Let us note that this set E has a limit point (namely, $z = 0$) but that no point of E is a limit point of E; we wish to stress the difference between having a limit point and containing one.

(f) The set of all complex numbers (that is, $\mathbb{R}^2$).

(g) The segment (a, b).

Let us note that (d),(e),(g)can be regarded also as subsets of R'. Some properties of these sets are tabulated below:

	Closed	Open	Perfect	Bounded
(a)	No	Yes	No	Yes
(b)	Yes	No	Yes	Yes
(c)	Yes	No	No	Yes
(d)	Yes	No	No	No
(e)	No	No	No	Yes
(f)	Yes	Yes	Yes	No
(g)	No		No	Yes

In (g), we left the second entry blank. The reason is that the segment (a, b) is not open if we regard it as a subset of $\mathbb{R}^2$, but it is an open subset of $\mathbb{R}^1$.

mynotes:
根据定义, 复数集既是闭集又是开集...

Theorem 2.22. *Let $\{E_\alpha\}$ be a (finite or infinite) collection of sets E_α. Then*

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c) \tag{2.17}$$

Theorem 2.23. *A set E is open if and only if its complement is closed.*

证明. First, suppose E^c is closed. Choose $x \in E$. Then $x \notin E^c$, and x is not a limit point of E^c. Hence there exists a neighborhood N of x such that $E^c \cap N$ is empty, that is, $N \subset E$. Thus x is an interior point of E, and E is open.

Next, suppose E is open. Let x be a limit point of E^c. Then every neighborhood of x contains a point of E^c, so that x is not an interior point of E. Since E is open, this means that $x \in E^c$. It follows that E is closed. □

Corollary. A set F is closed if and only if its complement is open.

mynotes:
这里使用新的定义得到的开集与闭集保持了原有的性质: 开集的补集是闭集, 闭集的补集是开集

Theorem 2.24. *(a) For any collection $\{G_\alpha\}$ of open sets, $\cup_\alpha G_\alpha$ is open.*

(b) For any collection $\{F_\alpha\}$ of closed sets, $\cap_\alpha F_\alpha$ is closed.

(c) For anyfinite collection $G_1, ..., G_n$ of open sets, $\cap_{i=1}^n G_i$ is open.

(d) For anyfinite collection $F_1, ..., F_n$ of closed sets, $\cup_{i=1}^n F_i$ is closed.

mynotes:
无限开集的并仍是开集, 有限开集的交仍是开集
无限闭集的交仍是闭集, 有限闭集的并仍是闭集
下面给出一个反例

Example 2.25. In parts (c)and (d) of the preceding theorem, the finiteness of the collections is essential.

$$G_n = \left(-\frac{1}{n}, \frac{1}{n}\ (n = 1, 2, 3, \dots).\right)$$

$G = \cap_{n=1}^\infty G_n$ Then G consists of a single point (namely, $x = 0$) and is therefore not an open subset of $\mathbb{R}$.

Thus the intersection of an infinite collection of open sets need not be open. Similarly, the union of an infinite collection of closed sets need not be closed.

Definition 2.26. If X is a metric space, if $E \subset X$, and if E' denotes the set of all limit points of E in X, then the *closure* of E is the set $\bar{E} = E \cup E'$.

Theorem 2.27. *If X is a metric space and $E \subset X$, then*

(a) E is closed,

(b) $E = \bar{E}$ if and only if E is closed,

(c) $E \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

By (a) and (c), E 1s the smallest closed subset of X that contains E.

Theorem 2.28. *Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup E$. Then $y \in \bar{E}$. Hence $y \in E$ if E is closed.*

Remark 2.29. Suppose $E \subset Y \subset X$, where X is a metric space. To say that E is an open subset of X means that to each point $p \in E$ there is associated a positive number r such that the conditions $d(p, q) < r$, $g \in X$ imply that $q \in E$. But we have already observed (Sec. 2.16) that Y is also a metric space, so that our definitions may equally well be made within Y. To be quite explicit, let us say that

E is *open relative* to Y if to each $p \in E$ there is associated an $r > 0$ such that $q \in E$ whenever $d(p, q) < r$ and $g \in Y$.

Example 2.21(g) showed that a set may be open relative to Y without being an open subset of X. However, there is a simple relation between these concepts, which we now state.

Theorem 2.30. *Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.*

2.3 Compact sets

Definition 2.31. By an *open cover* of a set E in a metric space X we mean a collection $\{G_\alpha\}$ of open subsets of X such that $E \subset \cup_\alpha G_\alpha$.

Definition 2.32. A subset K of a metric space X is said to be *compact* if every open cover of K contains a *finite* subcover.

More explicitly, the requirement is that if $\{G_\alpha\}$ is an open cover of K, then there are finitely many indices $\alpha_1, ..., \alpha_n$ such that

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}.$$

The notion of compactness is of great importance in analysis, especially in connection with continuity (Chap. 4).

It is clear that every finite set is compact. The existence of a large class of infinite compact sets in $\mathbb{R}^k$ will follow from Theorem 2.41.

We observed earlier (in Sec. 2.29) that if $E \subset Y \subset X$, then E may be open relative to Y without being open relative to X. The property of being open thus depends on the space in which E is embedded. The same is true of the property of being closed.

Compactness, however, behaves better, as we shall now see. To formulate the next theorem, let us say, temporarily, that K is compact relative to X if the requirements of Definition 2.32 are met.

Theorem 2.33. *Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.*

证明. Suppose K is compact relative to X, and let $\{V_\alpha\}$ be a collection of sets, open relative to Y, such that $K \subset \cup_\alpha V_\alpha$ theorem 2.30, there are sets G_α, open relative to X, such that $V_\alpha = Y \cap G_\alpha$, for all α; and since K is compact relative to X, we have

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}. \tag{2.18}$$

for some choice of finitely many indices $\alpha_1 ..., \alpha_n$. Since $K \subset Y$, 2.18 implies

$$K \subset V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}. \tag{2.19}$$

This proves that K is compact relative to Y.

Conversely, suppose K is compact relative to Y, let G_α be a collection of open subsets of X which covers K, and put $V_\alpha = Y \cap G_\alpha$. Then 2.19 will hold for some choice of $\alpha_1, ..., \alpha_n$; and since $V_\alpha = G_\alpha$, 2.19 implies 2.18.

This completes the proof. □

Theorem 2.34. *Compact subsets of metric spaces are closed.*

Theorem 2.35. *Closed subsets of compact sets are compact.*

Corollary. If F is closed and K is compact, then $F \cap K$ is compact.

Theorem 2.36. *If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\cap K_\alpha$ is nonempty.*

Corollary. If $\{K_\alpha\}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1} (n = 1, 2, 3, ...)$, then $\cap_1^\infty K_n$ is not empty.

Theorem 2.37. *If E is an infinite subset of a compact set K, then E has a limit point in K.*

Theorem 2.38. *If $\{I_n\}$ is a sequence of intervals in $\mathbb{R}^1$, such that $I_n \supset I_{n+1}, (n = 1, 2, 3, ...)$, then $\cap_1^\infty I_n$ is not empty.*

Theorem 2.39. *Let k be a positive integer. If I_n is a sequence of k-cells such that $I_n \supset I_{n+1}, (n = 1, 2, 3, ...)$, then $\cap_1^\infty I_n$ is not empty.*

Theorem 2.40. *Every k-cell is compact.*

Theorem 2.41. *If a set E in $\mathbb{R}^k$ has one of the following three properties, then it has the other two:*

(a) E is closed and bounded.

(b) E is compact.

(c) Every infinite subset of E has a limit point in E.

Theorem 2.42. *Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.*

证明. Being bounded, the set E in question is a subset of a k-cell $I \subset \mathbb{R}^k$. By Theorem 2.40, I is compact, and so E has a limit point in I, by Theorem 2.37. □

2.4 Perfect sets

Theorem 2.43. *Let P be a nonempty perfect set in $\mathbb{R}^k$. Then P is uncountable.*

Corollary. Every interval $[a, b](a < b)$ is uncountable. In particular, the set all real numbers is uncountable.

Definition 2.44. The Cantor set The set which we are now going to construct shows that there exist perfect sets in R! which contain no segment.

Let E_0 be the interval $[0, 1]$. Remove the segment $(\frac{1}{3}, \frac{2}{3})$, and let E_1 be the union of the intervals

$$[0, \frac{1}{3}] \quad [\frac{2}{3}, 1]$$

Remove the middle thirds of these intervals, and let E_2 be the union of the intervals

$$[0, \frac{1}{9}] \quad [\frac{2}{9}, \frac{3}{9}] \quad [\frac{6}{9}, \frac{7}{9}] \quad [\frac{8}{9}, 1]$$

Continuing in this way, we obtain a sequence of compact sets E_n, such that

(a) $E_1 \supset E_2 \supset E_3 \supset \dots$;

(b) E_n is the union of 2^n intervals, each of length 3^{-n}.

The set

$$P = \bigcap_{n=1}^{\infty} E_n$$

is called the *Cantor set*. P is clearly compact, and Theorem 2.36 shows that P is not empty.

No segment of the form

$$\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right) \tag{2.20}$$

where k and m are positive integers, has a point in common with P. Since every segment (α, β) contains a segment of the form (24), if

$$3^{-m} < \frac{\beta - \alpha}{6},$$

P contains no segment.

To show that P is perfect, it is enough to show that P contains no isolated point. Let $x \in P$, and let S be any segment containing x. Let I_n be that interval of E_n which contains x. Choose n large enough, so that $I_n \subset S$. Let x_n be an endpoint of I_n, such that $x_n \neq x$.

It follows from the construction of P that $x_n \in P$. Hence x is a limit point of P, and P is perfect.

One of the most interesting properties of the Cantor set is that it provides us with an example of an uncountable set of measure zero (the concept of measure will be discussed in Chap. 11).

2.5 connected sets

Definition 2.45. Two subsets A and B of a metric space X are said to be separated if both $A \cap \bar{B}$ and $\bar{A} \cap B$ are empty, i.e., if no point of A lies in the closure of B and no point of B lies in the closure of A.

A set $E \subset X$ is said to be connected if E is not a union of two nonempty separated sets.

Remark 2.46. Separated sets are of course disjoint, but disjoint sets need not be separated. For example, the interval $[0, 1]$ and the segment $(1, 2)$ are not separated, since 1 is a limit point of $(1, 2)$. However, the segments $(0, 1)$ and (1.2) are separated.

The connected subsets of the line have a particularly simple structure:

Theorem 2.47. *A subset E of the real line $\mathbb{R}^1$ is connected if and only if it has the following property: If $x \in E$, $y \in E$, and $x < z < y$, then $z \in E$.*

证明. If there exist $x \in E$, $y \in E$, and some $z \in (x, y)$ such that $z \notin E$, then $E = A_z \cup B_z$ where

$$A_z = E \cap (-\infty, z), \quad B_z = E \cap (z, \infty).$$

Since $x \in A_z$ and $y \in B_z$, A and B are nonempty. Since $A_z = (-\infty, z)$ and $B_z = (z, \infty)$, they are separated. Hence E is not connected.

To prove the converse, suppose E is not connected. Then there are nonempty separated sets A and B such that $A \cup B = E$. Pick $x \in A_z$ $y \in B_z$ and assume (without loss of generality) that $x < y$. Define

$$z = \sup(A \cap [x, y]).$$

By Theorem 2.28, $z \in \bar{A}$; hence $z \notin B$. In particular, $x \le z < y$.

If $z \notin A$, it follows that $x < z < y$ and $z \notin E$.

If $z \in A$, then $z \notin B$, hence there exists z, such that $z < z_1 < y$ and $z_1 \notin B$. Then $x < z_1 < y$ and $z_1 \notin E$. $\square$

mynotes

自己的笔记还是需要有自己的思考在里面

2.18 中的定义结合 2.21 的例子是非常重要的。

Definition 2.5.1. neighborhood

neighborhood $N_r(p)$, $\forall q, \exists r > 0$ s.t. $d(p, r) < r$.

Definition 2.5.2. limit point, isolated point, interior point

limit point, $p \in E$, $\forall N_r(p)$, $\exists q \in N_r(p)$, $q \neq p$ s.t. $q \in E$.

isolated point, $p \in E$, p is not a limit point.

interior point, for a point $p \in E$, $\exists N_r(p) \subset E$.

Definition 2.5.3. closed

E is *closed* if every limit point of E is a point of E.

Definition 2.5.4. open

E is *open* if every point of E is an interior point of E.

Theorem 2.5.1. *Every neighborhood is an open set.*

Definition 2.5.5. complement

The *complement* of E (denoted by E^c)is the set of all points $p \in X$ such that $p \notin E$.

Definition 2.5.6. perfect

(h) E is *perfect* if E is closed and if every point of E is a limit point of E.

Definition 2.5.7. bounded

(i) E is *bounded* if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$.

Definition 2.5.8. dense

(j) E is *dense* in X if every point of X is a limit point of E, or a point of E (or both).

定理 2.23 表明，虽然使用了看似不相关的定义，这里得到的开集与闭集仍然满足 “开集的补集是闭集，闭集的补集是开集” 这样直观的定理

Theorem 2.5.2. *A set E is open if and only if its complement is closed.*

Definition 2.5.9. open relative

E is open relative to Y, $\forall p \in E$, $\exists r > 0$, s.t. $q \in E$, $d(p, q) < r$, $q \in Y$.

任给 E 中一点, 存在邻域 N, N 是 Y 的子集. 称 E 对 Y 而言是开集.

2022.11.10

Definition 2.5.10. *Derived Set*

The limit points of a set P, denoted $P^{'}$.

Definition 2.5.11. *Perfect Set*

A set P is called perfect if $P = P^{'}$, where $P^{'}$ is the derived set of P.

`https://mathworld.wolfram.com/PerfectSet.html`

Definition 2.5.12. *Complete Space*

A space of functions comprising a complete biorthogonal system.

Definition 2.5.13. *Complete Metric Space*

A complete metric space is a metric space in which every Cauchy sequence is convergent.

Examples include the real numbers with the usual metric, the complex numbers, finite-dimensional real and complex vector spaces, the space of square-integrable functions on the unit interval $L^2([0, 1])$, and the p-adic numbers.

Chapter 3

Numerical sequences and series

3.1 Convergent sequences

Definition 3.1. A sequences $\{p_n\}$ in metric space X is said to converge if there is a point $p \in X$ with the following property:

For every $\varepsilon > 0$ there is an integer N s.t. $n \geq N$ implies that $d(p_n, p) < \varepsilon$. (Here d denotes the distance in X.)

In this case we also say that $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$. [see Th 3.2(b)], and we write $p_n \to p$, or

$$\lim_{n\to\infty} p_n = p.$$

if $\{p_n\}$ does not converge, it is said to diverge.

our definition of "convergent sequence" depends not only on $\{p_n\}$ but also on X. For instance, the sequence $\{1/n\}$ converges in $\mathbb{R}^1$(to 0), but fails to converge in the set of all positive real numbers [with $d(x, y) = |x - y|$]. In cases of possible ambiguity, we can be more precise and specify "convergent in X" rather than "convergent".

we recall that the set of all points $p_n (n = 1, 2, 3, ...)$ is the range of $\{p_n\}$. The range of a sequence may be a finite set, or it may be infinite. The sequence $\{p_n\}$ is said to be bounded if its range is bounded.

As examples, consider the following sequences of complex numbers (that is, $X = \mathbb{R}^2$):

(a) If $s_n = 1/n$, then $\lim_{n\to\infty} s_n = 0$; the range is infinite, and the sequence is bounded.

(b) If $s_n = n^2$ the sequence $\{s_n\}$ is unbounded, is divergent, and has infinite range.

(c) If $s_n = 1 + [(-1)^n/n]$, the sequence $\{s_n\}$ converges to 1, is bounded, and has infinite range.

(d) If $s_n = i^n$ the sequence $\{s_n\}$ is divergent, is bounded, and has finite range.

(e) If $s_n = 1 (n = 1, 2, 3, ...)$, then $\{s_n\}$ converges to 1, is bounded, and has finite range.

Theorem 3.2. *Let* $\{p_n\}$ *be a sequence in a metric space* X.

(a) $\{p_n\}$ *converges to* $p \in X$ *if and only if every neighborhood of* p *contains* p_n *for all but finitely many* n.

(b) If $p \in X$, $p' \in X$, *and if* $\{p_n\}$ *converges to* p *and to* p', *then* $p' = p$.

(c) If $\{p_n\}$ *converges, then* $\{p_n\}$ *is bounded.*

(d) If $E \subset X$ *and if* p *is a limit point of* E, *then there is a sequence* $\{p_n\}$ *in* E *such that* $p = \lim_{n\to\infty} p_n$.

证明. (d) For each positive integer n, there is a point $p_n \in E$ such that $d(p_n, p) < 1/n$. Given $\varepsilon > 0$, choose N so that $N\varepsilon > 1$. If $n > N$, it follows that $d(p_n, p) < \varepsilon$. Hence $p_n \to p$. □

Theorem 3.3. *Suppose* $\{s_n\}, \{t_n\}$ *are complex sequences, and* $\lim_{n\to\infty} s_n = s$, $\lim_{n\to\infty} t_n = t$. *Then*

(a) $\lim_{n\to\infty}(s_n + t_n) = s + t$;

(b) $\lim_{n\to\infty} cs_n = cs$, $\lim_{n\to\infty}(c + s_n) = c + s$, *for any number* c;

(c) $\lim_{n\to\infty} s_n t_n = st$;

(d) $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$, *provided* $s_n \neq 0 (n = 1, 2, 3, \dots)$, *and* $s \neq 0$.

Theorem 3.4. *(a) Suppose* $\mathbf{x}_n \in R^k (n = 1, 2, 3, \dots)$ *and*

$$\mathbf{x_n} = (\alpha_{1,n}, \dots \alpha_{k,n}).$$

Then $\{\mathbf{x}_n\}$ *converges to* $\mathbf{x} = (\alpha_1, \dots, \alpha_k)$ *if and only if*

$$\lim_{n\to\infty} \alpha_{j,n} = \alpha_j \qquad (1 \leq j \leq k). \tag{3.1}$$

(b) Suppose $\{\mathbf{x}_n\}$, $\{\mathbf{y}_n\}$ *are sequences in* $\mathbb{R}^k$, $\{\beta_n\}$ *is a sequence of real numbers, and* $\mathbf{x}_n \to \mathbf{x}$, $\mathbf{y}_n \to \mathbf{y}$, $\beta_n \to \beta$. *Then*

$$\lim_{n\to\infty}(\mathbf{x_n} + \mathbf{y_n}) = \mathbf{x} + \mathbf{y}, \quad \lim_{n\to\infty} \mathbf{x_n} \cdot \mathbf{y_n} = \mathbf{x} \cdot \mathbf{y}, \quad \lim_{n\to\infty} \beta_n \mathbf{x_n} = \beta \mathbf{x}.$$

3.2 Subsequences

Definition 3.5. Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of positive integers, such that $n_1 < n_2 < n_3 < \dots$. Then the sequence $\{p_{n_i}\}$ is called a *subsequence* of $\{p_n\}$. If $\{p_{n_i}\}$ converges, its limit is called a subsequential limit of $\{p_n\}$.

It is clear that $\{p_n\}$ converges to p if and only if every subsequence of $\{p_n\}$ converges to p. We leave the details of the proof to the reader.

Theorem 3.6. *(a) If* $\{p_n\}$ *is a sequence in a compact metric space* X, *then some subsequence of* $\{p_n\}$ *converges to a point of* X.

(b) Every bounded sequence in $\mathbb{R}^k$ *contains a convergent subsequence.*

Theorem 3.7. *The subsequential limits of a sequence* $\{p_n\}$ *in a metric space* X *form a closed subset of* X.

3.3 Cauchy sequences

Definition 3.8. A sequence $\{p_n\}$ in a metric space X is said to be a *Cauchy sequence* if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < e$ if $n \geq N$ and $m \geq N$.

In our discussion of Cauchy sequences, as well as in other situations which will arise later, the following geometric concept will be useful.

Definition 3.9. Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of the form $d(p, q)$, with $p \in E$ and $q \in E$. The sup of S is called the diameter of E.

If $\{p_n\}$ is a sequence in X and if E_N consists of the points $p_N, p_{N+1}, p_{N+2}, \dots$, it is clear from the two preceding definitions that $\{p_n\}$ is a *Cauchy sequence if and only if*

$$\lim_{N\to\infty} \text{diam } E_N = 0.$$

Theorem 3.10. *(a) If $\bar{E}$ is the closure of a set E in a metric space X, then*

$$\text{diam E} = \text{diam E}.$$

(b) If K_n is a sequence of compact sets in X such that $K_n \supset K_{n+1}$ $(n = 1, 2, 3, ...)$and if

$$\lim_{n\to\infty} \text{diam K}_\text{n} = 0,$$

then $\cap_1^\infty K_n$ consists of exactly one point.

Theorem 3.11. *(a) Inany metric space X, every convergent sequence is a Cauchy sequence.*

(b) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point of X.

(c) In $\mathbb{R}^k$, every Cauchy sequence converges.

Note: The difference between the definition of convergence and the definition of a Cauchy sequence is that the limit is explicitly involved in the former, but not in the latter. Thus Theorem 3.11(b) may enable us to decide whether or not a given sequence converges without knowledge of the limit to which it may converge.

The fact (contained in Theorem 3.11) that a sequence converges in $\mathbb{R}^k$ if and only if it is a Cauchy sequence is usually called the *Cauchy criterion* for convergence.

Definition 3.12. A metric space in which every Cauchy sequence converges is said to be *complete*.

Thus Theorem 3.11 says that *all compact metric spaces and all Euclidean spaces are complete*. Theorem 3.11 implies also that every closed subset E of a complete metric space X is complete. (Every Cauchy sequence in E is a Cauchy sequence in X. hence it converges to some $p \in X$, and actually $p \in E$ since E is closed.) An example of a metric space which is not complete is the space of all rational numbers, with $d(x, y) = |x - y|$.

Theorem 3.2(c) and example (d) of Definition 3.1 show that convergent sequences are bounded, but that bounded sequences in $\mathbb{R}^k$ need not converge. However, there is one important case in which convergence is equivalent to boundedness; this happens for monotonic sequences in $\mathbb{R}^1$.

Definition 3.13. A sequence $\{s_n\}$ of real numbers is said to be

(a) monotonically increasing if $s_n \leq s_{n+1}$ $(n = 1, 2, 3, ...)$;

(b) monotonically decreasing if $s_n \geq s_{n+1}$ $(n = 1, 2, 3, ...)$.

Theorem 3.14. *Theorem Suppose $\{s_n\}$ is monotonic. Then $\{s_n\}$ converges if and only if it is bounded.*

3.4 Upper and lower limits

Definition 3.15. Let $\{s_n\}$ be a sequence of real numbers with the following property: For every real M there is an integer N such that $n > N$ implies $s_n \geq M$. We then write

$$s_n \to +\infty.$$

Similarly, if for every real M there is an integer N such that $n > N$ implies $s_n \leq M$, we write

$$s_n \to -\infty.$$

It should be noted that we now use the symbol $rightarrow$ (introduced in Definition 3.1) for certain types of divergent sequences, as well as for convergent sequences, but that the definitions of convergence and of limit, given in Definition 3.1, are in no way changed.

Definition 3.16. Let $\{s_n\}$ be a sequence of real numbers. Let E be the set of numbers x (in the extended real number system) such that $s_n rightarrow x$ for some subsequence $\{s_n\}$. This set E contains all subsequential limits as defined in Definition 3.5, plus possibly the numbers $+\infty$, $-\infty$.

We now recall Definitions 1.8 and 1.23 and put

$$s^* = \sup E, s_* = \inf E,$$

The numbers s^*, s_*, are called the upper and lower limits of $\{s_n\}$; we use the notation

$$\limsup_{n\to\infty} s_n = s^*, \quad \liminf_{n\to\infty} s_n = s_*.$$

Theorem 3.17. *Let $\{s_n\}$ be a sequence of real numbers. Let E and s^* have the same meaning as in Definition 3.16. Then $s*$ has the following two properties:*

(a) $s^* \in E$.

(b) If $x > s^*$, there is an integer N such that $n > N$ implies $s_n < x$.

Moreover, s^* is the only number with the properties (a) and (b).

Of course, an analogous result is true for s_*.

Example 3.18. (a) Let $\{s_n\}$ be a sequence containing all rationals. Then every real number is a subsequential limit, and

$$\limsup_{n\to\infty} = +\infty, \quad \liminf_{n\to\infty} = -\infty.$$

(b) Let $s_n = (-1^n)/[1+(1/n)]$. Then

$$\limsup_{n\to\infty} = 1, \quad \liminf_{n\to\infty} = -1.$$

(c) For a real-valued sequence $\{s_n\}$, $\lim_{n\to\infty} s_n = s$ if and only if

$$\limsup_{n\to\infty} = \liminf_{n\to\infty} = s.$$

Theorem 3.19. *If $s_n \le t_n$ for $n \ge N$, where N is fixed, then*

$$\liminf_{n\to\infty} s_n \le \liminf_{n\to\infty} t_n,$$
$$\limsup_{n\to\infty} s_n \le \limsup_{n\to\infty} t_n.$$

3.5 Some special sequences

some sequences occur frequently. remark: If $0 \le x_n \le s_n$ for $n \ge N$, where N is some fixed number, and if $s_n \to 0$, then $x_n \to 0$.

Theorem 3.20. *(a) If $p > 0$, then $\lim_{n\to\infty} \frac{1}{n^p} = 0$.*

(b) If $p > 0$, then $\lim_{n\to\infty} \sqrt[n]{p} = 1$.

(c) $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

(d) If $p > 0$ and α is real, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$.

(e) If $|x| < 1$, then $\lim_{n\to\infty} x^n = 0$.

3.6 Series

Consider complex-valued sequences and series

Definition 3.21. Given a sequence $\{a_n\}$, we use the notation

$$\sum_{n=p}^{q} a_n \quad (p \le q)$$

to denote the sum $a_p + a_{p+1} + \cdots + a_q$. With $\{a_n\}$ we associate a sequence $\{s_n\}$, where

$$s_n = \sum_{k=1}^{n} a_k.$$

For $\{s_n\}$ we also use the symbolic expression

$$a_1 + a_2 + a_3 + \ldots$$

or, more concisely

$$\sum_{n=1}^{\infty} a_n. \tag{3.2}$$

we call this *infinite series*, or just a *series*. The numbers $\{s_n\}$ are called the *partial sums* of the series. If $\{s_n\}$ converges to s, we say that that the series *converges*, and write

$$\sum_{n=1}^{\infty} a_n = s.$$

The number s is called the sum of the series; but it should be clearly understood that s is the *limit of a sequence of sums*, and is not obtained simply by addition.

If $\{s_n\}$ diverges, the series is said to diverge.

Sometimes, for convenience of notation, we shall consider series of the form

$$\sum_{n=0}^{\infty} a_n. \tag{3.3}$$

And frequently, when there is no possible ambiguity, or when the distinction is immaterial, we shall simply write $\sum a_n$, in place of (4) or (5).

It is clear that every theorem about sequences can be stated in terms of series (putting $a_1 = s_1$, and $a_n = s_n - s_{n-1}$ for $n > 1$), and vice versa. But it is nevertheless useful to consider both concepts.

The Cauchy criterion (Theorem 3.11) can be restated in the following form:

Theorem 3.22. *$\sum a_n$ converges if and only if for every $\varepsilon \in> 0$ there is an integer N such that*

$$\left| \sum_{k=n}^{m} a_k \right| \leq \varepsilon \tag{3.4}$$

if $m \geq n \geq N$.

In particular, by taking $m = n$, (6) becomes

$$|a_n| \leq \varepsilon \quad (n \geq N).$$

Theorem 3.23. *If $\sum a_n$ converges, then* $\lim_{n\to\infty} a_n = 0$.

The condition $a_n \to 0$ is not sufficient to ensure convergence of $\sum a_n$. For instance, the series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges; for the proof we refer to Theorem 3.28.

Theorem 3.14, concerning monotonic sequences, also has an immediate counterpart for series.

Theorem 3.24. *A series of nonnegative terms converges if and only if its partial sums form a bounded sequence.*

Theorem 3.25. *(a) If $|a_n| \le c_n$, for $n \ge N_0$, where N_0 is some fixed integer, and if $\sum c_n$ converges, then $\sum a_n$ converges.*

(b) If $a_n \ge d_n \ge 0$ for $n \ge N_0$, and if $\sum d_n$, diverges, then $\sum a_n$ diverges.

mynotes:
比较审敛法

3.7 Series of nonnegative terms

Theorem 3.26. *If $0 \le x < 1$, then*

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

If $x \ge 1$, the series diverges.

mynotes:
几何级数收敛条件

mynotes:
Cauchy use "thin" subsequence of $\{a_n\}$ determines the convergence or divergence of $\sum a_n$

Theorem 3.27. *Suppose $a_1 \ge a_2 \ge a_3 \ge \cdots \ge 0$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series*

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \dots \tag{3.5}$$

converges.

证明. By Theorem 3.24, it suffices to consider boundedness of the partial sums. Let

$$s_n = a_1 + a_2 + \cdots + a_n,$$
$$t_n = a_1 + 2a_2 + \cdots + 2^k a_{2^k}.$$

For $n < 2^k$,

$$\begin{aligned} s_n &\le a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\le a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\ &= t_k, \end{aligned}$$

so that

$$s_n \le t_k. \tag{3.6}$$

On the other hand, if $n > 2^k$,

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\geq \frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{k+1}a_{2^k} \\ &= \frac{1}{2}t_k, \end{aligned}$$

so that

$$2s_n \geq t_k. \tag{3.7}$$

$\{s_n\}$, $\{t_n\}$ are both bounded or both unbounded. $\square$

Theorem 3.28. $\sum \frac{1}{n^p}$ *converges if* $p > 1$ *and diverges if* $p \leq 1$.

Theorem 3.29. *If* $p > 1$ *,*

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p} \tag{3.8}$$

converges; if $p \leq 1$ *, the series diverges.*

"$\log n$" the logarithm of n to the base e (compare Exercise 7, Chap. 1); the number e will be defined in a moment (see Def 3.30). We let the series start with $n = 2$, since $\log 1 = 0$.

证明. The monotonicity of the logarithmic function (which will be discussed in more detail in Chap. 8) implies that ($\log n$) increase. Hence ($1/n \log n$) decreases, and we can apply Theorem 3.27 to (3.8); this leads us the series

$$\sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k(\log 2^k)^p} = \sum_{k=1}^{\infty} \cdot \frac{1}{(k\log 2)^p} = \frac{1}{(\log 2)^p} \sum_{k=1}^{\infty} \cdot \frac{1}{k^p} \tag{3.9}$$

and Theorem 3.29 follows from Theorem 3.28. $\square$

This procedure may evidently be continued. For instance,

$$\sum_{n=3}^{\infty} \frac{1}{n \log n \log\log n} \tag{3.10}$$

diverges, whereas

$$\sum_{n=3}^{\infty} \frac{1}{n \log n (\log\log n)^2} \tag{3.11}$$

converges.

Series (3.10) differ very littel from (3.11). Still, one diverges, the other converges. If we continue the process which led us from Theorem 3.28 to Theorem 3.29, we get pairs of convergent and divergent series whose terms differ even less than those of (3.10) and (3.11). One might thus be led to the conjecture that there is a limiting situation of some sort, a "boundary" with all convergent series on one side, all divergent series on the other side — at least as far as series with monotonic coefficients are converned. This notion of "boundary" is of course quite vague. The point we wish to make is this: No matter how we make this notion precise, the conjecture is false. Exercises 11(b) and 12(b) may serve as illustrations.

More deeper aspect of convergence theory can refer to Knopp's *"Theory and Application of Infinite Series"*, Chap IX, particularly Sec. 41.

3.8 The number e

Definition 3.30. $e = \sum_{n=0}^{\infty} \frac{1}{n!}$.

Here $n! = 1 \cdot 2 \cdot 3 \cdots n$ if $n \geq 1$, and $0! = 1$.

Since

$$\begin{aligned} s_n &= 1 + 1 + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \cdots + \frac{1}{1 \cdot 2 \cdots n} \\ &< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} < 3 \end{aligned}$$

The series converges, and the definition makes sense. In fact, the series converges very rapidly and allows us to compute e with great accuracy.

It is of interest to note that e can also be difined by means of another limit process; the proof provides a good illustration of operations with limits:

Theorem 3.31. $\lim_{n\to\infty}(1 + 1/n)^n = e$.

mynotes:
Is this equation found by Bernoulli?

证明. Let

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}, \quad t_n = \sum_{k=0}^{n} (1 + \frac{1}{n})^n.$$

by the binomial theorem,

$$\begin{aligned} t_n = 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots \\ + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{n-1}{n}\right). \end{aligned}$$

Hence $t_n \leq s_n$, so that

$$\limsup_{n\to\infty} t_n \leq e, \tag{3.12}$$

by Theorem 3.19. Next, if $n \geq m$,

$$t_n \geq 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \cdots + \frac{1}{m!}\left(1 - \frac{1}{n}\right) + \cdots + \frac{1}{m!}\left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right).$$

Let $n \to \infty$, kepping m fixed. We get

$$\liminf_{n\to\infty} t_n \geq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!},$$

so that

$$s_m \leq \liminf_{n\to\infty} t_n,$$

Letting $m \to \infty$, we finally get

$$e \leq \liminf_{n\to\infty} t_n. \tag{3.13}$$

The Theorem follows from (3.12) and (3.13). □

The rapidly with which the series $\sum 1/n!$ converges can be estimated as follows: If s_n has the same meaning as above, we have

$$\begin{aligned} e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots \\ &< \frac{1}{(n+1)!}\left\{1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right\} = \frac{1}{n!n} \end{aligned}$$

so that

$$0 < e - s_n < \frac{1}{n!n}. \tag{3.14}$$

Thus s_{10}, for instance, approximates e with an error less than 10^{-7}. The inequality (3.14) is of theoretical interest as well, since it enables us to prove the irrationality of e very easily.

Theorem 3.32. *e is irrational.*

证明. Suppose e is rational. Then $e = p/q$, where p and q are positive integers. By (3.14),

$$0 < q!(e - s_q) < \frac{1}{q}. \tag{3.15}$$

By our assumption, $q!e$ is an integer. Since

$$q!s_q = q!\left(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{q!}\right)$$

is an integer, we see that $q!(e - s_q)$ is an integer.

Since $q \geq 1$, (3.15) implies the existence of an integer between 0 and 1. We have thus reached a contradiction. □

Actually, e is not even an algebraic number. For a simple proof of this, see page 25 of Niven's book, or page 176 of Herstein's, cited in the Bibliography.

3.9 The root and ratio tests

Theorem 3.33. *(Root test) Given* $\sum a_n$ *, put* $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.

Then

(a) if $\alpha < 1$, $\sum a_n$ *converges;*

(a) if $\alpha > 1$, $\sum a_n$ *diverges;*

(a) if $\alpha = 1$, *The test gives no information.*

Theorem 3.34. *(Ratio test) The series* $\sum a_n$

(a) converges if $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$.

(b) diverges if $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ *for all* $n \geq n_0$, *where* n_0 *is some fixed integer.*

mynotes:
3.33 为根值审敛法，3.34 为比值审敛法

Note: The knowledge that $\lim a_{n+1}/a_n = 1$ implies nothing about the convergence of $\sum a_n$. The series $\sum 1/n$ and $\sum 1/n^2$ demonstrate this.

Example 3.35. (a) Consider the series

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{2^4} + \frac{1}{3^4} + \cdots ,$$

for which

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0,$$

$$\liminf_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[2n]{\frac{1}{3^n}} = \frac{1}{\sqrt{3}},$$

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{2}\left(\frac{3}{3}\right)^n = +\infty,$$

$$\limsup_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \sqrt[2n]{\frac{1}{2^n}} = \frac{1}{\sqrt{2}}.$$

The root test indicates convergence: the ratio test does not apply.

(b) The same is true for the series

$$\frac{1}{2} + 1 + \frac{1}{8} + \frac{1}{4} + \frac{1}{32} + \frac{1}{16} + \frac{1}{128} + \frac{1}{64} + \cdots ,$$

where

$$\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{8},$$

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = 2,$$

but

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \frac{1}{2}.$$

Remark 3.36. The ratio test is frequently easier to apply than the root test, since it is usually easier to compute ratios than nth roots. However, the root test has wider scope. More precisely: Whenever the ratio test shows convergence, the root test does too; whenever the root test is inconclusive, the ratio test is too. This is a consequence of Theorem 3.37, and is illustrated by the above examples.

Neither of the two tests is subtle with regard to divergence. Both deduce divergence from the fact that a_n does not tend to zero as $n \to \infty$.

Theorem 3.37. *For any sequence $\{c_n\}$ of positive numbers,*

$$\liminf_{n\to\infty} \frac{c_{n+1}}{c_n} \le \liminf_{n\to\infty} \sqrt[n]{c_n},$$
$$\limsup_{n\to\infty} \sqrt[n]{c_n} \le \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

证明. We shall prove the second inequality; the proof of the first is quite similar. Put

$$\alpha = \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

If $\alpha = +\infty$, there is nothing to prove. If α is finite , choose $\beta > \alpha$. There is an integer N such that

$$\frac{c_{n+1}}{c_n} \le \beta$$

for $n \ge N$. In particular, for any $p > 0$,

$$c_{N+k+1} \le \beta c_{N+k} \quad (k = 0, 1, \ldots, p-1).$$

Multiplying these inequalities, we obtain

$$c_{N+p} \le \beta^p c_N,$$

or

$$c_n \le c_N \beta^{-N} \cdot \beta^n \quad (n \ge N).$$

Hence

$$\sqrt[n]{c_n} \le \sqrt[n]{c_N \beta^{-N}} \cdot \beta,$$

so that

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \le \beta, \tag{3.16}$$

by Theorem 3.20(b). Since (3.16) is true for every $\beta > \alpha$, we have

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \le \alpha.$$

□

3.10 Power series

Definition 3.38. Given a sequence $\{c_n\}$ of complex numbers, the series

$$\sum_{n=0}^{\infty} c_n z^n \tag{3.17}$$

is called a *power series*. The numbers c_n are called the *coefficients* of the series; z is a complex number.

In general, the series will converge or diverge, depending on the choice of z. More specifically, with every power series there is associated a circle, the circle of convergence, such that (3.17) converges if z is in the interior of the circle and diverges if z is in the exterior (to cover all cases, we have to consider the plane as the interior of a circle of infinite radius, and a point as a circle of radius zero). The behavior on the circle of convergence is much more varied and cannot be described so simply.

Theorem 3.39. *Given the power series $\sum c_n z^n$, put*

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|}, \quad R = \frac{1}{\alpha}.$$

(If $\alpha = 0$, $R = +\infty$; If $\alpha = +\infty$, $R = 0$.) Then $\sum c_n z^n$ converges of $|z| < R$, and diverges if $|z| > R$.

证明. Put $a_n = c_n z^n$, and apply the root test:

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{|z|}{R}.$$

□

Note: R is called the radius of convergence of $\sum c_n z^n$.

Example 3.40. (a) The series $\sum n^n z^n$ has $R = 0$.

(b) The series $\sum z^n/n!$ has $R = +\infty$. (In this case the ratio test is easier to apply than the root test.)

(c) The series $\sum z^n$ has $R = 1$. If $|z| = 1$, the series diverges, since $\{z^n\}$ does not tend to 0 as $n \to \infty$.

(d) The series $\sum z^n/n$ has $R = 1$. It diverges if $z = 1$. It converges for all other z with $|z| = 1$. (The last assertion will be proved in Theorem 3.44)

(e) The series $\sum z^n/n^2$ has $R = 1$. It converges for all z with $z = 1$, by the comparison test, since $|z^n/n^2| = 1/n^2$.

3.11 Summation by parts

Theorem 3.41. *Given two sequences $\{a_n\}$, $\{b_n\}$, put*

$$A_n = \sum_{k=0}^{n} a_k$$

if $n \geq 0$; put $A_{-1} = 0$. Then if $0 \leq p \leq q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p. \tag{3.18}$$

证明.
$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q}(A_n - A_{n-1})b_n = \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1}$$
and the last expression on the right is clearly equal to the right side of (3.18). □

Formula (20), the so-called "partial summation formula," is useful in the investigation of series of the form $\sum a_n b_n$, particularly when $\{b_n\}$ is monotonic. We shall now give applications.

Theorem 3.42. *Suppose*
(a) the partial sums A_n of $\sum a_n$ form a bounded sequence;
(b) $b_0 \geq b_1 \geq b_2 \geq \cdots$;
(c) $\lim_{n\to\infty} b_n = 0$.
Then $\sum a_n b_n$ converges.

证明. Choose M such that $|A_n| \leq M$ for all n. Given $\varepsilon > 0$, there is an integer N such that $b_N \leq (\varepsilon/2M)$. For $N \leq p \leq q$, we have

$$\begin{aligned}
\left|\sum_{n=p}^{q} a_n b_n\right| &= \left|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p\right| \\
&\leq M \left|\sum_{n=p}^{q-1}(b_n - b_{n+1}) + b_q + b_p\right| \\
&= 2Mb_p \leq 2Mb_N \leq \varepsilon.
\end{aligned}$$

Convergence now follows from the Cauchy criterion. We note that the first inequality in the above chain depends of course on the fact that $b_n - b_{n+1} \geq 0$. □

Theorem 3.43. *Suppose*
(a) $c_1 \geq c_2 \geq c_3 \geq \cdots$;
(b) $c_{2m-1} \geq 0, c_{2m} \leq 0 \quad (m = 1, 2, 3, \ldots)$;
(c) $\lim_{n\to\infty} c_n = 0$.
Then $\sum c_n$ converges.

Series for which (b) holds are called “alternating series”; the theorem was known to Leibnitz.

证明. Apply Theorem 3.43, with $a_n = (-1)^{n+1}$, $b_n = |c_n|$. □

Theorem 3.44. *Suppose the radius of convergence of* $\sum c_n z^n$ *is* 1*, and suppose* $c_0 \geq c_1 \geq c_2 \geq \cdots$, $\lim_{n\to\infty} c_n = 0$. *Then* $\sum c_n z^n$ *converges at every point on the circle* $|z| = 1$*, except possibly at* $z = 1$.

证明. Put $a_n = z^n$, $b_n = c_n$. The hypotheses of Theorem 3.42 are then satisfied, since

$$|A_n| = \left|\sum_{m=0}^{n} z^m\right| = \left|\frac{1-z^{n+1}}{1-z}\right| \leq \frac{2}{|1-z|},$$

if $|z| = 1, z \neq 1$. □

3.12 Absolute convergence

The series $\sum a_n$ i said to *converge absolutely* if the series $\sum |a_n|$ converges.

Theorem 3.45. *if* $\sum a_n$ *converges absolutely, then* $\sum a_n$ *converges.*

证明. The assertion follow from the inequality

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k|,$$

plus the Cauchy criterion. □

Remark 3.46. For series of positive terms, absolute convergence is the same as convergence.

If $\sum a_n$ converges, but $\sum |a_n|$ diverges, we say that $\sum a_n$ converges *nonabsolutely*. For instance, the series

$$\sum \frac{(-1)^n}{n}$$

converges nonabsolutely (Theorem 3.43).

The comparison test, as well as the root and ratio tests, is really a test for absolute convergence, and therefore cannot give any information about nonabsolutely convergent series. Summation by parts can sometimes be used to handle the latter. In particular, power series converge absolutely in the interior of the circle of convergence.

We shall see that we may operate with absolutely convergent series very much as with finite sums. We may multiply thm term by term and we may change the order in which the additions are carried out, without affecting the sum of series. But for nonabsolutely convergent series this is no longer true, and more care has to be taken when dealing with them.

3.13 Addition and multiplication of series

Theorem 3.47. *If $\sum a_n = A$, and $\sum b_n = B$, then $\sum(a_n + b_n) = A + B$, and $\sum ca_n = cA$, for any fixed c.*

证明. Let

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} b_k.$$

Then

$$A_n + B_n = \sum_{k=0}^{n}(a_k + b_k).$$

Since $\lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} B_n = B$, we see that

$$\lim_{n\to\infty} (A_n + B_n) = A + B.$$

The proof of the second assertion is even simpler. □

Thus two convergent series may be added term by term, and the resulting series converges to the sum pf the two series. The situation becomes more complicated when we consider multiplication of two series. To begin with, we have to define the product. This can be done in several ways; we shall consider the so-called "Cauchy product".

Definition 3.48. Given $\sum a_n$ and $\sum b_n$, we put

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} \quad (n = 0, 1, 2, \dots)$$

and call $\sum c_n$ the *product* of the two given series.

This definition may be motivated as follows. If we take two power series $\sum a_n z^n$ and $\sum b_n z^n$, multiply them term by term, and collect ter,s containing the same power of z, we get

$$\begin{aligned}\sum_{n=0}^{\infty} a_n z^n \cdot \sum_{n=0}^{\infty} b_n z^n &= (a_0 + a_1 z + a_2 z^2 + \cdots)(b_0 + b_1 z + b_2 z^2 + \cdots) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) z + (a_0 b_2 + a_1 b_1 + a_2 b_0) z^2 + \cdots \\ &= c_0 + c_1 z + c_2 z^2 + \cdots .\end{aligned}$$

Setting $z = 1$, we arrive at the above definition.

Example 3.49. If

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} b_k, \quad C_n = \sum_{k=0}^{n} c_k,$$

and $A_n \to A$, $B_n \to B$, then it is not an all clear that $\{C_n\}$ will converge to AB, since we do not have $C_n = A_n B_n$. The dependence of $\{C_n\}$ on $\{A_n\}$ and $\{B_n\}$ is quite a complicated one (see the proof of Theorem 3.50). We shall now show that the product of two convergent series may actually diverge.

The series

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}} = 1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \cdots$$

converges (Theorem 3.43). we form the product of this series with itself and obtain

$$\sum_{n=0}^{\infty} c_n = 1 - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{3} + \frac{1}{\sqrt{2}\sqrt{2}} + \frac{1}{\sqrt{3}}}\right) + \left(\frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}\sqrt{2}} + \frac{1}{\sqrt{2}\sqrt{3}} + \frac{1}{\sqrt{4}}\right) + \cdots,$$

so that

$$c_n = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(n-k+1)(k+1)}}.$$

Since

$$(n-k+1)(k+1) = \left(\frac{n}{2}+1\right)^2 - \left(\frac{n}{2}-k\right)^2 \leq \left(\frac{n}{2}+1\right)^2.$$

we have

$$|c_n| \geq \sum_{k=0}^{n} \frac{2}{n+2} = \frac{2(n+1)}{n+2},$$

so that the condition $c_n \to 0$, which is necessary for the Convergence of $\sum c_n$, is not satisfied.

In the view of the next theorem, due to Mertens, we note that we have here considered the product of two nonabsolutely convergent series.

Theorem 3.50. *Suppose*
(a) $\sum_{n=0}^{\infty} a_n$ *converges absolutely,*
(b) $\sum_{n=0}^{\infty} a_n = A$*,*
(c) $\sum_{n=0}^{\infty} b_n = B$*,*
(d) $c_n = \sum_{k=0}^{n} a_k b_{n-k} \qquad (n = 0, 1, 2, \dots)$. *Then*

$$\sum_{n=0}^{\infty} c_n = AB.$$

That is, the product of two convergent series Converges, and to the right value, if at least one of the two series converges absolutely.

证明. Put

$$A_n = \sum_{k=0}^{n} a_k, \quad B_n = \sum_{k=0}^{n} b_k, \quad C_n = \sum_{k=0}^{n} c_k, \quad \beta_n = B_n - B.$$

Then

$$
\begin{aligned}
C_n &= a_0b_0 + (a_0b_1 + a_1b_0) + \cdots + (a_0b_n + a_1b_{n-1} + \cdots + a_nb_0) \\
&= a_0B_n + a_1B_{n-1} + \cdots + a_nB_0 \\
&= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \cdots + a_n(B + \beta_0) \\
&= A_nB + a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0
\end{aligned}
$$

Put

$$\gamma_n = a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0.$$

We wish to show that $C_n \to AB$. Since $A_nB \to AB$, it suffices to show that

$$\lim_{n\to\infty} \gamma_n = 0. \tag{3.19}$$

Put

$$\alpha = \sum_{n=0}^{\infty} |a_n|.$$

[It's here that we use (a).] Let $\varepsilon > 0$ be given. By (c), $\beta_n \to 0$. Hence we can choose N such that $|\beta_n| \le \varepsilon$ for $n \ge N$, in which case

$$
\begin{aligned}
|\gamma_n| &\le |\beta_0a_n + \cdots + \beta_Na_{n-N}| + |\beta_{N+1}a_{n-N-1} + \cdots + \beta_na_0| \\
&\le |\beta_0a_n + \cdots + \beta_Na_{n-N}| + \varepsilon\alpha.
\end{aligned}
$$

Keeping N fixed, and letting $n \to \infty$, we get

$$\limsup_{n\to\infty} |\gamma_n| \le \varepsilon\alpha,$$

since $a_k \to 0$ as $k \to \infty$. Since ε is arbitrary, (3.19) follows. □

Another question which may be asked is whether the series $\sum c_n$, if convergent, must have the sum AB. Abel showed that the answer is in the affirmative.

Theorem 3.51. *If the series $\sum a_n$, $\sum b_n$, $\sum c_n$ Converge to A, B, C , and $c_n = a_0b_n + \cdots + a_nb_0$, then $C = AB$.*

Here no assumption is made concerning absolute convergence. We shall give a simple proof (which depends on the continuity of power series) after Theorem 8.2.

3.14 Rearrangements

Definition 3.52. Let $\{k_n\}, n = 1, 2, 3, \dots$, be a sequence in which every positive integer appears once and only once (that is, $\{k_n\}$ is a 1-1 function from J onto J, in the notation of Definition 2.2). Putting

$$a'_n = a_{k_n} \quad (n = 1, 2, 3, \dots),$$

we say that $\sum a'_n$ is a rearrangements of $\sum a_n$.

If $\{s_n\}$, $\{s'_n\}$ are the sequences of partial sums of $\{a_n\}$, $\{a'_n\}$, it is easily seen that, in general, these two sequences consist of entirely different numbers. We are thus led to the problem of determining under what conditions all rearrangements of a convergent series will converge and whether the sums aare necessarily the same.

Example 3.53. Consider the convergent series

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}-\frac{1}{6}+\cdots \tag{3.20}$$

and one of its rearrangements

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\frac{1}{9}+\frac{1}{11}-\frac{1}{6}+\cdots \tag{3.21}$$

in which two positive terms are always followed by one negative. If s is the sum of (3.20), then

$$s<1-\frac{1}{2}+\frac{1}{3}=\frac{5}{6}.$$

Since

$$\frac{1}{4k-3}+\frac{1}{4k-1}-\frac{1}{2k}>0$$

for $k\geq 1$, we see that $s'_3<s'_6<s'_9<\cdots$, where s'_n is nth partial sum of (3.21).
Hence

$$\limsup_{n\to\infty} s'_n>s'_3=\frac{5}{6},$$

so that (3.21) certainly does not converge to s. [verigy (3.21) converge] my:

$$\begin{aligned}\frac{1}{4k-3}+\frac{1}{4k-1}-\frac{1}{2k}&=\frac{(4k-3)+(4k-1)}{(4k-3)(4k-1)}-\frac{1}{2k}\\&=\frac{(8k-4)}{(4k-3)(4k-1)}-\frac{1}{2k}\\&=\frac{(8k-4)2k-(4k-3)(4k-1)}{(4k-3)(4k-1)2k}\\&=\frac{8k-3}{(4k-3)(4k-1)2k}\\&<\frac{4}{(4k-3)(4k-1)}\\&<\frac{4}{(4(k-1))^2}=\frac{}{4(k-1)^2}\end{aligned}$$

because $\sum 1/n^2$ converge, (3.21) converge.

This example illustrates the following theorem, due to Riemann.

Theorem 3.54. *Let $\sum a_n$ be a series of real numbers which converges, but not absolutely. Suppose*

$$-\infty\leq\alpha\leq\beta\leq\infty.$$

Then there exists a rearrangement $\sum a'_n$ with partial sum $\sum s'_n$ such that

$$\liminf_{n\to\infty} s'_n=\alpha,\quad \limsup_{n\to\infty} s'_n=\beta. \tag{3.22}$$

证明. Let

$$p_n = \frac{|a_n| + a_n}{2}, \quad q_n = \frac{|a_n| - a_n}{2}, \quad (n = 1, 2, 3, \dots).$$

Then $p_n - q_n = a_n$, $p_n + q_n = a_n$, $p_n \geq 0$, $q_n \geq 0$. The series $\sum p_n, \sum q_n$ must both diverge.

For if both were convergent, then

$$\sum (p_n + q_n) = \sum |a_n|$$

would converge, contrary to hypothesis. Since

$$\sum_{n=1}^{N} a_n = \sum_{n=1}^{N} (p_n - q_n) = \sum_{n=1}^{N} p_n - \sum_{n=1}^{N} q_n,$$

divergence of $\sum p_n$ and convergence of $\sum q_n$ (or vice versa) implies divergence of $\sum a_n$, again contrary to hypothesis.

Now let $P_1, P_2, P_3, \dots$ denote the nonnegative terms of $\sum a_n$, in the order in which they occur, and let $Q_1, Q_2, Q_3, \dots$ be the absolute values of the negative terms of $\sum a_n$, also in their original order.

The series $\sum P_n$, $\sum Q_n$ differ from $\sum p_n$, $\sum q_n$ only by zero terms, and are therefore divergent.

We shall construct sequences $\{m_n\}$, $\{k_n\}$, such that the series

$$\begin{aligned} &P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} \\ &\quad + P_{m_1+1} + \cdots + P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} + \cdots , \end{aligned} \tag{3.23}$$

which clearly is a rearrangement of $\sum a_n$, satisfies (3.22).

Choose real-valued sequence $\{\alpha_n\}, \{\beta_n\}$ such that $\alpha_n \to \alpha$, $\beta_n \to \beta$, $\alpha_n < \beta_n$, $\beta_1 > 0$.

Let m_1, k_1 be the smallest integers such that

$$\begin{aligned} &P_1 + \cdots + P_{m_1} > \beta_1, \\ &P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} < \alpha_1; \end{aligned}$$

let m_2, k_2 be the smallest integers such that

$$\begin{aligned} &P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} > \beta_2, \\ &P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} - Q_{k_1+1} - \cdots - Q_{k_2} < \alpha_2; \end{aligned}$$

and continues in this way. This is possible since $\sum P_n$ and $\sum Q_n$ diverge.

If x_n, y_n denote the partial sums of (3.23) whose last terms are P_{m_n} , $-Q_{k_n}$, then

$$|x_n - \beta_n| \leq P_{m_n}, \quad |y_n - \alpha_n| \leq Q_{k_n}.$$

Since $P_n \to 0$ and $Q_n \to 0$ as $n \to \infty$, we see that $x_n \to \beta, y_n \to \alpha$.

Finally, it is clearly that no number less than α or greater than β can be a subsequential limit of the partial sums of (3.23). □

Theorem 3.55. *If $\sum a_n$ is a series of complex numbers which converges absolutely, then every rearrangement of $\sum a_n$ converges, and they all converge to the same sum.*

证明. Let $\sum a'_n$ be a rearrangement, with partial sums $\sum s'_n$. Given $\varepsilon > 0$, there exists an integer N such that $m \geq n \geq N$ implies

$$\sum_{i=n}^{m} |a_i| \leq \varepsilon. \tag{3.24}$$

Now choose p such that the integers $1, 2, \ldots, N$ are all contained in the set $k_1, k_2, \ldots, k_p$ (we use the notation of Definition 3.52). Then if $n > p$, the numbers $a_1, a_2, \ldots, a_N$ will cancel in the difference $s_n - s'_n$, so that $|s_n - s'_n| \leq \varepsilon$, by (3.24). Hence $\{s'_n\}$ converges to the same sum as $\{s_n\}$. □

Chapter 4

Continuity

The function concept and some of the related terminology were introduced in Definitions 2.1 and 2.2. Although we shall (in later chapters) be mainly interested in real and complex functions (i.e., in functions whose values are real or complex numbers) we shall also discuss vector-valued functions (i.e., functions with values in $\mathbb{R}^k$) and functions with values in an arbitrary metric space. The theorems we shall discuss in this general setting would not become any easier if we restricted ourselves to real functions, for instance, and it actually simplifies and clarifies the picture to discard unnecessary hypotheses and to state and prove theorems in an appropriately general context.

The domains of definition of our functions will also be metric spaces, suitably specialized in various instances.

4.1 Limits of functions

Definition 4.1. Let X and Y be metric spaces; suppose $E \subset X$, f maps E into Y, and p is a limit point of E. We write $f(x) \to q$ as $x \to p$, or

$$\lim_{x \to p} f(x) = q \tag{4.1}$$

if there is a point $q \in Y$ with the following property: For every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), q) < \varepsilon \tag{4.2}$$

for all points $x \in E$ for which

$$0 < d_X(x, p) < \delta. \tag{4.3}$$

The symbols d_X and d_Y refer to the distances in X and Y, respectively.

If X and/or Y are replaced by the real line, the complex plane, or by some euclidean space $\mathbb{R}^k$, the distances d_X, d_Y are of course replaced by absolute values, or by norms of differences (see Sec. 2.16).

It should be noted that $p \in X$, but that p need not be a point of E in the above definition. Moreover, even if $p \in E$, we may very well have $f(p) \neq \lim_{x \to p} f(x)$ □ .

We can recast this definition in terms of limits of sequences:

Theorem 4.2. *Let X, Y, E, f , and p be as in Definition 4.1. Then*

$$\lim_{x \to p} f(x) = q \tag{4.4}$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q \tag{4.5}$$

for every sequence $\{p_n\}$ in E such that

$$p_n \neq p, \quad \lim_{n \to \infty} p_n = p. \tag{4.6}$$

证明. Suppose (4.4) holds. Choose $\{p_n\}$ in E satisfying (4.6). Let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that $d_Y(f(x), q) < \varepsilon$ if $x \in E$ and $0 < d_X(x, p) < \delta$. Also, there exists N such that $n > N$ implies $0 < d_X(p_n, p) < \delta$. Thus, for $n > N$, we have $d_Y(f(p_n), q) < \delta$, which shows that (4.5) holds.

Conversely, suppose (4.4) is false. Then there exists some $\varepsilon > 0$ such that for every $\delta > 0$ there exists a point $x \in E$ (depending on δ), for which $d_Y(f(x), q) \geq \varepsilon$ but $0 < d_X(x, p) < \delta$. Taking $\delta_n = 1/n (n = 1, 2, 3, ...)$, we thus find a sequence in E satisfying (4.6) for which (4.5) is false. □

Corollary. If f has a limit at p, this limit is unique.

Definition 4.3. Suppose we have two complex functions, f and g, both defined on E. By $f + g$ we mean the function which assigns to each point x of E the number $f(x) + g(x)$. Similarly we define the difference $f - g$, the product fg, and the quotient f/g of the two functions, with the understanding that the quotient is defined only at those points x of E at which $g(x) \neq 0$. If f assigns to each point x of E the same number c, then f is said to be a constant function, or simply a constant, and we write $f = c$. If f and g are real functions, and if $f(x) \geq g(x)$ for every $x \in E$, we shall sometimes write $f \geq g$, for brevity.

Similarly, if $\mathbf{f}$ and $\mathbf{g}$ map E into $\mathbb{R}^k$, we define $\mathbf{f} + \mathbf{g}$ and $\mathbf{f} \cdot \mathbf{g}$ by

$$(\mathbf{f} + \mathbf{g})(x) = \mathbf{f}(x) + \mathbf{g}(x), \quad (\mathbf{f} \cdot \mathbf{g})(x) = \mathbf{f}(x) \cdot \mathbf{g}(x);$$

and if λ is a real number, $(\lambda \mathbf{f})(x) = \lambda \mathbf{f}(x)$.

Theorem 4.4. *Suppose $E \subset X$, a metric space, p is a limit point of E, f and g are complex functions on E, and*

$$\lim_{x \to p} f(x) = A, \quad \lim_{x \to p} g(x) = B.$$

Then

(a) $\lim_{x \to p} (f + g)(x) = A + B$;

(b) $\lim_{x\to p}(fg)(x) = AB$*;*
(b) $\lim_{x\to p}(\frac{f}{g})(x) = \frac{A}{B}$*, if* $B \neq 0$*.*

证明. In view of Theorem 4.2, these assertions follow immediately from the analogous properties of sequences (Theorem 3.3). □

Remark 4.5. If f and g map E into $\mathbb{R}^k$, then (a) remains true, and (b) becomes
(b') $\lim_{x\to p}(\mathbf{f}\cdot\mathbf{g})(x) = \mathbf{A}\cdot\mathbf{B}$;

(Compare Theorem 3.4.)

4.2 Continuous functions

Definition 4.6. Suppose X and Y are metric spaces, $E \subset X$, $p \in E$, and f maps E into Y. Then f is said to be *continuous at* p if for every $\varepsilon > 0$ there exists a $\varepsilon > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon$$

for all points $x \in E$ for which $d_X(x, p) < \delta$.

If f is continuous at every point of E, then f is said to be *continuous on* E.

It should be noted that f has to be defined at the point p in order to be continuous at p. (Compare this with the remark following Definition 4.1.)

If p is an isolated point of E, then our definition implies that every function f which has E as its domain of definition is continuous at p. For, no matter which $\varepsilon > 0$ we choose, we can pick $\delta > 0$ so that the only point $x \in E$ for which $d_X(x, p) < \delta$ is $x = p$; then

$$d_Y(f(x), f(p)) = 0 < \varepsilon.$$

Theorem 4.7. *In the situation given in Definition 4.6, assume also that* p *is a limit point of* E*. Then* f *is continuous at* p *if and only if* $\lim_{x\to p} f(x) = f(p)$*.*

证明. This is clear if we compare Definitions 4.1 and 4.6. □

We now turn to compositions of functions. A brief statement of the following theorem is that a continuous function of a continuous function is continuous.

Theorem 4.8. *Suppose* X, Y, Z *are metric spaces,* $E \subset X$*,* f *maps* E *into* Y*,* g *maps the range of* f*,* $f(E)$*, into* Z*, and* h *is the mapping of* E *into* Z *defined by*

$$h(x) = g(f(x)) \quad (x \in E).$$

If f *is continuous at a point* $p \in E$ *and if* g *is continuous at the point* $f(p)$*, then* h *is continuous at* p*.*

This function his called the composition or the composite off and g. The notation

$$h = g \circ f$$

is frequently used in this context.

证明. Let $\varepsilon > 0$ be given. Since g is continuous at $f(p)$, there exists $\eta > 0$ such that

$$d_Z(g(y), g(f(p))) < \varepsilon \text{ if } d_Y(y, f(p)) < \eta, \text{ and } y \in f(E).$$

Since f is continuous at p, there exists $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \eta, \text{ if } d_X(x, p) < \delta \text{ and } x \in E.$$

It follows that

$$d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \varepsilon$$

if $d_X(x, p) < \delta$ and $x \in E$. Thus h is continuous at p. □

Theorem 4.9. *A mapping f of a metric space X into a metric space Y is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y.*

(Inverse images are defined in Definition 2.2.) This is a very useful characterization of continuity.

Corollary. ?? A mapping f of a metric space X into a metric space Y is continuous if and only if $f^{-1}(C)$ is closed in X for every closed set C in Y.

This follows from the theorem, since a set is closed if and only if its complement is open, and since $f^{-1}(E^c) = [f^{-1}(E)]^c$ for every $E \subset Y$.

We now turn to complex-valued and vector-valued functions, and to functions defined on subsets of $\mathbb{R}^k$.

Theorem 4.10. *Let f and g be complex continuous functions on a metric space X. Then $f+g$, fg, and f/g are continuous on X.*

In the last case, we must of course assume that $g(x) \neq 0$, for all $x \in X$.

证明. At isolated points of X there is nothing to prove. At limit points, the statement follows from Theorems 4.4 and 4.7. □

Theorem 4.11. *(a) Let $f_1, \ldots, f_k$ be real functions on a metric space X, and let $\mathbf{f}$ be the mapping of X into $\mathbb{R}^k$ defined by*

$$\mathbf{f}(x) = (f_1(x), ..., f_k(x)) \quad (x \in X); \tag{4.7}$$

then $\mathbf{f}$ is continuous if and only if each of the functions $f_1, ..., f_k$ is continuous.

(b) If $\mathbf{f}$ and $\mathbf{g}$ are continuous mappings of X into $\mathbb{R}^k$, then $\mathbf{f} + \mathbf{g}$ and $\mathbf{f} \cdot \mathbf{g}$ are continuous on X.

The functions $f_1, ..., f_k$ are called the *components* of $\mathbf{f}$. Note that $\mathbf{f}+\mathbf{g}$ is a mapping into $\mathbb{R}^k$, whereas $\mathbf{f}\cdot\mathbf{g}$ is a *real* function on X.

证明. Part (a) follows from the inequalities

$$|f_j(x) - f_j(x)| \leq |\mathbf{f}(x) - \mathbf{f}(y)| = \left\{\sum_{i=1}^{k} |f_i(x) - f_i(y)|^2\right\}^{\frac{1}{2}},$$

for $j = 1, 2, ..., k$. Part (b) follows form (a) and Theorem 4.10. □

Example 4.12.

$$\phi_i(\mathbf{x}) = x_i \quad (\mathbf{x} \in \mathbb{R}^k) \tag{4.8}$$

$$x_1^{n_1} x_2^{n_2} \dots x_k^{n_k} \tag{4.9}$$

$$P(\mathbf{x}) = \sum c_{n_1 \dots n_k} x_1^{n_1} x_2^{n_2} \dots x_k^{n_k} \quad (\mathbf{x} \in \mathbb{R}^k) \tag{4.10}$$

$$||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}| \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^k) \tag{4.11}$$

Remark 4.13. We defined the notion of continuity for functions defined on a subset E of a metric space X. However, the complement of E in X plays no role whatever in this definition (note that the situation was somewhat different for limits of functions). Accordingly, we lose nothing of interest by discarding the complement of the domain off This means that we may just as well talk only about continuous mappings of one metric space into another, rather than of mappings of subsets. This simplifies statements and proofs of some theorems. We have already made use of this principle in Theorems 4.9 to 4.11, and will continue to do so in the following section on compactness.