Supplementary Notes for
W. Rudin: Principles of Mathematical Analysis

SIGURDUR HELGASON

In 18.100B it is customary to cover Chapters 1–7 in Rudin's book. Experience shows that this requires careful planning especially since Chapter 2 is quite condensed. These notes include solutions of Exercises 23–26, Chapter 2 because these help in understanding the abstract compactness notion in 2.32, and makes it more useful in analysis.

Since the course is preceded by 18.034 which deals with the techniques of differential equations without proofs it seems very desirable to go through Exercise 27 in Ch. 5 and Exercise 25, Ch. 7. This gives concrete applications of the general theory in the course, consolidating 18.034.

For reasons of time some omissions seem advisable. This includes the Appendix in Ch. 1, and the section on Rectifiable Curves in Ch. 6. Also taking $\alpha(x) = x$ throughout Chapter 6 seems practical. The notions are then familiar and quite a bit of time is saved.

The book has a number of exercises of considerable theoretical interest. However, they are sometimes too difficult to assign to students and as a result they are often omitted altogether. Occasional unproved statements in the text may also baffle the beginner. It is the purpose of these notes to remedy this situation to some extent. Some of this material, like Lusin's Example, related to 3.44, is of course optional.

Chapter 1

On p. 11 we define for each $x \in R$, $x > 0$ the set

$$E = \{n_0 + \tfrac{n_1}{10} + \cdots + \tfrac{n_k}{10^k} : \; k = 0, 1, 2 \ldots\},$$

where for each k, n_k is the largest integer such that

$$n_0 + \tfrac{n_1}{10} + \cdots + \tfrac{n_k}{10^k} \leq x\,.$$

Proposition. $x = \sup E$.

(Proof not given on p.11): We must prove that if $y < x$ then y is not an upper bound of E. Let

$$r_k = x - \left[n_0 + \tfrac{n_1}{10} + \cdots + \tfrac{n_k}{10^k}\right]\,.$$

Then $r_k \geq 0$. We shall prove by induction on k that

$$r_k < \tfrac{1}{10^k}\,. \tag{1}$$

This is clear if $k = 0$. Assume (1) for k. If $n_{k+1} < 9$ then $r_{k+1} < \frac{1}{10^{k+1}}$ (proving (1)) because otherwise we could replace n_{k+1} by $n_{k+1} + 1$. If $n_{k+1} = 9$ then

$$r_{k+1} = r_k - \tfrac{9}{10^{k+1}} < \tfrac{1}{10^k} - \tfrac{9}{10^{k+1}} = \tfrac{1}{10^{k+1}}$$

proving (1) in this case too. Now choose $p \in Q$ such that $0 < p < x - y$ and then an integer $k > 0$ such that $10^{-k} < p < x - y$. Then $y < x - 10^{-k} < x - r_k$. Since $x - r_k \in E$, y is not an upper bound of E.

The following exercise is an analog of Exercise 5.

Exercise 5a.

Let A be a bounded set of real number $x > 0$ and $A^{-1} = \{x^{-1} : x = A\}$. Then

$$\sup A = 1/\inf(A^{-1})\,. \tag{2}$$

Proof: The upper bounds of A are just the reciprocals of the lower bounds of A^{-1}. The least of the former are thus the reciprocals of the largest of the latter. Thus (2) holds.

Exercise 6. Solution

(a) If $m > 0$ then $p > 0$ and

$$(b^m)^q = (b^p)^n = b^{mq}\,. \tag{3}$$

If $m < 0$ then $p < 0$ and $b^m = \left(\frac{1}{b}\right)^{-m}$ so by (3)

$$(b^m)^q = \left(\left(\tfrac{1}{b}\right)^{-m}\right)^q = \left(\tfrac{1}{b}\right)^{-mq} = b^{mq}\,.$$

Similarly $(b^p)^n = b^{np}$. Thus $(b^m)^{\frac{1}{n}}$ and $(b^p)^{\frac{1}{q}}$ have the same nq^{th} power so must coincide.

(b) By definition if $p, q \in J$, $\alpha^{\frac{p}{q}} = (\alpha^p)^{\frac{1}{q}}$ so

$$\left(\alpha^{\frac{p}{q}}\right)^q = \alpha^p\,. \tag{4}$$

Let $r = m/n$, $s = p/q$ and raise the numbers $b^{\frac{m}{n}+\frac{1}{q}} = b^{\frac{mq+np}{nq}}$ and $b^{\frac{m}{n}}b^{\frac{p}{q}}$ to the nq^{th} power. Using (4) the result is the same in both cases. By the uniqueness in Theorem 1.21,

$$b^{r+s} = b^r b^s\,. \tag{5}$$

(c) Since $b > 1$ we have $b^{\frac{m}{n}} > 1$ for $m, n \in J$ because $b^{\frac{m}{n}} \leq 1$ leads to $b^m \leq 1$. Thus if $r \in Q$, $t \in Q\,, t \leq r$ we have by (5)

$$b^r = b^t b^{r-t} \geq b^t \tag{6}$$

so $\sup B(r) = b^r$. Hence we define for $x \in R$

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

We shall prove

$$\sup_{t \leq x, t \in Q} b^t = \inf_{r \geq x, r \in Q} b^r\,. \tag{7}$$

We have $\le$ by (6). Let L be the difference between the right and left side in (7). Suppose $L > 0$. Select $n \in J$ so large that

$$b^x\left(b^{\frac{1}{n}} - 1\right) < L$$

(see p. 58). Then select $r, t \in Q$ such that

$$x < r < \tfrac{1}{2n} + x\,, \quad x - \tfrac{1}{2n} < t < x\,.$$

Then $r - t < \frac{1}{n}$ so

$$b^r - b^t = b^t(b^{r-t} - 1) \le b^x(b^{r-t} - 1) < L\,,$$

which is a contradiction, proving (7).

(d) We have

$$\sup B(x+y) = \sup\{b^t : t \in Q,\ t \le x+y\}\,.$$

The last set contains $\{b^{r+s} : r \le x, s \le y, r, s \in Q\}$ so by (5)

$$\sup B(x+y) \ge \sup\{b^r b^s : r \le x, s \le y\}\,,$$

and since r and s are independent this equals

$$\sup\{b^r : r \le x\} \sup\{b^s : s \le y\}\,.$$

Hence

$$b^{x+y} \ge b^x b^y\,. \tag{8}$$

By Exercise 5a above and (7) (s, t being in Q)

$$b^{-x} = \sup_{t \le x} b^t = \frac{1}{\inf_{t \le -x} b^{-t}} = \frac{1}{\inf_{s \ge x} b^s} = \frac{1}{b^x}\,. \tag{9}$$

If $b^x b^y$ were $< b^{x+y}$ then by (8) (9)

$$b^y = b^{-x}(b^x b^y) < b^{-x}(b^{x+y}) \le b^y\,.$$

This contradiction proves $b^{x+y} = b^x b^y$.

Chapter 2

In connection with Theorem 2.22 the following result is useful. The proof is quite similar.

Theorem 2.22a *If $E \subset X$ and $E_\alpha \subset X$ for each α, then*

$$E \cap (\cup_\alpha E_\alpha) = \cup_\alpha (E \cap E_\alpha), \quad E \cup (\cap_\alpha E_\alpha) = \cap_\alpha (E \cup E_\alpha)\,.$$

If $f : X \to Y$ is a mapping then

$$f(\cup_\alpha E_\alpha) = \cup_\alpha f(E_\alpha), \quad f(\cap_\alpha E_\alpha) \subset \cap_\alpha f(E_\alpha)\,.$$

If $F \subset Y$, $F_\beta \subset Y$ for each β then

$$f^{-1}(\cup_\beta F_\beta) = \cup_\beta f^{-1}(F_\beta), \quad f^{-1}(\cap_\beta F_\beta) = \cap_\beta f^{-1}(F_\beta)$$

$$f^{-1}(F^c) = (f^{-1}(F))^c, \quad f(f^{-1}(F)) \subset F, \quad E \subset f^{-1}(f(E))\,.$$

Remark There are examples such that

(i) $f(A \cap B) \neq f(A) \cap f(B)$,

(ii) $f(E^c) \neq (f(E))^c$,

(iii) $f(f^{-1}(F)) \neq F$,

(iv) $f^{-1}(f(E)) \neq E$.

For (i) take a case when $A \cap B = \emptyset$. For (ii) take a case when f maps X onto Y. For (iii) take a case when $F = Y$ and f a constant map. For (iv) take a case when E is a single point and f a constant map.

Notation If $A \subset X$, $B \subset X$, we write $A - B = \{x \in A : \ x \notin B\}$.
Theorem 2.27 (Improved version.) *Let X be a metric space and $E \subset X$. Then*

(a) $p \in \overline{E}$ *if and only if each neighborhood N of p intersects E (i.e., $N \cap E \neq \emptyset$).*

(b) $\overline{E}$ *is closed.*

(c) $E = \overline{E}$ *if and only if E is closed.*

(d) $\overline{E} \subset F$ *for every closed set $F \subset X$ such that $E \subset F$.*

(e) *Anticipating Definition 3.1 we have*

$$p \in \overline{E} \text{ if and only if there exists a sequence } (p_n) \subset E \text{ such that } p_n \to p\,.$$

Proof

(a) If $p \in \overline{E}$ then either $p \in E$ or p is a limit point of E. In both cases each neighborhood N of p intersects E. On the other hand suppose $p \notin \overline{E}$. Then $p \notin E'$. Thus there exists a neighborhood N of p such that $N \cap (E - p) = \emptyset$. But $p \notin E$ so $N \cap E = \emptyset$ proving a).

(b) If $p \notin \overline{E}$ then by a) there exists a neighborhood N of p such that $N \cap E = \emptyset$. We claim that $N \cap \overline{E} = \emptyset$ (which would show $X - \overline{E}$ open). Otherwise there is a $q \in \overline{E} \cap N$. But then N is an open set containing q so by a) $N \cap E \neq \emptyset$ which is a contradiction.

(c) If $E = \overline{E}$ then by b), E is closed. If E is closed then by definition $E' \subset E$ so $\overline{E} \subset E$.

(d) If F is closed and $F \supset E$, then $F \supset F'$ so $F \supset E'$ whence $F \supset \overline{E}$.

(e) Let $p \in \overline{E}$. Using a) we can for each $n \in J$ pick a point $p_n \in N_{1/n}(p) \cap E$. Then $d(p_n, p) < \frac{1}{n}$ so $p_n \to p$. On the other hand if (p_n) is a sequence in E such that $p_n \to p$ then $p \in \overline{E}$ by a).

Theorem 2.37 (Improved version.) *A space X is compact if and only if each infinite subset has a limit point.*
The "only if" part is proved in the text. The "if" part is Exercise 26 which we shall prove along with Exercises 23, 24.

Exercise 23(Stronger version.)

Let X be a metric space. Then X is separable if and only if X has a countable base.

Proof: First assume X has a countable dense subset $P = \{p_1, p_2, \ldots\}$. Consider the countable collection $\mathcal{N}$ of neighborhoods $N_r(p_i)$, $r \in Q$, $i \in J$. Let G be open and $x \in G$. Then $N_\rho(x) \subset G$ for some $\rho \in Q$. Since P is dense in X, i.e., $\overline{P}$=X, $P \cap N_{\rho/2}(x)$ contains some p_i. Then $d(p_i, x) < \rho/2$ so $x \in N_{\rho/2}(p_i)$. If $y \in N_{\rho/2}(p_i)$ then $d(x,y) \leq d(x,p_i) + d(p_i,y) < \rho$ so $y \in G$. Thus $x \in N_{\rho/2}(p_i) \subset G$ so G is the union of some members of $\mathcal{N}$. Thus $\mathcal{N}$ is a countable base.

Conversely, suppose $\{V_i\}_{i\in J}$ is any countable base for X. Pick $x_i \in V_i$ for each i and let $F = \{x_1, x_2, \ldots\}$. Then $\overline{F} = X$ because otherwise the open set $X - \overline{F}$ would be the union of some of the V_i contradicting $x_i \in F$.

Exercise 24.

Let X be a metric space in which each infinite subset has a limit point. Then X is separable.

Proof: Let $\delta > 0$ and let $x_1 \in X$. Pick $x_2 \in X$ (if possible) such that $d(x_1, x_2) \geq \delta$. Having chosen $x_1, \ldots, x_j$ choose x_{j+1} (if possible) such that $d(x_i, x_{j+1}) \geq \delta$ for $1 \leq i \leq j$. This process has to stop because otherwise the set $E = (x_1, x_2, \ldots)$ would be an infinite subset of X without limit points. Consequently X is covered by finitely many sets $N_\delta(x_i)$. Now take $\delta = \frac{1}{n}$ $(n = 1, 2, \ldots)$. The construction gives:

$$n = 1, \quad x_{11}, \ldots, x_{1j_1} \quad X = \bigcup_{i=1}^{j_1} N_1(x_{1i})$$

$$n = 2, \quad x_{21}, \ldots, x_{2j_2} \quad X = \bigcup_{i=1}^{j_2} N_{\frac{1}{2}}(x_{2i})\,,$$

etc. Let S consist of these centers x_{ki} $k \in J$, $1 \leq i \leq j_k$. Then S is dense in X (i.e., $\overline{S} = X$). In fact let $x \in X$ and $N_r(x)$ any neighborhood of x. Choose k such that $\frac{1}{k} < r$ and x_{ki} such that $x \in N_{\frac{1}{k}}(x_{ki})$. Then $x_{ki} \in N_{\frac{1}{k}}(x) \subset N_r(x)$ so $N_r(x) \cap S \neq \emptyset$ as claimed.

Exercise 26.

Let X be a metric space in which every infinite subset has a limit point. Then X is compact.

Proof:

By Exercises 23 and 24, X has a countable base $\{B_i\}_{i\in J}$. Let $\{G_\alpha\}_{\alpha\in A}$ be any open covering of X. If $x \in X$ then $x \in G_\alpha$ for some α and thus there is a B_i such that $x \in B_i \subset G_\alpha$. Let $\{B_i\}_{i\in I}$ be the sub-collection of $\{B_i\}_{i\in J}$ consisting of those B_i which are contained in some G_α. For each $B_i (i \in I)$ let G_i be one of the G_α containing B_i. Then

$$X = \bigcup_{i\in I} G_i\,. \tag{10}$$

Thus we have obtained a countable subcovering of the $\{G_\alpha\}_{\alpha\in A}$. We shall now show that a suitable finite collection of the G_i covers X, thus proving the compactness of X. Let $I = \{i_1, i_2, \ldots\}$. If no finite sub-collection of $\{G_i\}_{i\in I}$ covers X then for each n

$$F_n = (G_{i_1} \cup \cdots \cup G_{i_n})^c \neq \emptyset\,,$$

whereas by (10)

$$\bigcap_1^\infty F_n = \emptyset\,. \tag{11}$$

Of course we have

$$F_1 \supset F_2 \supset \cdots \supset F_n \supset \cdots . \tag{12}$$

We select $x_i \in F_i$ for each i and let $E = \{x_1, x_2, \ldots\}$. For each $x \in E$ let N_x denote the set of indices p for which $x \in F_p$. Then

$$\bigcup_{x \in E} N_x = J .$$

If E were finite some N_x would be infinite and x would belong to F_p for infinitely many p whence by (12) x would belong to F_p for all p. This would contradict (11) so we conclude that E is an infinite subset of X. By assumption it has a limit point z. Let N be any neighborhood of z. Then $N \cap (E - z) \neq \emptyset$ so $N \cap E \neq \emptyset$. Thus N intersects some F_n so by (12) $N \cap F_1 \neq \emptyset$. Since N was arbitrary we have $z \in \overline{F}_1$. But F_1 is closed so $z \in F_1$. Now consider the sequence

$$F_2 \supset F_3 \supset \cdots \supset F_n \supset \cdots$$

(with F_1 removed) and $E_1 = E - x_1 = \{x_2, x_3, \ldots\}$. The point z is still a limit point of E_1 so by the argument above $z \in F_2$. Continuing we deduce $z \in \cap_1^\infty F_i$ contradicting (11). This proves that X is compact.

Corollary. *Let X be a metric space. The following conditions are equivalent:*

(i) X is compact.

(ii) Each sequence (x_n) in X has a convergent subsequence.

(iii) Each infinite subset of X has a limit point.

Proof: (i) $\Rightarrow$ (ii): This is Theorem 3.6(a). (ii) $\Rightarrow$ (iii): Let $E \subset X$ be an infinite subset. Let (x_n) be a sequence of *different* points in E. By (ii) it has a subsequence (x_{n_i}) converging to a limit x. Then each $N_r(x)$ contains infinitely many x_{n_i}. Since $x_{n_i} = x$ for at most one i this implies $N_r(x) \cap (E - x) \neq \emptyset$ so x is a limit point of E. Part (iii) $\Rightarrow$ (i) was Exercise 26.

Chapter 3

Exercise.

Every sequence (x_n) in $\mathbf{R}$ has a monotone subsequence.

Proof: This is clear if (x_n) is unbounded. If x_n is bounded it has a subsequence converging to say x. Considering $x_n - x$ we may assume $x = 0$. Considering signs we may assume all terms positive. A positive sequence converging to 0 clearly has a decreasing subsequence.

Exercise 3. Solution: We have $s_1 = \sqrt{2}$, $s_{n+1} = \sqrt{2 + \sqrt{s_n}}$.
(s_n) is increasing by induction: If $s_{n+1} - s_n \geq 0$ then

$$s_{n+2}^2 - s_{n+1}^2 = (2 + \sqrt{s_{n+1}}) - (2 + \sqrt{s_n}) \geq 0 .$$

Also $s_n < 2$ by induction since $\sqrt{2 + \sqrt{2}} < 2$.

Exercise 16. Solution: We have $\alpha > 0$, $x_1 > \sqrt{\alpha}$ and

$$x_{n+1} = \tfrac{1}{2}\left(x_n + \tfrac{\alpha}{x_n}\right) \qquad n = 2, 3, \dots .$$

Show (x_n) is decreasing and that $x_n \to \sqrt{\alpha}$. For this note that

$$x_{n+1} = \tfrac{1}{2}\left(\sqrt{x_n} - \sqrt{\tfrac{\alpha}{x_n}}\right)^2 + \sqrt{\alpha}\,, \quad x_n - x_{n+1} = \tfrac{1}{2}\left(x_n - \tfrac{\alpha}{x_n}\right).$$

The first relation shows by induction that $x_n > \sqrt{\alpha}$ and them the second relation shows that (x_n) is decreasing. The definition then shows $x_n \to \sqrt{\alpha}$.

Exercise 17. Solution:
Here one proceeds similarly using the relations

$$\begin{aligned} x_{n+1} - \sqrt{\alpha} &= \frac{(\sqrt{\alpha}-1)(\sqrt{\alpha}-x_n)}{1+x_n} \\ x_{n+2} - x_n &= \frac{2}{1+\alpha+2x_n}(\alpha - x_n^2)\,. \end{aligned}$$

Lusin's Example (Compare Theorem 3.44.)

There exists a power series $\Sigma a_n z^n$ with $a_n \to 0$ and radius of convergence 1, diverging everywhere on the circle $|z| = 1$.

Proof: Let

$$\begin{aligned} g_m(z) &= 1 + z + \cdots z^{m-1} \\ h_m(z) &= g_m(z) + z^m g_m(e^{-2\pi i/m}z) + \cdots + z^{mk} g_m(e^{-2\pi ik/m}z) + \cdots z^{m(m-1)} g_m(e^{-2\pi i(m-1)/m}z)\,. \end{aligned}$$

Notice that when the individual terms in $h_m(z)$ are written out as polynomials in z there is no overlap. Thus $h_m(z)$ becomes a polynomial in z (of degree $(m+1)(m-1)$) in which each coefficient has absolute value 1. Also if $e^{i\varphi} \neq 1$

$$|g_m(e^{i\varphi})| = \frac{|1 - e^{mi\varphi}|}{|1 - e^{i\varphi}|} = \frac{|\sin\frac{m\varphi}{2}|}{|\sin\frac{\varphi}{2}|}\,.$$

The graph of $\sin x$ shows

$$\tfrac{2}{\pi}x \le \sin x \le x \qquad \text{for } 0 \le x \le \tfrac{\pi}{2}\,.$$

Thus

$$|g_m(e^{i\varphi})| \ge \frac{\frac{2}{\pi}|\frac{m\varphi}{2}|}{|\frac{\varphi}{2}|} = \tfrac{2}{\pi}m \quad \text{for } |\varphi| \le \tfrac{\pi}{m}\,.$$

Consider now the circular arc

$$A = \{e^{i\varphi} : |\varphi| \le \tfrac{\pi}{m}\}\,.$$

For each point $e^{i\varphi}$ on $|z| = 1$ we can find an integer k such that

$$e^{-\frac{2\pi ki}{m}} e^{i\varphi} \in A\,.$$

In fact the arcs $e^{\frac{2\pi ki}{m}}A$ $(0 \le k \le m)$ cover the circle. It follows that

$$\text{Max}_{\,0\le k\le m} \left|g_m\left(e^{-\frac{2\pi k}{m}i}z\right)\right| \ge \frac{2}{\pi}m \qquad \text{for } |z| = 1\,. \tag{13}$$

Consider now the series

$$\sum_{m=1}^{\infty} \tfrac{1}{\sqrt{m}} z^{1^2+2^2+\cdots+(m-1)^2} h_m(z) = \sum_n a_n z^n \, .$$

Again there is no overlap for different m because the m^{th} term contains powers of z of degree at most $1^2 + 2^2 + \cdots + (m-1)^2 + (m+1)(m-1)$ which is less than $1^2 + 2^2 + \cdots + m^2$ which is the lowest degree of z appearing in the $(m+1)^{\text{th}}$ term.

Now suppose the series did converge at a point z with $|z| = 1$. Then the sequence of terms $a_n z^n$ must converge to 0. This remains true for each subsequence. The absence of overlaps noted thus implies

$$\lim_{m\to\infty} \frac{1}{\sqrt{m}} z^{1^2+2^2+\cdots+(m-1)^2} \operatorname{Max}_{0\leq k\leq m} |z^{mk} g_m \left(e^{-\frac{2\pi ki}{m}} z\right) | = 0 \, .$$

This would contradict (13).

Corollary. *There exists a power series $\sum_n c_n z^n$ with radius of convergence 1 which converges for $z = 1$ but diverges everywhere else on the circle $|z| = 1$.*

Proof: With $\Sigma a_n z^n$ as in Lusin's example put

$$g(z) = a_0 - a_0 z + a_1 z^2 - a_1 z^3 + a_2 z^4 - a_2 z^5 + \cdots \, .$$

Clearly we have convergence for $z = 1$. If we had convergence at a point $z \neq 1$ then the series

$$a_0(1-z) + a_1 z^2 (1-z) + \cdots$$

would also converge and so would the series

$$a_0 + a_1 z^2 + a_2 z^4 + \cdots$$

contradicting the theorem.

Exercise 19. Solution (cf. Mattuck: Introduction to Analysis).

The Cantor set C is given as follows:

$$C_0 = [0,1], \; C_1 = C_0 - \left(\tfrac{1}{3}, \tfrac{2}{3}\right), \; C_2 = C_1 - \left(\tfrac{1}{9}, \tfrac{2}{9}\right) - \left(\tfrac{7}{9}, \tfrac{8}{9}\right), \text{ etc.}$$

and $C = \cap_n C_n$. Now represent each $x = C_0$ by infinite decimal expansion with base 3, i.e.,

$$x = \cdot\alpha_1\alpha_2\cdots = \tfrac{\alpha_1}{3} + \tfrac{\alpha_2}{3^2} + \cdots \, .$$

Here we avoid 1 as much as possible: we replace each ternary decimal ending with 1 by an equivalent ternary decimal ending with an infinite string of 2. Then

$$\alpha_1 = \begin{cases} 0 & \text{if} \quad 0 \leq x \leq \frac{1}{3} \\ 1 & \text{if} \quad \frac{1}{3} < x < \frac{2}{3} \\ 2 & \text{if} \quad \frac{2}{3} \leq x \leq 1 \end{cases} \, .$$

Thus $C_1 = \{x \in C_0 : \alpha_1 \neq 1\}$. Similarly α_2 is found by taking the interval of length $1/3$ containing x and dividing it into three subintervals of length $1/9$. Then $\alpha_2 = 0, 1$ or 2 according to which of these intervals contain x just as above. Thus

$$C_2 = \{x \in C_0 : \alpha_1 \neq 1, \alpha_2 \neq 1\}.$$

Continuing in this way we see that

$$C = \{x | x = \frac{\alpha_1}{3} + \frac{\alpha_2}{3^2} + \cdots + \frac{\alpha_n}{3^n} + \cdots, \quad \text{each } \alpha_i \neq 1\}.$$

Exercise 24. (Completion of a metric space.) We prove Part d) before Part c).

(a) Easy from the definition.

(b) If $\{p_n'\} \sim \{p_n\}$ and $\{q_n'\} \sim \{q_n\}$ then

$$d(p_m', q_n') \leq d(p_n', p_n) + d(p_n, q_n) + d(q_n, q_n')$$

and exchanging p_n' and p_n, q_n' and q_n

$$|d(p_n', q_n') - d(p_n, q_n)| \leq d(p_n', p_n) + d(q_n', q_n)$$

so $\Delta(P, Q)$ is well defined. Note that if $\Delta(P, Q) = 0$ then $\{p_n\} \sim \{q_n\}$ so $P = Q$. The inequality $\Delta(P, R) \leq \Delta(P, Q) + \Delta(Q, R)$ is easy.

(d) With the notation indicated we have

$$\Delta(P_p, P_q) = \lim_{n \to \infty} d(p, q) = d(p, q),$$

so the mapping $\varphi : p \to P_p$ from X into X^* is distance preserving. If $\{p_n\} \in P$ then by b)

$$\Delta(P, P_p) = \lim_{n \to \infty} d(p, p_n).$$

In particular,

$$\Delta(P_{p_m}, P) = \lim_{n \to \infty} d(p_m, p_n)$$

so since $\{p_n\}$ is Cauchy sequence,

$$\lim_{m \to \infty} \Delta(P_{p_m}, P) = 0. \tag{14}$$

Thus given any Cauchy sequence $\{p_n\}$ and $\epsilon > 0$ there is a *constant* sequence whose equivalence class is within ϵ of the equivalence class of $\{p_n\}$. This means that $\varphi(X)$ is dense in X^*.

(c) X^* is complete. Let $P_1, P_2, \ldots$ be a Cauchy sequence in X^*. Given $\epsilon > 0$ choose for each n (by (14)) a p_n such that $\Delta(P_{p_n}, P_n) < \epsilon$. We now prove that $P_{p_1}, P_{p_2}, \ldots$ is a Cauchy sequence in X^*. In fact

$$\begin{aligned} \Delta(P_{p_m}, P_{p_r}) &\leq \Delta(P_{p_m}, P_m) + \Delta(P_m, P_r) + \Delta(P_r, P_{p_r}) \\ &\leq \epsilon + \Delta(P_m, P_r) + \epsilon. \end{aligned}$$

The left hand side equals $d(p_m, p_r)$ so the inequality shows that (p_n) is a Cauchy sequence in X. Let P denote its equivalence class. Since

$$\Delta(P, P_m) \leq \Delta(P, P_{p_m}) + \Delta(P_{p_m}, P_m)$$

we deduce from (14) that $\Delta(P, P_m) \to 0$ proving c).

(e) If X was complete to begin with let $P \in X^*$ and select $\{p_n\} \in P$. Then $\{p_n\}$ is a Cauchy sequence in X so $p_n \to p$ for some $p \in X$. Then $\{p\} \sim \{p_n\}$, so $P_p = P$, i.e., $\varphi(p) = P$.

Chapter 4

Remark 4.3.1

Given any countable subset E of (a,b) there exists a function f monotonic on (a,b) discontinuous at each point of E and at no other point of (a,b). In fact if $\sum_n c_n$ is convergent and $c_n > 0$ then the text states that

$$f(x) = \sum_{x_n < x} c_n \,,$$

where E consists of the points $x_1, x_2, \ldots$, is such a function. We shall prove this in the following steps:

(a) f is increasing.

(b) f is continuous at each $x \in (a,b)$ outside E.

(c) $f(x) = f(x^-)$ for all x.

(d) $f(x_n^+) - f(x_n^-) = c_n$.

(a) This part is obvious.

(b) Here we separate two cases.

  (i) $x \notin E$ and x is not a limit point of E.

  (ii) $x \notin E$ and x is a limit point of E.

  In case (i) there is a neighborhood $N_\delta(x)$ such that $N_\delta(x) \cap E = \emptyset$. Then $f(y) = \sum_{x_n<y} c_n$ is constant for $y \in (x-\delta, x+\delta)$, hence continuous at x.

  In case (ii) let $\epsilon > 0$ and N a number such that $\sum_{N+1}^{\infty} c_n < \epsilon$. Let $\delta > 0$ be such that

$$0 < \delta < |x - x_n| \quad \text{for } n = 1, \ldots, N \,.$$

  Then if $x_k \in (x-\delta, x+\delta)$ we have $k > N$. Let $y \in (x-\delta, x+\delta)$. Then

$$f(y) = \sum_{x_n<y} c_n \,, \quad f(x) = \sum_{x_n<x} c_n \,.$$

  Since $x_k \in (x-\delta, x+\delta)$ implies $k > N$ the sums above differ by at most $\sum_{k>N} c_k$ which is $< \epsilon$. Thus

$$|f(x) - f(y)| < \epsilon \quad \text{if} \quad y \in (x-\delta, x+\delta)$$

  so f is continuous at x.

(c) Here the relation remains to be proved for $x = x_n \in E$. Again we consider the two cases

  (i) x_n is an isolated point of E.

  (ii) x_n is a limit point of E.

In case (i) choose $\delta > 0$ such that $N_\delta(x_n) \cap E = x_n$. Then if $x_n - \delta < y < x_n$ we have

$$f(x_n) = \sum_{x_k < x_n} c_k = \sum_{x_k < y} c_k = f(y)$$

so $f(x_n) = f(x_n^-)$.

In case (ii) let $\epsilon > 0$ be given and choose N such that $\sum_{k>N} c_k < \epsilon$ and then choose $\delta > 0$ such that

$$0 < \delta < |x_n - x_m| \text{ for } m = 1, 2, \ldots, N, \quad m \neq n\,.$$

Then if $x_k \in E \cap (x_n - \delta, x_n)$ we have $k > N$. Thus if $y \in (x_n - \delta, x_n)$

$$\begin{aligned} f(x_n) = \sum_{x_k < x_n} c_k &= \sum_{x_k < y} c_k + \sum_{y \le x_k < x_n} c_k \\ &\le f(y) + \sum_{k > N} c_k \le f(y) + \epsilon\,. \end{aligned}$$

Hence $0 \le f(x_n) - f(y) \le \epsilon$ so $f(x_n) = f(x_n^-)$ also in this case.

(d) Again we consider two separate cases:

(i) x_n isolated in E.

(ii) x_n limit point of E.

In case (i) select $\delta > 0$ such that $N_\delta(x_n) \cap E = \emptyset$. Then if $y \in (x_n, x_n + \delta)$

$$f(y) = \sum_{x_k < y} c_k = c_n + \sum_{x_k < x_n} c_k = c_n + f(x_n)$$

so

$$f(x_n^+) = c_n + f(x_n) = c_n + f(x_n^-)\,.$$

In the case (ii) let $\epsilon > 0$ be given and select N such that $\sum_{k>N} c_k < \epsilon$. Again select $\delta > 0$ such that

$$\delta < |x_n - x_m| \text{ for } m = 1, 2, \ldots, N\,, \quad m \neq n\,.$$

If $x_k \in E \cap (x_n, x_n + \delta)$ then $k > N$. Thus if $y \in (x_n, x_n + \delta)$

$$\begin{aligned} f(y) &= \sum_{x_k < y} c_k = \sum_{x_k < x_n} c_k + c_n + \sum_{x_n < x_k < y} c_k \\ &= f(x_n) + c_n + \eta \text{ where } |\eta| < \epsilon \end{aligned}$$

so again $f(x_n^+) = f(x_n) + c_n$.

Exercise 15. A continuous open mapping of R into R is monotonic.

Proof: First we show that f is one-to-one. If not suppose $x_1 < x_2$ and $f(x_1) = f(x_2)$. Let y and z be points in $[x_1, x_2]$ where f takes its maximum and minimum, respectively. If $y \in (x_1, x_2)$ then $f(y)$ is an end point of $f((x_1, x_2))$ contradicting the openness of f. Similarly $z \in (x_1, x_2)$ is impossible. Thus both y and z are end points of $[x_1, x_2]$ so since $f(x_1) = f(x_2)$, f is constant in $[x_1, x_2]$, contradicting the openness of f.

Thus f is one-to-one. Consider any $x_1 < x_2$ and assume $f(x_1) < f(x_2)$. We shall prove f increasing on $[x_1, x_2]$. If not, select $x_3 < x_4$ in the interval such that $f(x_3) > f(x_4)$. Put $y_i = f(x_i)$ $(1 \leq i \leq 4)$. If $y_3 \leq y_2$ then $y_4 < y_2$ so by Theorem 4.23 f takes the value y_3 on (x_4, x_2). This contradicts the one-to-one property. Thus $y_3 > y_2 > y_1$ so f takes the value y_2 on (x_1, x_3) contradicting $x_3 < x_2$.

The argument applies to $x_1 = -n$, $x_2 = n$ and shows f monotonic on $(-n, n)$. Hence f is monotonic on R.

Chapter 5

Exercise 22. (Generalized.)

Let X be a complete metric space and $f : X \to X$ a contractive map, i.e., there is a constant $k < 1$ such that

$$d(f(x), f(y)) \leq k\, d(x, y) \text{ for all } x, y \in X\,.$$

Then f has a fixed point.

Proof (Richard Palais): By the triangle inequality

$$d(x, y) \quad \leq \quad d(x, f(x)) + d(f(x), f(y)) + d(f(y), y)$$

so

$$d(x, y) \quad \leq \quad \frac{1}{1-k}[d(x, f(x)) + d(f(y), y)]\,.$$

Replace here x by the iterate $f^{(n)}(x)$, y by $f^{(m)}(x)$. Then

$$\begin{aligned} d(f^{(n)}(x), f^{(m)}(x)) \quad &\leq \quad \frac{1}{1-k}[d(f^{(n)}(x), f^{(n+1)}(x)) + d(f^{(m)}(x), f^{(m+1)}(x))] \\ &\leq \quad \frac{1}{1-k}(k^n + k^m)\, d(x, f(x))\,. \end{aligned}$$

Thus for each $x \in X$ $f^{(n)}(x)$ is a Cauchy sequence and its limit is a fixed point for f.

Exercise 25.

b) Here one has only $x_{n+1} \leq x_n$ (not necessarily $x_{n+1} < x_n$ as the text states). In fact the case $f(x) = x$ shows this.

In b) one can use the fact that $f''(x) \geq 0$ implies that the curve always lies above its tangent: In fact by the mean value theorem there exists some ξ between x and x_1 and some η between ξ and x_1 such that

$$\begin{aligned} f(x) - [f(x_1) + f'(x_1)(x - x_1)] \quad &= \quad (f'(\xi) - f'(x_1))(x - x_1) \\ &= \quad f''(\eta)(\xi - x_1)(x - x_1) \geq 0\,. \end{aligned}$$

Thus $f(x_2) \geq 0$ and by induction $f(x_n) \geq 0$.

Chapter 7

Remark on (38) p.154. The sign determination in (38) is made as follows: Consider the intervals

$$\begin{aligned} I_1: &\quad 4^m x - 4^m|\delta_m| \le t < 4^m x \\ I_2: &\quad 4^m x \le t < 4^m x + 4^m|\delta_m| \end{aligned}$$

whose (disjoint) union is the half-open interval

$$I: \quad 4^m x - 4^m|\delta_m| \le t < 4^m x + 4^m|\delta_m|$$

which has length 1. Exactly one of the intervals I_1, I_2 contains an integer; if I_1 does we take $\delta_m > 0$; if I_2 does we take $d_m < 0$. In the first case the interval

$$[4^m x, 4^m x + 4^m \delta_m) \text{ contains no integer}$$

and in the second case the interval

$$[4^m x + 4^m \delta_m,\ 4^m x) \text{ contains no integer.}$$

In both cases φ is linear on the interval in question and this results in the crucial fact that $|\gamma_m| = 4^m$.

Exercise 4. $f(x) = \sum_1^\infty \frac{1}{1+n^2x}$

(a) If $x = 0$ series diverges.

(b) If $x = -\frac{1}{k^2}$ for some $k \in J$ some term is undefined so we have divergence.

(c) If $-\frac{1}{k^2} < x < -\frac{1}{(k+1)^2}$ for some $k = 1, 2, 3$ we write

$$f(x) = \sum_{n=1}^{k+1} \frac{1}{1+n^2x} + \frac{1}{x}\sum_{k+2}^{\infty} \frac{1}{\frac{1}{x}+n^2}.$$

Since

$$-(k+1)^2 < \tfrac{1}{x} < -k^2 \tag{15}$$

we have

$$|n^2 + \frac{1}{x}| \ge n^2 - (k+1)^2$$

and $\sum_{k+2}^\infty \frac{1}{n^2-(k+1)^2}$ converges. By Theorem 7.10 we have uniform convergence on each interval (15).

(d) If $-\infty < x < -1$ we write

$$f(x) = \frac{1}{1+x} + \frac{1}{x}\sum_{2}^{\infty} \frac{1}{n^2+\frac{1}{x}}. \tag{16}$$

Again

$$|n^2 + \tfrac{1}{x}| \ge n^2 - |\tfrac{1}{x}| \ge n^2 - 1$$

so by Theorem 7.10 we have uniform convergence on $-\infty < x < -1$.

(e) Writing f in the form (16) it is clear that the series

$$\sum_2^\infty \frac{1}{n^2 + \frac{1}{x}}$$

converges uniformly on $(0, \infty)$ since $n^2 + \frac{1}{x} > n^2$. But what about the series

$$\sum_1^\infty \frac{1}{1 + n^2 x} \,? \tag{17}$$

For this put

$$f_n(x) = \sum_{k=1}^{n} \tfrac{1}{1+k^2 x} \,.$$

If we had uniform convergence of (17) on $(0, \infty)$ then by Theorem 7.11 the sequence (A_n) given by

$$A_n = \lim_{x \to 0} f_n(x) \tag{18}$$

would converge. However $A_n = n$ so this is not the case. Thus (17) converges uniformly on each interval $[\alpha, \infty)$, $\alpha > 0$ but not on $(0, \infty)$.

(f) Since the proofs above always involved absolute values we have absolute convergence except for $x = -k^{-2}$, $(k = 1, 2, \ldots)$ and $x = 0$. The function f is continuous except at these points.

(g) Is $f(x)$ bounded? No, if $|f(x)| < K$ for all x then for $x > 0$

$$f_n(x) \leq K \text{ for each } n$$

so by (18) $A_n \leq K$ for each n which is false.

Exercise 14. To see that $f(3^k t_0) = a_k$ we write

$$t_0 = \sum_{i=0}^{k-1} \frac{2a_i}{3^{i+1}} + \frac{2a_k}{3^{k+1}} + \sum_{k+1}^{\infty} \frac{2a_i}{3^{i+1}} \,.$$

Then

$$3^k t_0 = 2p + \tfrac{2a_k}{3} + z$$

where p is an integer and $0 \leq z \leq \frac{1}{3}$. If $a_k = 0$ then $f(3^k t_0) = 0 = a_k$ and if $a_k = 1$ then $f(3^k t_0) = 1 = a_k$. Referring now to Exercise 19 in Chapter 3 we see that Φ does indeed map the Cantor set onto the unit square.

Exercise 25 (Remark on Part b.)

In general if $\mathcal{F}$ is a family of functions whose derivatives are uniformly bounded then by the mean-value theorem (5.10) the family $\mathcal{F}$ is equicontinuous. This does not quite work here since f_n is not differentiable at the points $x_i = i/n$.

However, we can proceed as follows: Let $0 \le x < y \le 1$. Select the smallest i for which $x \le x_i$ and the largest k such that $x_{i+k} \le y$. (If $y \le x_i$ the relation $|f_n(x) - f_n(y)| \le M|y-x|$ is obvious.) Then

$$f_n(x) - f_n(y) = f_n(x) - f_n(x_i) + f_n(x_i) - f_n(x_{i+1}) + \cdots + f_n(x_{i+k}) - f_n(y)\,.$$

We can use the mean-value theorem on each difference on the right giving an expression

$$f_n'(\xi_i)(x - x_i) + f_n'(\xi_{i+1})(x_i - x_{i+1}) + \cdots + f_n'(\xi_{i+k+1})(x_{i+k} - y)$$

which is bounded by $M|x-y|$.