HOMEWORK 3

Sequences, Series, and Continuity

MATH S4061 | Introduction to Modern Analysis I

Rudin Chapters 3 and 4 — Sequences, Series, and Continuity

Problems in This Set

Rudin Chapter 3 – Numerical Sequences and Series

Problem 3.1 Rudin Ch. 3, Ex. 1

Problem 3.2 Rudin Ch. 3, Ex. 2

Problem 3.3 Rudin Ch. 3, Ex. 3

Problem 3.4 Rudin Ch. 3, Ex. 4

Problem 3.5 Rudin Ch. 3, Ex. 5

Problem 3.6 Rudin Ch. 3, Ex. 6

Problem 3.7 Rudin Ch. 3, Ex. 7

Problem 3.8 Rudin Ch. 3, Ex. 8

Problem 3.9 Rudin Ch. 3, Ex. 9

Problem 3.10 Rudin Ch. 3, Ex. 10

Problem 3.11 Rudin Ch. 3, Ex. 11

Problem 3.12 Rudin Ch. 3, Ex. 12

Problem 3.13 Rudin Ch. 3, Ex. 13

Problem 3.14 Rudin Ch. 3, Ex. 14

Problem 3.15 Rudin Ch. 3, Ex. 15

Problem 3.16 Rudin Ch. 3, Ex. 16

Problem 3.17 Rudin Ch. 3, Ex. 17

Problem 3.18 Rudin Ch. 3, Ex. 18

Problem 3.19 Rudin Ch. 3, Ex. 19

Problem 3.20 Rudin Ch. 3, Ex. 20

Problem 3.21 Rudin Ch. 3, Ex. 21

Problem 3.22 Rudin Ch. 3, Ex. 22

Problem 3.23 Rudin Ch. 3, Ex. 23

Problem 3.24 Rudin Ch. 3, Ex. 24

Problem 3.25 Rudin Ch. 3, Ex. 25

Rudin Chapter 4 – Continuity

Rudin Chapter 3 – Numerical Sequences and Series

Problem 3.1 *Rudin, Chapter 3, Exercise 1*

Prove that convergence of $\{s_n\}$ implies convergence of $\{|s_n|\}$. Is the converse true?

SOLUTION.

> The sequence $\{s_n\}$ lives in $\mathbb{R}$ (or $\mathbb{C}$), with $|\cdot|$ the absolute value (modulus); a sequence $x_n \to x$ when for every $\varepsilon > 0$ there is an N with $|x_n - x| < \varepsilon$ for all $n \geq N$ (5.1.1).

Suppose $s_n \to s$. The reverse triangle inequality gives, for every n,

$$\big|\,|s_n| - |s|\,\big| \le |s_n - s|.$$

Fix $\varepsilon > 0$. By convergence there is an N with $|s_n - s| < \varepsilon$ for all $n \ge N$, and then $\big|\,|s_n| - |s|\,\big| < \varepsilon$ for all $n \ge N$. Hence $|s_n| \to |s|$, so $\{|s_n|\}$ converges.

The converse is false. Take $s_n = (-1)^n$ in $\mathbb{R}$. Then $|s_n| = 1$ for every n, so $\{|s_n|\}$ is the constant sequence 1 and converges to 1. But $\{s_n\}$ diverges: it has a subsequence $s_{2n} = 1 \to 1$ and a subsequence $s_{2n+1} = -1 \to -1$, and a convergent sequence has a single limit (at most one limit; 5.1.3), so two subsequences with different limits forbid convergence. Thus $\{|s_n|\}$ can converge while $\{s_n\}$ does not.

■

Reflections.

The statement is an implication between two convergence claims, so the word doing the work is *convergence* itself, and the lecture object it names is the ε–N definition of a limit (5.1.1). Reading the hypothesis and the conclusion side by side shows what the proof must supply: the hypothesis hands you control of $|s_n - s|$, and the conclusion asks for control of $\big|\,|s_n| - |s|\,\big|$, so the entire task is to bound the second quantity by the first. The clue that such a bound exists is the absolute-value bars wrapped around the terms—whenever a problem replaces s_n by $|s_n|$, the reverse triangle inequality $\big|\,|a| - |b|\,\big| \le |a - b|$ is the tool that compares the two, because it says taking absolute values can only shrink distances, never enlarge them.

It helps to see why that inequality should hold before leaning on it. The ordinary triangle inequality $|a| \le |a-b| + |b|$ rearranges to $|a| - |b| \le |a-b|$, and swapping the roles of a and b gives $|b| - |a| \le |a - b|$; together these two say $\big|\,|a| - |b|\,\big| \le |a - b|$. So the map $x \mapsto |x|$ is 1-Lipschitz, and a 1-Lipschitz map cannot pull a convergent sequence apart.

With that in hand the forward direction is a single clean estimate:

(1) the reverse triangle inequality bounds $\big|\,|s_n| - |s|\,\big|$ by $|s_n - s|$ for every n, turning a question about $\{|s_n|\}$ into the question about $\{s_n\}$ the hypothesis already answers;

(2) given $\varepsilon > 0$, convergence $s_n \to s$ supplies an N past which $|s_n - s| < \varepsilon$ (5.1.1), and chaining the two inequalities makes $\big|\,|s_n| - |s|\,\big| < \varepsilon$ for all $n \ge N$, which is exactly $|s_n| \to |s|$.

The same N that works for $\{s_n\}$ works verbatim for $\{|s_n|\}$, since the bound holds termwise; no new tolerance has to be manufactured. Phrased through the tail, every ball about s eventually swallows the s_n (all but finitely many terms), and the Lipschitz bound carries that tail straight to a ball about $|s|$.

The converse question is answered by a counterexample, and the reason one is even worth hunting for is that the forward proof used *both* the size and the sign of s_n only through $|s_n - s|$; the modulus $|s_n|$ throws the sign away, and information thrown away cannot in general be recovered. The cleanest witness is $s_n = (-1)^n$: its moduli are all 1, a constant convergent sequence, while the signs oscillate forever. The rigor of "$\{s_n\}$ diverges" rests on uniqueness of limits (a sequence has at most one limit; 5.1.3)—two subsequences converging to 1 and to -1 cannot both be tails of one convergent sequence. So the implication is strictly one-directional: passing to $|s_n|$ is a genuine loss, not an equivalence.

The transferable move is to recognise that any 1-Lipschitz (or even continuous) operation applied termwise preserves convergence, the absolute value being the first instance; the matching lesson is that such an operation can collapse distinct inputs, so the converse fails exactly when the operation is not injective on the limit's neighbourhood. The borderline case is worth keeping in mind: if $s_n \to 0$ then $|s_n| \to 0$, and here the converse *does* hold, since $|s_n| \to 0$ forces $s_n \to 0$ by the very same termwise bound $|s_n - 0| = \big|\,|s_n| - 0\,\big|$—the sign cannot hide near 0.

Problem 3.2 *Rudin, Chapter 3, Exercise 2*

Calculate $\lim_{n\to\infty}\left(\sqrt{n^2+n}-n\right)$.

SOLUTION.

The limit is $\frac{1}{2}$. For $n \geq 1$ write $a_n = \sqrt{n^2+n} - n$. Rationalising the difference by its conjugate,

$$a_n = (\sqrt{n^2+n} - n)\cdot\frac{\sqrt{n^2+n}+n}{\sqrt{n^2+n}+n} = \frac{(n^2+n)-n^2}{\sqrt{n^2+n}+n} = \frac{n}{\sqrt{n^2+n}+n},$$

the denominator being positive, so no division is illegitimate. Dividing numerator and denominator by $n > 0$,

$$a_n = \frac{1}{\sqrt{1+\frac{1}{n}}+1}.$$

As $n \to \infty$ the standard limit $\frac{1}{n} \to 0$ (6.5.1; Archimedean) gives $1+\frac{1}{n} \to 1$, and the square root is continuous at 1, so $\sqrt{1+\frac{1}{n}} \to 1$; by the algebra of limits (5.2.1; limits respect the algebraic operations) the denominator tends to $1+1 = 2 \neq 0$, and the quotient tends to $\frac{1}{1+1} = \frac{1}{2}$.

To keep the appeal to the square root self-contained, the same value follows from a squeeze. Since $\left(n+\frac{1}{2}\right)^2 = n^2+n+\frac{1}{4} > n^2+n > n^2$, taking positive roots gives $n < \sqrt{n^2+n} < n+\frac{1}{2}$, hence $2n < \sqrt{n^2+n}+n < 2n+\frac{1}{2}$. Inverting and multiplying by $n > 0$,

$$\frac{n}{2n+\frac{1}{2}} < a_n < \frac{n}{2n} = \frac{1}{2}, \qquad \text{that is} \qquad \frac{1}{2+\frac{1}{2n}} < a_n < \frac{1}{2}.$$

The left bound is $\dfrac{1}{2+1/(2n)} \to \dfrac{1}{2+0} = \frac{1}{2}$ by the algebra of limits, and the right bound is the constant $\frac{1}{2}$; both endpoints converge to $\frac{1}{2}$, so by the squeeze theorem (6.3.3) $a_n \to \frac{1}{2}$. Either way,

$$\lim_{n\to\infty}\left(\sqrt{n^2+n}-n\right) = \frac{1}{2}.$$

■

Reflections.

The expression $\sqrt{n^2+n}-n$ is a difference of two quantities that each march off to $+\infty$, so the symbol $\infty-\infty$ it suggests is indeterminate—the limit laws (5.2.1; limits respect the algebraic operations) say nothing about a difference of two divergent sequences, and reading that is the first move. The clue is the lone square root standing against a polynomial: a root minus its near-polynomial is the textbook cue to multiply by the conjugate, because $(\sqrt{A}-B)(\sqrt{A}+B) = A-B^2$ erases the root and turns a fragile difference into a quotient the algebra of limits can actually touch.

It is worth believing the answer before proving it. The root $\sqrt{n^2+n}$ sits between n and $n+\frac{1}{2}$—it is the geometric-mean-like object just above n—so the gap $\sqrt{n^2+n}-n$ hovers just under $\frac{1}{2}$ and ought to settle there; the perturbation $+n$ under the root contributes exactly half a unit in the limit, not a full one, because the root dilutes it.

The computation then runs in a few controlled steps, each resting on a named result:

(1) multiply by the conjugate $\sqrt{n^2+n}+n$ over itself, legitimate since that factor is strictly positive, collapsing the numerator to $(n^2+n)-n^2 = n$ and producing the quotient $a_n = \dfrac{n}{\sqrt{n^2+n}+n}$;

(2) divide top and bottom by n to expose $a_n = \dfrac{1}{\sqrt{1+1/n}+1}$, the form in which the $n\to\infty$ behaviour is visible;

(3) send $\frac{1}{n}\to 0$, the standard limit (6.5.1) that is the Archimedean property in disguise (Archimedean), so the denominator tends to $1+1=2$, a nonzero value, and the quotient to $\frac{1}{2}$ by the algebra of limits (5.2.1);

(4) to avoid leaning on continuity of the root, bracket $n < \sqrt{n^2+n} < n+\frac{1}{2}$ from $(n+\frac{1}{2})^2 > n^2+n > n^2$, which pins a_n between $\dfrac{1}{2+1/(2n)}$ and $\frac{1}{2}$, both tending to $\frac{1}{2}$, so the squeeze theorem (6.3.3) delivers the same value.

The two routes are not redundant: they isolate exactly the one fact that must be supplied from outside the algebra of limits. The quotient form needs $\sqrt{1+1/n}\to 1$, which is continuity of the square root at 1; the squeeze form replaces that with the bare algebraic

inequality $(n+\frac{1}{2})^2 > n^2+n$ and the order facts behind a squeeze (6.3.1), so it assumes nothing about roots beyond their monotonicity. Both need the denominator's limit to be nonzero—the one hypothesis of the quotient law (5.2.1) that is genuinely load-bearing here; were it 0 the algebra of limits would not apply and the squeeze would be the only honest path. Nothing is circular, since $\frac{1}{n} \to 0$ is proved independently of this limit.

The transferable move is the conjugate trick for any $\infty - \infty$ built from a root: rationalise to convert the indeterminate difference into a determinate quotient, then strip the dominant growth (here a factor of n) so the limit laws can finish. The recurring discipline is to refuse the limit laws on a difference of divergent sequences and instead reshape the expression until every sub-limit it contains exists—only then is the algebra of limits (limits respect the algebraic operations) licensed.

Problem 3.3 *Rudin, Chapter 3, Exercise 3*

If $s_1 = \sqrt{2}$, and

$$s_{n+1} = \sqrt{2+\sqrt{s_n}} \qquad (n = 1, 2, 3, \ldots),$$

prove that $\{s_n\}$ converges, and that $s_n < 2$ for $n = 1, 2, 3, \ldots$.

SOLUTION.

Every term is a square root of a positive number, so $s_n > 0$ for all n, and the recursion is the single increasing map $f(x) = \sqrt{2+\sqrt{x}}$ applied repeatedly: $s_{n+1} = f(s_n)$. The plan is to show $\{s_n\}$ is increasing and bounded above by 2, after which monotone convergence forces a limit.

First, $s_n < 2$ for every n, by induction on n. The base case is $s_1 = \sqrt{2} < 2$. Assume $s_n < 2$. Then $\sqrt{s_n} < \sqrt{2} < 2$, so $2 + \sqrt{s_n} < 4$, and taking square roots (which preserves the order on nonnegative reals) gives

$$s_{n+1} = \sqrt{2+\sqrt{s_n}} < \sqrt{4} = 2.$$

By induction $s_n < 2$ for all $n = 1, 2, 3, \ldots$.

Next, $\{s_n\}$ is increasing, i.e. $s_n < s_{n+1}$ for every n, again by induction. For the base case compare the squares: $s_1^2 = (\sqrt{2})^2 = 2$, while $s_2^2 = 2 + \sqrt{\sqrt{2}} > 2$ since $\sqrt{\sqrt{2}} > 0$. Both s_1 and s_2 are positive, so $s_2^2 > s_1^2$ gives $s_2 > s_1$. For the step, assume $s_n < s_{n+1}$. The map f is increasing on $[0, \infty)$: it is the composition $x \mapsto \sqrt{x} \mapsto 2 + \sqrt{x} \mapsto \sqrt{2+\sqrt{x}}$ of order-preserving steps on the nonnegative reals. Applying f to $s_n < s_{n+1}$ therefore preserves the inequality,

$$s_{n+1} = f(s_n) < f(s_{n+1}) = s_{n+2},$$

which is the statement at $n+1$. By induction $s_n < s_{n+1}$ for all n.

Thus $\{s_n\}$ is a monotonically increasing sequence of real numbers bounded above (by

2). By monotone convergence (a bounded monotonic real sequence converges; 6.2.2) the sequence converges to a limit L, with $s_n \leq L \leq 2$ for every n. In particular $\{s_n\}$ converges and $s_n < 2$ for $n = 1, 2, 3, \ldots$.

■

Reflections.

The statement hands you a sequence not by a formula for s_n but by a rule that builds each term from the one before, $s_{n+1} = \sqrt{2 + \sqrt{s_n}}$—a recursively defined sequence. There is no closed form to take a limit of, so the convergence test that needs no formula and no advance knowledge of the limit is the one to reach for: monotone convergence (6.2.2), which certifies a limit from two qualitative facts, monotonicity and boundedness (6.2.1). That the problem also asks for the explicit ceiling $s_n < 2$ is the second clue: it is not a side request but the very bound the convergence proof will run on, so the two halves of the task are really one.

Believe the picture before proving it. Plotting the first terms, $s_1 \approx 1.414$, then 1.786, 1.827, 1.831, …, the values climb but slow down and crowd against a ceiling a little below 1.832, never crossing 2. A sequence that rises yet is penned beneath a fixed wall has nowhere to go but toward a limit—this is exactly the intuition monotone convergence makes rigorous.

The argument splits into three moves, each leaning on induction (0.6.2) or the convergence theorem:

(1) the ceiling $s_n < 2$ comes by induction: $s_1 = \sqrt{2} < 2$, and if $s_n < 2$ then $\sqrt{s_n} < \sqrt{2} < 2$ forces $2 + \sqrt{s_n} < 4$ and hence $s_{n+1} < \sqrt{4} = 2$—the bound reproduces itself because 2 is a fixed point of the estimate;

(2) monotonicity $s_n < s_{n+1}$ comes by a second induction, but the cleanest engine is that the generating map $f(x) = \sqrt{2 + \sqrt{x}}$ is *increasing*, being a composition of order-preserving steps; once $s_2 > s_1$ is checked by hand (squaring turns it into $2 + \sqrt{\sqrt{2}} > 2$), feeding $s_n < s_{n+1}$ through the increasing f yields $s_{n+1} < s_{n+2}$ for free;

(3) with the sequence increasing and bounded above by 2, monotone convergence (bounded monotonic real sequence; 6.2.2) delivers a limit L, and order-preservation in the limit (6.3.1) keeps $L \leq 2$.

The rigor rests on the two inductions being genuinely independent and each self-propagating. The bound in (1) is what makes the increasing step in (2) more than wishful—a sequence can rise without bound, and "increasing" alone never gives convergence—while monotonicity is what rules out the bounded sequence merely oscillating below 2 without settling. Drop either hypothesis and the theorem has nothing to stand on: that is the

content of 6.2.2, and it is why exhibiting both is the whole proof. The increasing-map argument for (2) deserves a second look, since the alternative—comparing s_{n+1} and s_n by squaring at every step—drowns in nested radicals; reading the recursion as "apply one fixed increasing function" converts monotonicity of the sequence into the single, checkable fact that f preserves order, and then induction propagates one inequality the length of the chain.

This problem also quietly answers "what is the limit?" even though it was not asked. Passing to the limit in $s_{n+1} = f(s_n)$, the limit L must satisfy $L = \sqrt{2+\sqrt{L}}$: the left side $s_{n+1} \to L$, while the right side $f(s_n) \to f(L)$ because f is continuous, and a sequence has one limit, so $L = f(L)$; thus L is the fixed point of f, the wall the terms approached. The transferable move is the template for any sequence given by $s_{n+1} = f(s_n)$: guess the trapping bound (often the fixed point of the estimate), prove it by induction, prove monotonicity—most painlessly by showing f is increasing and checking the first comparison—and let monotone convergence finish. The recursion never needs to be unwound; its limit is read off afterward by solving $L = f(L)$.

Problem 3.4 *Rudin, Chapter 3, Exercise 4*

Find the upper and lower limits of the sequence $\{s_n\}$ defined by

$$s_1 = 0; \qquad s_{2m} = \frac{s_{2m-1}}{2}; \qquad s_{2m+1} = \frac{1}{2} + s_{2m}.$$

SOLUTION.

The verdict: $\liminf_{n\to\infty} s_n = \frac{1}{2}$ and $\limsup_{n\to\infty} s_n = 1$.

Compute the first terms to read the pattern. From $s_1 = 0$ the recursion gives $s_2 = 0$, $s_3 = \frac{1}{2}$, $s_4 = \frac{1}{4}$, $s_5 = \frac{3}{4}$, $s_6 = \frac{3}{8}$, $s_7 = \frac{7}{8}$, $s_8 = \frac{7}{16}$, $s_9 = \frac{15}{16}$, which suggests two interleaved patterns. We prove, for every $m \geq 1$,

$$s_{2m+1} = 1 - \frac{1}{2^m}, \qquad s_{2m} = \frac{1}{2} - \frac{1}{2^m}.$$

Both formulas follow by induction on m. For $m = 1$: $s_2 = s_1/2 = 0 = \frac{1}{2} - \frac{1}{2}$, and $s_3 = \frac{1}{2} + s_2 = \frac{1}{2} = 1 - \frac{1}{2}$, so the base case holds. Assume both hold at some $m \geq 1$. The even recursion at $m+1$ reads $s_{2(m+1)} = s_{2m+1}/2$, and using $s_{2m+1} = 1 - 2^{-m}$,

$$s_{2m+2} = \frac{1 - 2^{-m}}{2} = \frac{1}{2} - \frac{1}{2^{m+1}}.$$

The odd recursion then gives

$$s_{2(m+1)+1} = \frac{1}{2} + s_{2m+2} = \frac{1}{2} + \left(\frac{1}{2} - \frac{1}{2^{m+1}}\right) = 1 - \frac{1}{2^{m+1}},$$

which is the pair of formulas at $m+1$. By induction they hold for all $m \geq 1$.

These closed forms split $\{s_n\}$ into its even- and odd-indexed parts. Since $2^{-m} \to 0$ (Archimedean property; the standard limit 6.5.1), the even subsequence $\{s_{2m}\}_{m\geq 1}$ increases to $\frac{1}{2}$ and the odd subsequence $\{s_{2m+1}\}_{m\geq 1}$ increases to 1; each is bounded and monotone, so each converges (monotone convergence; 6.2.2). Thus $\frac{1}{2}$ and 1 are subsequential limits of $\{s_n\}$ (5.3.1).

No other value is a subsequential limit. Suppose $\{s_{n_k}\}$ converges to some ℓ. Among its indices n_k, either infinitely many are even or infinitely many are odd, so $\{s_{n_k}\}$ has a further subsequence lying entirely inside the even family $\{s_{2m}\}$ or entirely inside the odd family $\{s_{2m+1}\}$. A subsequence of a convergent sequence has the same limit—if a sequence converges to a point, every neighborhood of that point contains all but finitely many terms, hence all but finitely many terms of any subsequence (tail trapping; 5.1.3). So that further subsequence converges to $\frac{1}{2}$ or to 1; but it is also a subsequence of $\{s_{n_k}\}$, which converges to ℓ, so by the same principle its limit is ℓ. Hence $\ell \in \{\frac{1}{2}, 1\}$.

The set of subsequential limits is therefore exactly $E = \{\frac{1}{2}, 1\}$. The sequence is bounded (every term lies in $[0, 1]$), so the upper and lower limits are the largest and smallest members of E (6.4.5):

$$\liminf_{n\to\infty} s_n = \min E = \frac{1}{2}, \qquad \limsup_{n\to\infty} s_n = \max E = 1.$$

■

Reflections.

The phrase *upper and lower limits* names the lecture object outright: the lim sup and lim inf, defined through the tail extrema but most usable through the characterization that, for a bounded sequence, they are the largest and smallest of the *subsequential* limits (6.4.5). So the task is not to evaluate a single limit—there is none, the terms oscillate forever—but to find every value the sequence clusters at and then read off the top and bottom of that set. The recursion is the second clue: it defines s_{2m} and s_{2m+1} by separate rules, so the sequence is two interleaved threads, and "even index versus odd index" is the parity split that turns one oscillating sequence into two well-behaved subsequences (5.3.1).

Believe it first by listing terms. The even-indexed values $0, \frac{1}{4}, \frac{3}{8}, \frac{7}{16}, \ldots$ climb toward $\frac{1}{2}$, while the odd-indexed values $\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \ldots$ climb toward 1; the sequence is forever ratcheting up to 1 on odd steps and getting halved back down toward $\frac{1}{2}$ on even steps. Two accumulation heights, $\frac{1}{2}$ and 1, are visible before any formula is written.

The proof then proceeds in four movements, each resting on a named result:

(1) guess and verify closed forms $s_{2m+1} = 1 - 2^{-m}$ and $s_{2m} = \frac{1}{2} - 2^{-m}$ by induction on

m (6.2.1 for the monotone reading; the induction itself is 0.6.2), the even and odd recursions feeding each other so the two formulas advance together in one step;

(2) read off that each subsequence is bounded and increasing, hence convergent by monotone convergence (a bounded monotone sequence converges; 6.2.2), with limits $\frac{1}{2}$ and 1 since $2^{-m} \to 0$ (6.5.1);

(3) rule out any third cluster value: an arbitrary convergent subsequence must reuse infinitely many even *or* infinitely many odd indices, so it has a sub-subsequence trapped in one thread, and a subsequence inherits its parent's limit—every neighborhood of the limit holds all but finitely many terms, so it holds all but finitely many of any subsequence (tail trapping; 5.1.3)—pinning the limit to $\frac{1}{2}$ or 1;

(4) with the set of subsequential limits now exactly $E = \{\frac{1}{2}, 1\}$ and the sequence bounded, invoke the characterization $\liminf = \min E$, $\limsup = \max E$ (6.4.5).

Step (3) is the one that is easy to skip and the one that makes the answer honest. Exhibiting two subsequential limits proves only $\frac{1}{2}, 1 \in E$, hence $\liminf \leq \frac{1}{2}$ and $\limsup \geq 1$; it does *not* show these are the extremes, because a stray subsequence drawing terms from both threads could in principle settle somewhere else. The parity dichotomy closes that door: every subsequence is eventually dominated by one thread, so E can hold nothing beyond $\frac{1}{2}$ and 1. Without this upper bound on E one has computed two cluster points, not the upper and lower limits. The boundedness of $\{s_n\}$ is also load-bearing—it is the hypothesis under which $\limsup$ and $\liminf$ equal $\max E$ and $\min E$ rather than $\pm\infty$ (6.4.7), and here it is free since every term lands in $[0, 1]$.

The transferable move is to treat a sequence defined by alternating rules as a braid of subsequences indexed by the residue class: split on the index modulo the period, settle each thread on its own (monotone convergence does it here), and then assemble the upper and lower limits as the max and min of the threads' limits—valid precisely because every subsequence is eventually swallowed by one thread. The extremes are attained as actual subsequential limits (6.4.6), which is why the verdict is a clean $\frac{1}{2}$ and 1 rather than an unapproached bound.

Problem 3.5 *Rudin, Chapter 3, Exercise 5*

For any two real sequences $\{a_n\}$, $\{b_n\}$, prove that

$$\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n,$$

provided the sum on the right is not of the form $\infty - \infty$.

SOLUTION.

Read lim sup through its definition as a limit of tail suprema (6.4.5): for a real sequence $\{c_n\}$ write

$$C_n = \sup_{k \geq n} c_k \in (-\infty, +\infty],$$

the supremum of a nonempty set of reals, so C_n is never $-\infty$. The tails are nested, so C_n is non-increasing in n, and $\limsup_n c_n = \lim_n C_n = \inf_n C_n$ in the extended reals. Apply this to the three sequences in play, setting

$$A_n = \sup_{k \geq n} a_k, \qquad B_n = \sup_{k \geq n} b_k, \qquad S_n = \sup_{k \geq n} (a_k + b_k),$$

each lying in $(-\infty, +\infty]$, with $\alpha := \limsup a_n = \lim_n A_n$, $\beta := \limsup b_n = \lim_n B_n$, and $\sigma := \limsup(a_n + b_n) = \lim_n S_n$.

The inequality holds at every finite stage. Fix n. For each index $k \geq n$ one has $a_k \leq A_n$ and $b_k \leq B_n$, so $a_k + b_k \leq A_n + B_n$. Because $A_n, B_n > -\infty$, the right-hand side is a well-defined element of $(-\infty, +\infty]$ and is an upper bound for $\{a_k + b_k : k \geq n\}$; taking the supremum over $k \geq n$ gives

$$S_n \leq A_n + B_n \qquad \text{for every } n.$$

Now pass to the limit. As $n \to \infty$, $A_n \downarrow \alpha$ and $B_n \downarrow \beta$ with $\alpha, \beta \in [-\infty, +\infty]$. The hypothesis excludes $\alpha + \beta = \infty - \infty$, so the sum $\alpha + \beta$ is defined in the extended reals and $A_n + B_n \to \alpha + \beta$ (the limit of a sum is the sum of the limits whenever the latter is not of the form $\infty - \infty$). Letting $n \to \infty$ in $S_n \leq A_n + B_n$ and using that the order $\leq$ is preserved under limits (6.3.1) yields

$$\sigma = \lim_n S_n \leq \lim_n (A_n + B_n) = \alpha + \beta,$$

that is, $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$.

■

Reflections.

The single word that fixes the strategy is *limsup*, and reading the statement well means deciding *which* of its faces to use. Our notes carry two: the upper limit is the largest subsequential limit (6.4.6), and it is the limit of the tail suprema $\sup_{k \geq n} c_k$ (6.4.5). A subsequential-limit attack would have to extract a subsequence of $a_n + b_n$ converging to $\limsup(a_n + b_n)$ and then control a and b *along the same indices*—awkward, because the subsequence chasing $\limsup(a_n + b_n)$ need not chase either $\limsup a_n$ or $\limsup b_n$. The tail-supremum face has no such coupling: a supremum over a block of indices already dominates every term in the block at once, which is exactly the kind of crude, simultaneous bound an inequality between three separate lim sups wants.

Believe the inequality, and its slack, on a small case first. Take $a_n = (-1)^n$ and $b_n = (-1)^{n+1}$, so each oscillates between ± 1 with $\limsup a_n = \limsup b_n = 1$; the right side is 2. But $a_n + b_n = 0$ for every n, so $\limsup(a_n + b_n) = 0$. The gap $0 < 2$ is not a defect—it is the whole phenomenon: each sequence peaks at $+1$, but on *opposite* parities, so their peaks never coincide and the sum never sees a 2. The right-hand side optimizes a and b independently; the left-hand side is forced to pick a common index. That picture is why one expects $\leq$ and not $=$.

The proof is short once the tail-supremum reading is in hand, and each move rests on the one before:

(1) rewrite all three lim sups as limits of tail suprema A_n, B_n, S_n (6.4.5), and note each tail supremum, being a sup of reals, lies in $(-\infty, +\infty]$ and decreases as the tail shrinks—this nestedness is what makes $\lim_n$ exist as $\inf_n$;

(2) at a *fixed* n, every $k \geq n$ satisfies $a_k \leq A_n$ and $b_k \leq B_n$, hence $a_k + b_k \leq A_n + B_n$; this is the only genuine idea, the subadditivity of a supremum, and it is legitimate because $A_n, B_n > -\infty$ keeps the bound $A_n + B_n$ from ever being an undefined $\infty - \infty$;

(3) taking the supremum over $k \geq n$ promotes the term bound to $S_n \leq A_n + B_n$, true for every n;

(4) let $n \to \infty$: order survives the limit (6.3.1), each tail supremum being non-increasing so that its limit is its infimum in the extended reals—which is how the upper limit is defined and why it always exists (6.4.5, 6.4.7), no boundedness assumed— and $A_n + B_n \to \alpha + \beta$ because the limit of a sum is the sum of the limits in the extended reals—valid precisely when the sum is not $\infty - \infty$, which is where the lone hypothesis is spent.

The rigor turns entirely on extended-real bookkeeping, and there are two places to be careful. First, a tail supremum is never $-\infty$ (it is the sup of a nonempty set of real numbers), so it can only be a real or $+\infty$; this is what guarantees the finite-stage sum $A_n + B_n$ is always defined, even though $\alpha + \beta$ might not be. Second, the hypothesis "not $\infty - \infty$" is load-bearing exactly at the passage to the limit: drop it and one could have $\limsup a_n = +\infty$, $\limsup b_n = -\infty$, where the right side is meaningless while the left may be anything—e.g. $a_n = n$, $b_n = -n$ gives a right side $\infty - \infty$ and a left side 0, so the statement could not even be parsed, let alone proved. When the right side *is* defined and equals $+\infty$ the claim is vacuous, and the only substantive case is finite α, β, which the same chain covers without separate treatment.

The move to carry away is that an inequality between lim sups of *different* sequences is almost always cleanest through the tail-supremum definition, never the subsequential-limit one: a supremum bounds an entire block of indices simultaneously, so a per-term

inequality survives both the block supremum and the limit untouched, and one never has to align subsequences. The same template, run with inf in place of sup and the inequalities reversed, gives the companion $\liminf(a_n+b_n) \geq \liminf a_n+\liminf b_n$; and the strict-versus-equal question raised by the ± 1 example—when do the peaks line up?—is the upper-limit bookkeeping that drives the oscillating sequence of Problem 3.4.

Problem 3.6 *Rudin, Chapter 3, Exercise 6*

Investigate the behavior (convergence or divergence) of $\sum a_n$ if

1. $a_n = \sqrt{n+1} - \sqrt{n}$;
2. $a_n = \dfrac{\sqrt{n+1} - \sqrt{n}}{n}$;
3. $a_n = (\sqrt[n]{n} - 1)^n$;
4. $a_n = \dfrac{1}{1+z^n}$, for complex values of z.

Solution.

Throughout, the rationalisation $\sqrt{n+1} - \sqrt{n} = \dfrac{1}{\sqrt{n+1}+\sqrt{n}}$, obtained by multiplying by the conjugate, converts each difference of roots into a positive fraction whose size is transparent.

For (a), the series *diverges*. Rationalising,

$$a_n = \frac{1}{\sqrt{n+1}+\sqrt{n}} \geq \frac{1}{2\sqrt{n+1}} > 0,$$

so by comparison (6.6.8) it suffices that $\sum 1/\sqrt{n+1}$ diverge, which it does: it is the p-series with $p = \frac{1}{2} \leq 1$ (6.6.5). One can also see it outright, since the partial sums telescope: $\sum_{n=0}^{N}(\sqrt{n+1} - \sqrt{n}) = \sqrt{N+1} \to \infty$.

For (b), the series *converges* (the terms begin at $n = 1$, where a_n is defined). Rationalising the numerator,

$$a_n = \frac{\sqrt{n+1} - \sqrt{n}}{n} = \frac{1}{n(\sqrt{n+1}+\sqrt{n})} \leq \frac{1}{n \cdot \sqrt{n}} = \frac{1}{n^{3/2}},$$

since $\sqrt{n+1} + \sqrt{n} > \sqrt{n}$. The bounding series $\sum 1/n^{3/2}$ is the p-series with $p = \frac{3}{2} > 1$, hence convergent (6.6.5); as $0 \leq a_n \leq n^{-3/2}$, comparison (6.6.8) gives convergence.

For (c), the series *converges*, by the root test. With $a_n = (\sqrt[n]{n} - 1)^n \geq 0$,

$$\sqrt[n]{a_n} = \sqrt[n]{n} - 1.$$

The standard limit $\sqrt[n]{n} \to 1$ (6.5.1) gives $\sqrt[n]{a_n} \to 0$, so $\limsup_n \sqrt[n]{a_n} = 0 < 1$, and the

root test (6.6.9) yields convergence.

For (d) the answer depends on $|z|$: the series *converges* when $|z| > 1$ and *diverges* when $|z| \leq 1$ (excluding any z with $z^n = -1$ for some n, where a term is undefined). The dividing line is whether the terms tend to 0, the necessary condition for convergence (6.6.3).

- If $|z| < 1$, then $z^n \to 0$, so $a_n = \dfrac{1}{1+z^n} \to 1 \neq 0$; the terms do not vanish, so the series diverges.
- If $|z| = 1$, then $|z^n| = 1$, so $|1 + z^n| \leq 1 + |z^n| = 2$ and hence $|a_n| = \dfrac{1}{|1+z^n|} \geq \frac{1}{2}$ whenever $1 + z^n \neq 0$; again the terms fail to tend to 0, and the series diverges.
- If $|z| > 1$, then $|z^n| = |z|^n \to \infty$, and the reverse triangle inequality gives $|1 + z^n| \geq |z|^n - 1 > 0$ for all large n, so

$$|a_n| = \frac{1}{|1+z^n|} \leq \frac{1}{|z|^n - 1}.$$

  Taking nth roots, $\left(|z|^n - 1\right)^{1/n} = |z|\,(1 - |z|^{-n})^{1/n} \to |z|$, so $\limsup_n \sqrt[n]{|a_n|} \leq 1/|z| < 1$; the root test (6.6.9) makes $\sum |a_n|$ converge, and absolute convergence forces convergence.

Thus $\sum a_n$ converges precisely for $|z| > 1$.

■

REFLECTIONS.

Every part asks the same question—does $\sum a_n$ converge?—and the word *investigate* is the instruction to first reach a verdict and only then justify it, so each part opens by deciding what to prove. The toolkit is fixed and small: a candidate series is weighed against one we already understand, either by the comparison test (6.6.8) against a p-series (6.6.5) or a geometric series (6.6.7), by the root test (6.6.9) when the nth term is built from an nth power, or—when nothing converges—by the cheapest disqualifier of all, that the terms must tend to 0 (6.6.3). Reading which clue each a_n carries is the whole skill the problem trains.

Parts (a) and (b) both wear a difference of square roots, and the reflex a radical difference should trigger is to multiply by the conjugate: $\sqrt{n+1} - \sqrt{n} = 1/(\sqrt{n+1} + \sqrt{n})$ turns a small difference of large numbers into a transparent positive fraction. Once rationalised, the size is plain. In (a) the term is $\asymp 1/(2\sqrt{n})$, a p-series with $p = \frac{1}{2}$, just on the divergent side of the boundary $p = 1$; in (b) the extra factor of n in the denominator pushes the exponent to $p = \frac{3}{2}$, just on the convergent side. The two parts are deliberately paired to sit on opposite shores of the same divide, and the lesson is that $\sum n^{-p}$ converges exactly when $p > 1$ (6.6.5)—the single fact both verdicts rest on.

It is worth believing (a) without any test at all, because its structure is a gift: the partial sum $\sum_{n=0}^{N}(\sqrt{n+1}-\sqrt{n})$ telescopes to $\sqrt{N+1}$, which marches off to infinity. A telescoping sum exposes the partial sums directly (6.6.1), and seeing $\sqrt{N+1}\to\infty$ is more convincing than any comparison.

Part (c) advertises its method in its shape: an nth term that is itself an nth power, $(\sqrt[n]{n}-1)^n$, is precisely the form the root test was built for (6.6.9), since taking $\sqrt[n]{\cdot}$ peels the power away cleanly. What remains, $\sqrt[n]{n}-1$, is governed by the standard limit $\sqrt[n]{n}\to 1$ (6.5.1); the base creeps down to 1, so the whole power collapses geometrically, and $\limsup\sqrt[n]{a_n}=0<1$ closes it. Trying the ratio test here would mean wrestling with $(\sqrt[n+1]{n+1}-1)^{n+1}/(\sqrt[n]{n}-1)^n$, a far uglier object—the nth power is the signpost pointing at the root test, not the ratio test.

Part (d) is the one that rewards reading the statement slowly: "for complex values of z" is an invitation to split on $|z|$, because the entire behaviour of z^n is dictated by its modulus—it shrinks to 0, stays on the unit circle, or blows up, according to whether $|z|$ is below, on, or above 1. The argument tracks those three regimes:

(1) for $|z|<1$ the powers $z^n\to 0$ (5.1.1), so $a_n\to 1\neq 0$ and the term test (6.6.3) kills convergence before any comparison is needed;

(2) for $|z|=1$ the powers ride the unit circle, so $|1+z^n|\le 2$ forces $|a_n|\ge\frac{1}{2}$, and the terms again fail to vanish—the term test does the work a second time, now via a lower bound rather than a limit;

(3) for $|z|>1$ the powers escape to infinity, the reverse triangle inequality gives $|1+z^n|\ge |z|^n-1$, and $|a_n|\le 1/(|z|^n-1)$ behaves like the convergent geometric tail $|z|^{-n}$; the root test (6.6.9), using $(|z|^n-1)^{1/n}\to|z|$, certifies $\limsup\sqrt[n]{|a_n|}\le 1/|z|<1$ and hence absolute convergence.

Two points keep (d) honest. The cases $|z|<1$ and $|z|=1$ are genuinely different arguments, not one with a shared limit: on the circle z^n need not converge at all (take $z=-1$, where z^n alternates and some terms are even undefined), so one cannot say $a_n\to$ anything—only that $|a_n|$ stays bounded below, which is all the term test needs. And the comparison in case (3) must run on the *moduli*; $\sum a_n$ is a complex series, so convergence is established through $\sum|a_n|$, the absolute-convergence detour that lets a real positive test settle a complex series. Across all four parts no step is circular, since each verdict is reached by comparison to a series whose fate is known independently.

The transferable instinct is to classify before you compute: a difference of roots wants its conjugate and a p-series comparison; an nth power of an nth term wants the root test; a parameter z entering only through powers wants a split on $|z|$; and whenever a comparison or root estimate would be laborious, first check the free necessary condition that the terms

tend to 0 (6.6.3)—it disqualifies a divergent series in one line, as it does twice in part (d).

Problem 3.7 *Rudin, Chapter 3, Exercise 7*

Prove that the convergence of $\sum a_n$ implies the convergence of

$$\sum \frac{\sqrt{a_n}}{n},$$

if $a_n \geq 0$.

SOLUTION.

The terms run over $n \geq 1$, so that $1/n$ is defined; every term $\sqrt{a_n}/n$ is ≥ 0, since $a_n \geq 0$.

The one inequality the proof needs is the bound between geometric and arithmetic means of two nonnegative reals: for $x, y \geq 0$,

$$\sqrt{xy} \leq \tfrac{1}{2}(x+y),$$

which follows from $\left(\sqrt{x}-\sqrt{y}\right)^2 \geq 0$, i.e. $x - 2\sqrt{xy} + y \geq 0$. Apply it with $x = a_n$ and $y = 1/n^2$, both ≥ 0:

$$\frac{\sqrt{a_n}}{n} = \sqrt{a_n \cdot \frac{1}{n^2}} \leq \frac{1}{2}\Big(a_n + \frac{1}{n^2}\Big).$$

The right-hand side is the term of a convergent series. The series $\sum a_n$ converges by hypothesis, and the p-series $\sum 1/n^2$ converges, being a p-series with $p = 2 > 1$ (6.6.5). Convergent series add term by term (5.2.1; algebra of limits applied to the partial sums), so

$$\sum_{n\geq 1} \frac{1}{2}\Big(a_n + \frac{1}{n^2}\Big) = \frac{1}{2}\sum_{n\geq 1} a_n + \frac{1}{2}\sum_{n\geq 1}\frac{1}{n^2}$$

converges.

Now the comparison test finishes it. The estimate $0 \leq \sqrt{a_n}/n \leq \frac{1}{2}(a_n + 1/n^2)$ holds for every $n \geq 1$ with a convergent majorant, so $\sum \sqrt{a_n}/n$ converges by comparison (6.6.8).

■

REFLECTIONS.

Two features of the statement set the whole approach. First, the hypothesis $a_n \geq 0$ together with $1/n > 0$ makes every term $\sqrt{a_n}/n$ nonnegative, and a nonnegative series is the friendly case: its partial sums increase, so it converges exactly when they stay bounded (6.6.6), and the entire arsenal of comparison-type tests is available (6.6.8). Second, the term $\sqrt{a_n}/n$ is a *product* of two pieces—$\sqrt{a_n}$, which the hypothesis controls only through $\sum a_n$, and $1/n$, which we know converges when squared. The reflex when an unknown quantity is multiplied against a known summable one is to split the product so each piece feeds a series we already understand, and the clean tool for splitting a product of nonnegatives

into a *sum* is the arithmetic–geometric mean inequality.

Believe it first by trusting the shape of the estimate. The expression $\sqrt{a_n}/n$ is the geometric mean $\sqrt{a_n \cdot 1/n^2}$ of a_n and $1/n^2$; the AM–GM bound trades that geometric mean for the average $\frac{1}{2}(a_n + 1/n^2)$, which is exactly a blend of the two series whose convergence we can certify. The square-root, which is what makes the term awkward to bound directly, is precisely what AM–GM is built to remove.

The argument runs in three steps, each resting on a named result:

(1) the pointwise bound $\sqrt{a_n}/n \le \frac{1}{2}(a_n + 1/n^2)$ comes from $(\sqrt{a_n} - 1/n)^2 \ge 0$, valid because $a_n \ge 0$ lets $\sqrt{a_n}$ exist—this is where the sign hypothesis is spent;

(2) the majorant converges, since $\sum a_n$ converges by hypothesis and $\sum 1/n^2$ is a p-series with $p = 2 > 1$ (6.6.5), and convergent series combine linearly term by term (5.2.1; the algebra of limits on the partial sums);

(3) the comparison test (6.6.8), whose hypothesis is exactly $0 \le b_n \le c_n$ with $\sum c_n$ convergent, transfers convergence from the majorant to $\sum \sqrt{a_n}/n$.

Each hypothesis of the comparison test is genuinely load-bearing, and dropping the sign assumption shows why. The test needs the terms to be nonnegative and dominated by a convergent series; the lower bound $0 \le \sqrt{a_n}/n$ is what rules out the cancellation that would otherwise let a bounded-above series still diverge. The hypothesis $a_n \ge 0$ is used twice over—once so that $\sqrt{a_n}$ is even defined, and once so that the comparison applies—and it is not cosmetic: were the a_n allowed to be negative, $\sqrt{a_n}$ would be meaningless and the statement would not parse. The factor $1/n$ is doing real work too; pairing a_n with $1/n^2$ rather than some larger weight is what keeps the auxiliary p-series convergent, and a heavier weight such as $1/\sqrt{n}$ (whose square $1/n$ gives the divergent harmonic series) would break step (2).

The transferable move is the splitting itself: to tame a product $\sqrt{a_n} \cdot w_n$ of a crudely-controlled factor against a known weight, use AM–GM to dominate it by $\frac{1}{2}(a_n + w_n^2)$ and sum the two halves separately. It is the discrete shadow of the Cauchy–Schwarz reflex, and it works precisely because $\sum w_n^2$—here the p-series $\sum 1/n^2$ (6.6.5)—converges, so the price of removing the square root is a series we have already paid for.

Problem 3.8 *Rudin, Chapter 3, Exercise 8*

If $\sum a_n$ converges, and if $\{b_n\}$ is monotonic and bounded, prove that $\sum a_n b_n$ converges.

Solution.

Write $A_n = \sum_{k=0}^{n} a_k$ for the partial sums of $\sum a_n$, with the convention $A_{-1} = 0$. Since $\sum a_n$ converges, the sequence $\{A_n\}$ converges, hence is bounded (a convergent sequence is bounded): there is M with $|A_n| \le M$ for all n.

The sequence $\{b_n\}$ is monotonic and bounded, so it converges, say $b_n \to b$ (a bounded monotonic sequence converges; 6.2.2). Setting $c_n = b_n - b$, the sequence $\{c_n\}$ is monotonic and $c_n \to 0$. Splitting the terms,

$$a_n b_n = a_n c_n + b\, a_n,$$

the series $\sum b\, a_n = b \sum a_n$ converges (a convergent series scaled by a constant), so it suffices to prove that $\sum a_n c_n$ converges. We verify the Cauchy criterion for it (6.6.2).

The tool is summation by parts. For $0 \le p \le q$, using $a_k = A_k - A_{k-1}$ and reindexing,

$$\sum_{k=p}^{q} a_k c_k = \sum_{k=p}^{q} (A_k - A_{k-1}) c_k = A_q c_q - A_{p-1} c_p + \sum_{k=p}^{q-1} A_k\,(c_k - c_{k+1}).$$

Bound each piece by $|A_k| \le M$. The boundary terms give $|A_q c_q| \le M|c_q|$ and $|A_{p-1} c_p| \le M|c_p|$. For the sum, $\{c_n\}$ is monotonic, so all the differences $c_k - c_{k+1}$ carry the same sign; hence

$$\sum_{k=p}^{q-1} |c_k - c_{k+1}| = \Big|\sum_{k=p}^{q-1} (c_k - c_{k+1})\Big| = |c_p - c_q| \le |c_p| + |c_q|,$$

the middle equality being a telescoping sum. Combining,

$$\begin{aligned} \Big|\sum_{k=p}^{q} a_k c_k\Big| &\le M|c_q| + M|c_p| + M \sum_{k=p}^{q-1} |c_k - c_{k+1}| \\ &\le M|c_q| + M|c_p| + M\big(|c_p| + |c_q|\big) = 2M\big(|c_p| + |c_q|\big). \end{aligned}$$

Now $c_n \to 0$, so given $\varepsilon > 0$ choose N with $|c_n| < \dfrac{\varepsilon}{4M+1}$ for all $n \ge N$ (the $+1$ guards the case $M = 0$). For $q \ge p \ge N$,

$$\Big|\sum_{k=p}^{q} a_k c_k\Big| \le 2M\big(|c_p| + |c_q|\big) < 2M \cdot \frac{2\varepsilon}{4M+1} = \frac{4M}{4M+1}\,\varepsilon < \varepsilon.$$

This is the Cauchy criterion for $\sum a_n c_n$, so that series converges (6.6.2). Adding back the convergent series $b \sum a_n$ gives that $\sum a_n b_n = \sum a_n c_n + b \sum a_n$ converges.

■

REFLECTIONS.

The hypotheses are oddly mismatched, and that mismatch is the whole clue. About $\sum a_n$ we are told only that it *converges*—nothing about its terms or its sign—while about $\{b_n\}$ we are handed two strong shape conditions, *monotonic* and *bounded*. When a weak control on one factor is paired with rigid monotonicity on the other, the pairing names a single technique: summation by parts, the discrete analogue of integration by parts, which trades

the unknown product $a_n b_n$ for partial sums A_n (which we *do* control) multiplied against the *differences* $b_k - b_{k+1}$ (which monotonicity makes well-behaved). The two words "monotonic and bounded" together are also the exact hypothesis of monotone convergence (a bounded monotonic sequence converges; 6.2.2), so $\{b_n\}$ has a limit b to peel off.

A reader should believe the shape of the answer before the inequalities. The convergence of $\sum a_n$ pins the partial sums A_n inside a bounded window (a convergent sequence is bounded); a monotonic $\{c_n\}$ sliding down to 0 supplies weights whose *total* variation is finite—in fact just $|c_p - c_q|$, because a one-signed difference telescopes. Bounded sums against finite-variation weights ought to converge, and that intuition is exactly what the algebra below makes precise. A concrete instance fitting the hypotheses: take the convergent alternating series $\sum a_n = \sum(-1)^n/(n+1)$ and weight it by $b_n = n/(n+1)$, which increases monotonically to 1 and stays bounded; the theorem promises $\sum a_n b_n$ converges, and it does. The degenerate weight $b_n \equiv 1$ recovers the bare convergence of $\sum a_n$, so the content is precisely that a monotone bounded reweighting cannot break a convergent series.

The proof runs in a few linked moves, each resting on the one before:

(1) convergence of $\sum a_n$ makes $\{A_n\}$ a convergent, hence bounded, sequence (convergent $\Rightarrow$ bounded), giving the uniform bound $|A_n| \le M$ that every later estimate leans on;

(2) "monotonic and bounded" lets monotone convergence (6.2.2) extract the limit b and split $b_n = b + c_n$ with $c_n \to 0$ monotonic; the constant part $b\sum a_n$ converges by the algebra of series (limits respect the algebraic operations; 5.2.1), reducing everything to the single series $\sum a_n c_n$;

(3) summation by parts rewrites $\sum_{k=p}^{q} a_k c_k$ via $a_k = A_k - A_{k-1}$ as two boundary terms plus $\sum A_k(c_k - c_{k+1})$—this is the step that swaps the uncontrolled a_k for the controlled A_k;

(4) monotonicity of $\{c_n\}$ forces the differences $c_k - c_{k+1}$ to share one sign, so their absolute values telescope to $|c_p - c_q|$, and the whole block is bounded by $2M(|c_p| + |c_q|)$;

(5) $c_n \to 0$ drives that bound below any ε for p, q large, which is precisely the Cauchy criterion for series (6.6.2), and completeness of $\mathbb{R}$ turns "Cauchy" into "convergent" (Cauchy sequences in $\mathbb{R}$ converge).

Each hypothesis is load-bearing, and dropping any one breaks the result. Without convergence of $\sum a_n$ there is no bound M, and the boundary term $A_q c_q$ can run off even as $c_q \to 0$ (take $a_n = 1$, $b_n = 1$: $\sum a_n b_n$ diverges). Without monotonicity of $\{b_n\}$ the telescoping collapses—the differences may alternate in sign and their absolute sum can grow without bound, so the estimate $\sum |c_k - c_{k+1}| = |c_p - c_q|$ fails; a bounded but wildly oscillating

$\{b_n\}$ can destroy convergence even when $\sum a_n$ converges. Boundedness of $\{b_n\}$ is what manufactures the limit b; a monotonic but unbounded $\{b_n\}$ (say $b_n = n$) has $c_n \to \pm\infty$ and the argument has nothing to drive to zero. The reasoning is not circular: the bound on A_n and the limit b are established independently before summation by parts is invoked, and the Cauchy criterion is verified directly, never assuming the conclusion.

The transferable move is summation by parts itself, and the principle it encodes: to control a series whose terms are a product, do not estimate the product head-on—resum so that the well-behaved factor appears as a bounded partial sum and the rigid factor appears only through its differences, whose total variation a monotone sequence keeps finite. This is the engine behind both Dirichlet's test (bounded partial sums against a monotone null sequence) and Abel's test (a convergent series against a monotone bounded sequence), and the present problem is exactly Abel's test. The same trade—swap a difference operator onto the tame factor and a summation operator onto the partial sums—is what integration by parts does for integrals, and recognising the discrete twin is the lasting lesson.

Problem 3.9 *Rudin, Chapter 3, Exercise 9*

Find the radius of convergence of each of the following power series:

(a) $\sum n^3 z^n$,

(b) $\sum \dfrac{2^n}{n!} z^n$,

(c) $\sum \dfrac{2^n}{n^2} z^n$,

(d) $\sum \dfrac{n^3}{3^n} z^n$.

SOLUTION.

For a power series $\sum c_n z^n$ the radius of convergence is $R = 1/\alpha$, where $\alpha = \limsup_{n\to\infty} |c_n|^{1/n}$, with the conventions $R = +\infty$ when $\alpha = 0$ and $R = 0$ when $\alpha = +\infty$ (6.6.11). The single standard limit

$$n^{1/n} \longrightarrow 1 \qquad (6.5.1)$$

settles all four cases, since each $|c_n|^{1/n}$ converges—so its lim sup is its ordinary limit (6.4.5). The verdicts are $R = 1,\ +\infty,\ \frac{1}{2},\ 3$.

For (a), $c_n = n^3$, so $|c_n|^{1/n} = \left(n^{1/n}\right)^3 \to 1^3 = 1$ by the algebra of limits (limits respect products). Hence $\alpha = 1$ and $R = 1$.

For (b), $c_n = 2^n/n!$. Here the ratio is cleaner than the root: with all terms positive,

$$\frac{|c_{n+1}|}{|c_n|} = \frac{2^{n+1}/(n+1)!}{2^n/n!} = \frac{2}{n+1} \longrightarrow 0.$$

Since $|c_{n+1}|/|c_n| \to 0$, the same value bounds $\limsup |c_n|^{1/n}$ from above (6.A.1, the inequality behind the fact that the root test subsumes the ratio test), so $\alpha = 0$ and $R = +\infty$: the series converges for every $z \in \mathbb{C}$.

For (c), $c_n = 2^n/n^2$, so

$$|c_n|^{1/n} = \frac{2}{\left(n^2\right)^{1/n}} = \frac{2}{\left(n^{1/n}\right)^2} \longrightarrow \frac{2}{1^2} = 2.$$

Hence $\alpha = 2$ and $R = \frac{1}{2}$.

For (d), $c_n = n^3/3^n$, so

$$|c_n|^{1/n} = \frac{\left(n^{1/n}\right)^3}{3} \longrightarrow \frac{1}{3}.$$

Hence $\alpha = \frac{1}{3}$ and $R = 3$.

■

Reflections.

The phrase *radius of convergence* names exactly one object in the notes and hands you its formula at the same time: for $\sum c_n z^n$ the radius is $R = 1/\alpha$ with $\alpha = \limsup_n |c_n|^{1/n}$ (6.6.11), the Cauchy–Hadamard recipe. So a problem that asks for four radii is really asking for four computations of $\limsup |c_n|^{1/n}$; the conceptual work is done before any algebra, and what remains is to extract an nth root and take a limit. The reason the root appears at all is that $\sum c_n z^n$ converges precisely where the root test (6.6.9) on the terms $c_n z^n$ returns a value below 1, and $\limsup |c_n z^n|^{1/n} = |z|\,\alpha < 1$ is the inequality $|z| < 1/\alpha$ in disguise.

A single fact does almost all the lifting, and recognising it is the point of the exercise: every coefficient here is a power of n divided by an exponential, and $\left(n^p\right)^{1/n} = \left(n^{1/n}\right)^p \to 1$ because $n^{1/n} \to 1$ (6.5.1). Believe that limit first on (a): the coefficients n^3 grow without bound, yet their nth roots $\left(n^{1/n}\right)^3$ slide down to 1, so a polynomial factor, however large its degree, is invisible to the root test—it contributes a factor tending to 1 and cannot move the radius. That is why (a) has $R = 1$ exactly as the bare geometric series $\sum z^n$ does.

The four computations then run on the same two moves:

(1) take the nth root and split it across the product or quotient, which is legitimate because limits respect the algebraic operations (the algebra of limits, 5.2.1); the polynomial part becomes a power of $n^{1/n} \to 1$ and disappears, leaving only the exponential base—1 in (a), 2 in (c), $\frac{1}{3}$ in (d)—as the value of α;

(2) read off $R = 1/\alpha$, with $\alpha = 0 \Rightarrow R = +\infty$ and $\alpha = +\infty \Rightarrow R = 0$ the two boundary conventions (6.6.11).

Part (b) is the one place a root is awkward—$(n!)^{1/n}$ is not a standard limit in our notes—so the ratio test is the right instrument: $|c_{n+1}|/|c_n| = 2/(n+1) \to 0$. The factorial crushes the geometric growth 2^n, the ratio tends to 0, and a ratio that tends to a limit forces the root lim sup to the same limit (6.A.1); hence $\alpha = 0$ and the series is entire, $R = +\infty$. This is the same mechanism that makes the exponential series converge everywhere (6.6.12).

Two points keep the argument rigorous. First, the formula is stated with a lim sup, not a plain limit, because a general coefficient sequence need not have $|c_n|^{1/n}$ converging; here every one of the four does converge, so the lim sup collapses to the ordinary limit (6.4.5) and no upper envelope is needed—worth saying explicitly rather than silently treating lim sup as lim. Second, the radius records only the open disk $|z| < R$ where convergence is absolute; on the circle $|z| = R$ the formula is silent, and behaviour there genuinely varies—$\sum z^n$, $\sum z^n/n$, and $\sum z^n/n^2$ all have $R = 1$ yet differ completely on $|z| = 1$—so a complete answer names R and stops, without claiming anything on the boundary.

The transferable reading is that the root test is the engine of the radius, and within it a polynomial coefficient is weightless: R is governed entirely by the exponential rate, the base b in a coefficient $\sim n^p b^n$ giving $\alpha = b$ and $R = 1/b$, while factorials send α to 0 and the radius to infinity. When the root is clean, compute $\limsup |c_n|^{1/n}$ directly; when a factorial or a stubborn product appears, switch to the ratio $|c_{n+1}/c_n|$, which agrees with the root whenever it converges (6.A.1).

Problem 3.10 *Rudin, Chapter 3, Exercise 10*

Suppose that the coefficients of the power series $\sum a_n z^n$ are integers, infinitely many of which are distinct from zero. Prove that the radius of convergence is at most 1.

SOLUTION.

The radius of convergence of $\sum a_n z^n$ is given by the Cauchy–Hadamard formula (6.6.11)

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}},$$

with the conventions $1/0 = +\infty$ and $1/\infty = 0$. To prove $R \le 1$ it is enough to prove $\limsup_{n\to\infty} |a_n|^{1/n} \ge 1$.

Let $S = \{n : a_n \neq 0\}$ be the set of indices of nonzero coefficients. By hypothesis S is infinite, so its elements form a strictly increasing sequence $n_1 < n_2 < n_3 < \cdots$. For each $n_k \in S$ the coefficient a_{n_k} is a *nonzero integer*, hence $|a_{n_k}| \ge 1$, and therefore

$$|a_{n_k}|^{1/n_k} \ge 1^{1/n_k} = 1 \qquad \text{for every } k.$$

Thus the subsequence $\left(|a_{n_k}|^{1/n_k}\right)_k$ of the sequence $\left(|a_n|^{1/n}\right)_n$ is bounded below by 1. Being a sequence in the (extended) reals, it has at least one subsequential limit L, and since each

of its terms is ≥ 1 and $[1, +\infty]$ is closed, that limit satisfies $L \geq 1$. The limit superior of $\left(|a_n|^{1/n}\right)_n$ is the largest of all its subsequential limits, and L is one of them (6.4.5, 6.4.6); hence

$$\limsup_{n\to\infty} |a_n|^{1/n} \;\geq\; L \;\geq\; 1.$$

Consequently

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}} \;\leq\; \frac{1}{1} = 1,$$

so the radius of convergence is at most 1.

■

Reflections.

The phrase *radius of convergence* fixes the entire approach: the one tool in our notes that turns a sequence of coefficients into a radius is the Cauchy–Hadamard formula $R = 1/\limsup |a_n|^{1/n}$ (6.6.11). Reading the goal $R \leq 1$ through that formula flips it into a statement about the denominator—$R \leq 1$ is the same as $\limsup |a_n|^{1/n} \geq 1$—and a lower bound on a lim sup is the cheapest kind of statement to establish, because the limit superior only has to be large *somewhere*, along a single subsequence, not everywhere.

That is where the second clue does its work. The hypothesis is oddly specific: the coefficients are not merely real but *integers*, and *infinitely many* are nonzero. Each word earns its place. A nonzero integer cannot be small—it has absolute value at least 1—so on the indices where $a_n \neq 0$ the quantity $|a_n|^{1/n}$ is pinned at or above 1; and "infinitely many" guarantees those indices form a genuine subsequence rather than a finite scatter that a lim sup could ignore. The integrality supplies the floor and the infinitude supplies the subsequence, and the limit superior is precisely the device built to hear an infinitely-recurring floor.

It helps to believe the mechanism on a degenerate case first. Take $a_n = 1$ for all n: every $|a_n|^{1/n} = 1$, so $\limsup = 1$ and $R = 1$, the geometric series $\sum z^n$ converging exactly on $|z| < 1$ (6.6.7). Padding with zeros, as in $\sum z^{n^2}$ where the nonzero coefficients are 1, changes nothing essential: the nonzero terms still force the lim sup to be 1, and the radius stays 1. The bound $R \leq 1$ is therefore sharp, attained by the simplest integer series there is.

The argument is short, and each move rests on the one before:

(1) rewrite the target through Cauchy–Hadamard (6.6.11): $R \leq 1$ is equivalent to $\limsup |a_n|^{1/n} \geq 1$, so the radius is never touched again and the whole problem becomes a statement about the coefficient sequence;

(2) collect the nonzero indices $S = \{n : a_n \neq 0\}$, infinite by hypothesis, and list them as a strictly increasing subsequence $n_1 < n_2 < \cdots$—this is the step that spends "infinitely

many," since only an infinite S yields a true subsequence;

(3) on that subsequence $|a_{n_k}| \geq 1$ because a nonzero integer has absolute value at least 1, and raising a number ≥ 1 to the positive power $1/n_k$ keeps it ≥ 1—this is where "integers" is spent, and it is the only arithmetic in the proof;

(4) read off the lim sup: the subsequence sitting at or above 1 has a cluster value, itself ≥ 1 because $[1, +\infty]$ is closed, and the limit superior, as the largest subsequential limit (6.4.5, 6.4.6), dominates that cluster value, hence $\limsup |a_n|^{1/n} \geq 1$ and $R \leq 1$.

The rigor turns on using lim sup rather than lim in step (4), and on seeing that both hypotheses are load-bearing. The ordinary limit of $|a_n|^{1/n}$ need not exist—between the nonzero indices the coefficients may be 0, where $|a_n|^{1/n} = 0$, so the sequence can oscillate between 0 and values ≥ 1 forever. The limit superior is exactly the right instrument because it ignores the dips to 0 and reports the recurring height ≥ 1; replacing it by a plain limit would be circular, assuming a convergence the problem does not give. Drop "integers" and the floor vanishes—coefficients like $a_n = 1/n^n$ are nonzero yet have $|a_n|^{1/n} = 1/n \to 0$, giving $R = \infty$, so the conclusion is simply false. Drop "infinitely many nonzero" and the surviving terms are a finite set the lim sup discards; a polynomial such as $1 + z$ has all but finitely many coefficients zero and radius $R = \infty$. Each hypothesis blocks exactly one of these escapes.

The transferable move is to read a radius-of-convergence question as a question about the limit superior of the coefficients, and then to control that lim sup along a well-chosen subsequence rather than the whole sequence. A lower bound on lim sup needs only one recurring floor; an upper bound needs the floor to hold eventually everywhere. Here a discreteness constraint—nonzero integers cannot be small—manufactures the floor, the same discreteness that makes $\mathbb{Z}$ a closed, spread-out set and that powers many "integers cannot crowd" arguments. Whenever you must keep $|a_n|^{1/n}$ from collapsing to 0, look for a reason the nonzero a_n stay bounded away from 0, and let the lim sup hear it.

Problem 3.11 *Rudin, Chapter 3, Exercise 11*

Suppose $a_n > 0$, $s_n = a_1 + \cdots + a_n$, and $\sum a_n$ diverges.

1. Prove that $\sum \dfrac{a_n}{1 + a_n}$ diverges.
2. Prove that
$$\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} \geq 1 - \frac{s_N}{s_{N+k}}$$
and deduce that $\sum \dfrac{a_n}{s_n}$ diverges.
3. Prove that
$$\frac{a_n}{s_n^2} \leq \frac{1}{s_{n-1}} - \frac{1}{s_n}$$

and deduce that $\sum \dfrac{a_n}{s_n^2}$ converges.

4. What can be said about

$$\sum \frac{a_n}{1+na_n} \quad \text{and} \quad \sum \frac{a_n}{1+n^2a_n}?$$

Solution.

Throughout, the terms are positive, so each partial sum s_n is strictly increasing; and because $\sum a_n$ diverges, the increasing sequence (s_n) is unbounded, that is $s_n \to +\infty$ (6.4.4). In particular $1/s_n \to 0$ (Archimedean).

(a) The series $\sum \frac{a_n}{1+a_n}$ diverges. Suppose, toward a contradiction, that it converges. Then its terms vanish (6.6.3), so $\frac{a_n}{1+a_n} \to 0$; solving $b_n = \frac{a_n}{1+a_n}$ for $a_n = \frac{b_n}{1-b_n}$ shows $a_n \to 0$ as well. Hence $a_n < 1$ for all n beyond some N, and there $1 + a_n < 2$, so

$$\frac{a_n}{1+a_n} > \frac{a_n}{2} \qquad (n \geq N).$$

Both series have positive terms, so by the comparison test (6.6.8) the convergence of $\sum \frac{a_n}{1+a_n}$ forces the convergence of $\sum \frac{a_n}{2}$, hence of $\sum a_n$. This contradicts the hypothesis that $\sum a_n$ diverges, so $\sum \frac{a_n}{1+a_n}$ diverges.

(b) Fix N and $k \geq 1$. For each index n with $N+1 \leq n \leq N+k$ the monotonicity $s_n \leq s_{N+k}$ gives $\frac{a_n}{s_n} \geq \frac{a_n}{s_{N+k}}$. Summing over these k indices,

$$\sum_{j=1}^{k} \frac{a_{N+j}}{s_{N+j}} \;\geq\; \frac{1}{s_{N+k}} \sum_{j=1}^{k} a_{N+j} \;=\; \frac{s_{N+k} - s_N}{s_{N+k}} \;=\; 1 - \frac{s_N}{s_{N+k}},$$

where $\sum_{j=1}^{k} a_{N+j} = s_{N+k} - s_N$ is the definition of the partial sums. This is the claimed inequality.

To deduce divergence, fix N and let $k \to \infty$. Since $s_{N+k} \to +\infty$ while s_N is fixed, $\frac{s_N}{s_{N+k}} \to 0$, so the right-hand side tends to 1; in particular there is a k with $1 - \frac{s_N}{s_{N+k}} > \frac{1}{2}$, hence

$$\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} > \tfrac{1}{2}.$$

As this holds for *every* starting index N, the partial sums of $\sum \frac{a_n}{s_n}$ are not Cauchy: no tail block can be made uniformly small. By the Cauchy criterion for series (6.6.2), $\sum \frac{a_n}{s_n}$ diverges.

(c) For $n \geq 2$, the difference of reciprocals telescopes a single term:

$$\frac{1}{s_{n-1}} - \frac{1}{s_n} = \frac{s_n - s_{n-1}}{s_{n-1}\, s_n} = \frac{a_n}{s_{n-1}\, s_n}.$$

Since $0 < s_{n-1} < s_n$, we have $s_{n-1}s_n < s_n^2$, so $\frac{a_n}{s_{n-1}s_n} \geq \frac{a_n}{s_n^2}$, which is the claimed inequality

$$\frac{a_n}{s_n^2} \leq \frac{1}{s_{n-1}} - \frac{1}{s_n} \qquad (n \geq 2).$$

The right-hand side is a telescoping series with nonnegative terms, whose partial sums are

$$\sum_{n=2}^{m}\Big(\frac{1}{s_{n-1}} - \frac{1}{s_n}\Big) = \frac{1}{s_1} - \frac{1}{s_m} < \frac{1}{s_1} = \frac{1}{a_1},$$

bounded above by $1/a_1$. By comparison (6.6.8) the partial sums of the nonnegative series $\sum_{n\geq 2} \frac{a_n}{s_n^2}$ are likewise bounded, and a nonnegative series with bounded partial sums converges (6.6.6). Adjoining the single term a_1/s_1^2 does not affect convergence, so $\sum \frac{a_n}{s_n^2}$ converges. (The bound in fact gives $\sum_{n\geq 2} \frac{a_n}{s_n^2} \leq 1/a_1$.)

(d) The second series *always* converges; the first one can do either, so nothing can be said about it in general.

For $\sum \frac{a_n}{1+n^2a_n}$, drop the 1 in the denominator: for $n \geq 1$,

$$\frac{a_n}{1+n^2a_n} < \frac{a_n}{n^2a_n} = \frac{1}{n^2}.$$

The series $\sum 1/n^2$ converges (a p-series with $p = 2$, 6.6.5), so by comparison (6.6.8) the nonnegative series $\sum \frac{a_n}{1+n^2a_n}$ converges, whatever the divergent $\sum a_n$ may be.

For $\sum \frac{a_n}{1+na_n}$ no verdict is possible, as two divergent choices of $\sum a_n$ show.

- Taking $a_n = 1$ gives the divergent $\sum 1$, and $\frac{a_n}{1+na_n} = \frac{1}{1+n}$, so $\sum \frac{a_n}{1+na_n} = \sum \frac{1}{1+n}$ diverges (the harmonic series, 6.6.4).
- Taking
$$a_n = \begin{cases} 1, & n = k^2 \text{ for some integer } k \geq 1, \\ 1/n^2, & \text{otherwise,} \end{cases}$$
the perfect-square terms alone sum to $\sum_{k\geq 1} 1 = \infty$, so $\sum a_n$ diverges. But $\sum \frac{a_n}{1+na_n}$ converges: over the perfect squares $n = k^2$ the terms are $\frac{1}{1+k^2} < \frac{1}{k^2}$, summable in k; off the perfect squares $\frac{a_n}{1+na_n} < a_n = \frac{1}{n^2}$, summable as well. Both pieces converge (6.6.8, 6.6.5), so the whole series does.

Thus $\sum \frac{a_n}{1+na_n}$ may converge or diverge under the same hypotheses, and no general statement holds.

■

REFLECTIONS.

The fixed cast—positive terms a_n, partial sums $s_n = a_1 + \cdots + a_n$, and a *divergent* $\sum a_n$—hands you two facts to lean on before any part is read. Positivity makes (s_n) strictly increasing, and divergence of a positive-term series means exactly that those increasing partial sums run off to $+\infty$ (6.4.4); together they give $s_n \to +\infty$ and hence $1/s_n \to 0$ (Archimedean), the single quantitative input each later part quietly consumes. The four parts then ask the recurring Chapter 3 question—convergence or divergence—about series built by feeding a_n and s_n through various damping denominators, and the whole exercise is a tour of the standard tests: comparison (6.6.8), the Cauchy criterion (6.6.2), telescoping against the bounded-partial-sums criterion (6.6.6), and the p-series benchmark (6.6.5).

The unifying reading is that each denominator is a throttle on the term a_n, and the verdict turns on how hard it throttles relative to the divergence it is fighting. A throttle of bounded size ($1 + a_n$ once $a_n \to 0$) cannot kill the divergence; a throttle that grows like s_n slows it only logarithmically and still loses; a throttle that grows like s_n^2 finally wins. Believe this on the cleanest instance $a_n = 1$, where $s_n = n$: the three series become $\sum \frac{1}{2}$ (diverges), $\sum \frac{1}{n}$ (diverges, harmonically), and $\sum \frac{1}{n^2}$ (converges)—the harmonic-versus-$p = 2$ threshold in miniature, and the template the general argument upgrades.

The four deductions each rest on a named result:

(1) part (a) is a contradiction (0.5.5): a convergent $\sum \frac{a_n}{1+a_n}$ would force its terms to vanish (6.6.3), hence $a_n \to 0$, hence $1+a_n < 2$ eventually, so the comparison $\frac{a_n}{1+a_n} > \frac{a_n}{2}$ (6.6.8) would drag $\sum a_n$ into convergence—the inversion $a_n = \frac{b_n}{1-b_n}$ is the only nonobvious move, and it is what recovers $a_n \to 0$ from $\frac{a_n}{1+a_n} \to 0$;

(2) part (b) bounds each s_n in a block by the largest, s_{N+k}, so the block sum exceeds $\frac{s_{N+k}-s_N}{s_{N+k}} = 1 - \frac{s_N}{s_{N+k}}$; letting $k \to \infty$ with N fixed pushes this past $\frac{1}{2}$ because $s_{N+k} \to +\infty$, so a tail block stays large no matter where it starts—precisely the failure of the Cauchy criterion (6.6.2) that certifies divergence;

(3) part (c) telescopes: $\frac{1}{s_{n-1}} - \frac{1}{s_n} = \frac{a_n}{s_{n-1}s_n}$, and $s_{n-1} < s_n$ replaces $s_{n-1}s_n$ by the larger s_n^2 to give the inequality; the telescoping partial sums collapse to $\frac{1}{s_1} - \frac{1}{s_m} < \frac{1}{a_1}$, so the dominated nonnegative series has bounded partial sums and converges (6.6.6, 6.6.8);

(4) part (d) splits in two—$\sum \frac{a_n}{1+n^2a_n}$ is killed outright by dropping the 1 to get $< \frac{1}{n^2}$ and comparing to the convergent $p = 2$ series (6.6.5, 6.6.8), while $\sum \frac{a_n}{1+na_n}$ is answered by two examples (0.5.8), $a_n = 1$ for divergence and a sparse "spike" sequence for convergence.

The rigor lives in details easy to skate past. In (a) the case $a_n \not\to 0$ is not separate but *subsumed*: if the series converged its terms would vanish, so the proof never needs to suppose $a_n \to 0$ as a hypothesis—it derives it, then uses $a_n < 1$ only eventually, which is all comparison needs. In (b) the order of quantifiers is the whole point: the block can

be made large for *every* N, and that "for every N" is exactly the negation of the Cauchy criterion; proving it for one N would prove nothing. In (c) the strict inequality $s_{n-1} < s_n$ (positivity again) is what lets $s_{n-1}s_n < s_n^2$, and the telescoping bound $1/s_1$ is finite only because $1/s_m \to 0$, i.e. because $\sum a_n$ diverges—curiously, the *divergence* hypothesis is what makes this series converge *to* a clean bound. In (d), positivity is what licenses dropping the 1 to enlarge a denominator and shrink a term, and the convergent example must have $\sum a_n$ genuinely divergent—the perfect-square subsum $\sum 1$ supplies that, while everything off the squares is already $p = 2$-summable.

The transferable move is to read a denominator as a competition of growth rates: a damped positive series $\sum \frac{a_n}{d_n}$ converges or diverges according to whether d_n outgrows the divergence of $\sum a_n$, and the way to test it is to replace d_n by the dominant comparison—the bounded constant in (a), the block-maximal s_{N+k} in (b), the telescoping product $s_{n-1}s_n$ in (c), the clean n^2 in (d). The deepest of these is the telescoping trick of (c): writing a term as a difference $f(n-1) - f(n)$ turns an opaque series into a collapsing sum whose convergence you can read off the limit of f, the same idea that powers the integral-test intuition and the p-series threshold itself (6.6.5). And part (d) is the lasting caution—"$\sum a_n$ diverges" constrains the s_n^2-throttled series completely but says nothing about the na_n-throttled one, because a divergent series can hide its mass in a sparse subsequence that the linear weight na_n tames while the quadratic weight does not.

Problem 3.12 *Rudin, Chapter 3, Exercise 12*

Suppose $a_n > 0$ and $\sum a_n$ converges. Put

$$r_n = \sum_{m=n}^{\infty} a_m.$$

1. Prove that
$$\frac{a_m}{r_m} + \cdots + \frac{a_n}{r_n} > 1 - \frac{r_n}{r_m}$$
if $m < n$, and deduce that $\sum \frac{a_n}{r_n}$ diverges.

2. Prove that
$$\frac{a_n}{\sqrt{r_n}} < 2\left(\sqrt{r_n} - \sqrt{r_{n+1}}\right)$$
and deduce that $\sum \frac{a_n}{\sqrt{r_n}}$ converges.

SOLUTION.

Because $\sum a_n$ converges, each tail $r_n = \sum_{m=n}^{\infty} a_m$ is a finite real, and since every $a_m > 0$ each $r_n > 0$. The tails decrease strictly: $r_n - r_{n+1} = a_n > 0$, so

$$a_n = r_n - r_{n+1}, \qquad r_0 > r_1 > r_2 > \cdots > 0, \qquad r_n \to 0,$$

the last because r_n is the tail of a convergent series.

Part (a). Fix $m < n$. For each index k with $m \le k \le n$ the denominator satisfies $r_k \le r_m$ (the tails decrease), so $\dfrac{a_k}{r_k} \ge \dfrac{a_k}{r_m}$. Summing over $k = m, \dots, n$ and using $a_k = r_k - r_{k+1}$,

$$\sum_{k=m}^{n} \frac{a_k}{r_k} \;\ge\; \frac{1}{r_m} \sum_{k=m}^{n} (r_k - r_{k+1}) \;=\; \frac{r_m - r_{n+1}}{r_m} \;=\; 1 - \frac{r_{n+1}}{r_m},$$

the middle sum telescoping. Since $r_{n+1} < r_n$, we have $1 - \dfrac{r_{n+1}}{r_m} > 1 - \dfrac{r_n}{r_m}$, and therefore

$$\frac{a_m}{r_m} + \cdots + \frac{a_n}{r_n} > 1 - \frac{r_n}{r_m} \qquad (m < n).$$

This forces $\sum a_n/r_n$ to fail the Cauchy criterion (6.6.2). Fix m. As $n \to \infty$ the tail $r_n \to 0$, so $r_n/r_m \to 0$, and the block sum $\sum_{k=m}^{n} a_k/r_k > 1 - r_n/r_m$ exceeds $\frac{1}{2}$ for all large n. Thus for every m there is an $n > m$ with $\sum_{k=m}^{n} a_k/r_k > \frac{1}{2}$; the partial sums are not Cauchy, so $\sum a_n/r_n$ diverges.

Part (b). Factor $a_n = r_n - r_{n+1} = \big(\sqrt{r_n} - \sqrt{r_{n+1}}\big)\big(\sqrt{r_n} + \sqrt{r_{n+1}}\big)$, legitimate since $r_n, r_{n+1} > 0$. Dividing by $\sqrt{r_n}$,

$$\frac{a_n}{\sqrt{r_n}} = \big(\sqrt{r_n} - \sqrt{r_{n+1}}\big) \cdot \frac{\sqrt{r_n} + \sqrt{r_{n+1}}}{\sqrt{r_n}}.$$

From $r_{n+1} < r_n$ comes $\sqrt{r_{n+1}} < \sqrt{r_n}$, hence $\sqrt{r_n} + \sqrt{r_{n+1}} < 2\sqrt{r_n}$, so the fraction on the right is < 2. As $\sqrt{r_n} - \sqrt{r_{n+1}} > 0$, multiplying preserves the inequality and

$$\frac{a_n}{\sqrt{r_n}} < 2\big(\sqrt{r_n} - \sqrt{r_{n+1}}\big).$$

The right-hand side telescopes, which bounds the partial sums. For $N \le M$,

$$\sum_{n=N}^{M} 2\big(\sqrt{r_n} - \sqrt{r_{n+1}}\big) = 2\big(\sqrt{r_N} - \sqrt{r_{M+1}}\big) \le 2\sqrt{r_N},$$

since $r_{M+1} > 0$. Hence the partial sums of $\sum a_n/\sqrt{r_n}$, a series of positive terms, satisfy

$$\sum_{n=0}^{M} \frac{a_n}{\sqrt{r_n}} < \sum_{n=0}^{M} 2\big(\sqrt{r_n} - \sqrt{r_{n+1}}\big) = 2\big(\sqrt{r_0} - \sqrt{r_{M+1}}\big) < 2\sqrt{r_0},$$

bounded above by $2\sqrt{r_0}$. A series of nonnegative terms with bounded partial sums converges (6.6.6), so $\sum a_n/\sqrt{r_n}$ converges.

■

Reflections.

The data are a convergent series of positive terms and its tails $r_n = \sum_{m \geq n} a_m$, and the first move is to read what those tails are. The relation $a_n = r_n - r_{n+1}$ rewrites each term as a *difference* of consecutive tails, and a sum of consecutive differences is a telescoping sum (6.6.1)—the structural clue that governs both parts. Two facts about r_n are then forced and used throughout: the tails decrease strictly, because $r_n - r_{n+1} = a_n > 0$, and they vanish, $r_n \to 0$, precisely because r_n is the tail of a convergent series (6.6.2). The problem is a study in contrasts—the *same* weighting of a_n by a power of r_n diverges in (a) and converges in (b)—so each part ends not with a generic test but with a hand-built estimate that the tail structure makes available.

The two parts are deliberately paired on opposite shores: dividing a_n by r_n leaves too much mass and the series diverges, while dividing by the gentler $\sqrt{r_n}$ leaves little enough to converge. Believe the mechanism on the model $a_n = 2^{-n}$, where $r_n = 2^{-n+1}$: then $a_n/r_n = \frac{1}{2}$ is constant, so $\sum a_n/r_n$ is a sum of constants and diverges outright, while $a_n/\sqrt{r_n} = \frac{1}{2}\sqrt{r_n}$ shrinks geometrically and sums. Dividing by the full tail strips away exactly the decay that made $\sum a_n$ converge; taking a square root returns half of it.

Part (a) runs in three moves, each resting on the tail structure:

(1) over the block $m \leq k \leq n$ every denominator obeys $r_k \leq r_m$ because the tails decrease, so $a_k/r_k \geq a_k/r_m$—replacing each variable denominator by the single largest one r_m is what makes the block summable in closed form;

(2) the resulting sum $\frac{1}{r_m}\sum_{k=m}^{n}(r_k - r_{k+1})$ telescopes to $\frac{1}{r_m}(r_m - r_{n+1}) = 1 - r_{n+1}/r_m$, the one place $a_k = r_k - r_{k+1}$ is spent;

(3) $r_{n+1} < r_n$ weakens this to the stated $1 - r_n/r_m$, and now $r_n \to 0$ does the killing: holding m fixed and pushing $n \to \infty$ drives the block sum above $\frac{1}{2}$, so the partial sums are not Cauchy (6.6.2; a convergent series is Cauchy) and $\sum a_n/r_n$ diverges.

The verdict-then-justify shape matters: the Cauchy criterion is the right tool precisely because the divergence is not visible term by term—the terms a_n/r_n may individually tend to 0, yet *blocks* of them stay large, and only a criterion that watches blocks rather than single terms can see that.

Part (b) is the conjugate trick reincarnated. The term $a_n/\sqrt{r_n}$ resists direct comparison, but $a_n = r_n - r_{n+1}$ is a difference of squares of square roots, so factoring $a_n = (\sqrt{r_n} - \sqrt{r_{n+1}})(\sqrt{r_n} + \sqrt{r_{n+1}})$ exposes the very telescoping factor $\sqrt{r_n} - \sqrt{r_{n+1}}$ one wants to sum. The remaining factor $(\sqrt{r_n} + \sqrt{r_{n+1}})/\sqrt{r_n}$ is bounded by 2 for the same reason as in (a)—the tails decrease—and that single bound converts the awkward term into something whose sum collapses. The deduction is then a bounded-partial-sums argument (6.6.6): the majorant $\sum 2(\sqrt{r_n} - \sqrt{r_{n+1}})$ telescopes to $2(\sqrt{r_0} - \sqrt{r_{M+1}}) < 2\sqrt{r_0}$, a ceiling independent of M, so a positive series sitting beneath it must converge—this is the comparison test

(6.6.8) against a convergent telescoping series, with the lower bound $a_n/\sqrt{r_n} > 0$ supplied free by $a_n > 0$.

The argument is honest on each hypothesis. Positivity of the a_n is load-bearing twice over: it makes every $r_n > 0$ so that $\sqrt{r_n}$ and the divisions even parse, and it makes the tails strictly decreasing, which is the inequality $r_k \le r_m$ and $r_{n+1} < r_n$ that both parts run on; drop it and the tails need not be monotone and the estimates collapse. Convergence of $\sum a_n$ is what gives $r_n \to 0$, and that limit is the whole engine of (a)—without it the block sums need not approach 1 and divergence is lost. The reasoning is not circular: (a) never assumes $\sum a_n/r_n$ converges, it assumes the contrary and breaks the Cauchy condition, while (b) builds a convergent majorant by hand rather than borrowing the conclusion.

The transferable instinct is to let the tail r_n be the variable you compute with, not the terms a_n. Writing $a_n = r_n - r_{n+1}$ turns weighted sums of a_n into telescoping sums in r_n, and a monotone tail lets you replace a moving denominator by a fixed extreme one. The same difference-of-tails identity recurs whenever a series is reweighted by a function of its own tail, and the lesson of the pairing is quantitative: $\sum a_n/r_n^p$ inherits convergence from $\sum a_n$ when $p < 1$ and loses it when $p \ge 1$, with the conjugate factoring of (b) handling the borderline-beating exponent $p = \frac{1}{2}$.

Problem 3.13 *Rudin, Chapter 3, Exercise 13*

Prove that the Cauchy product of two absolutely convergent series converges absolutely.

Solution.

Let $\sum_{n\ge0} a_n$ and $\sum_{n\ge0} b_n$ be absolutely convergent series of real (or complex) numbers, and write

$$A = \sum_{n=0}^{\infty} |a_n| < \infty, \qquad B = \sum_{n=0}^{\infty} |b_n| < \infty$$

for the (finite) sums of their absolute values. Their Cauchy product is $\sum_{n\ge0} c_n$ with

$$c_n = \sum_{k=0}^{n} a_k\, b_{n-k}.$$

The claim is that $\sum c_n$ converges absolutely, i.e. that $\sum_n |c_n|$ converges.

Bound a single term by the triangle inequality:

$$|c_n| = \Big|\sum_{k=0}^{n} a_k b_{n-k}\Big| \le \sum_{k=0}^{n} |a_k|\,|b_{n-k}|.$$

Summing this over $n = 0, 1, \ldots, N$ gives, with every term nonnegative,

$$\sum_{n=0}^{N} |c_n| \le \sum_{n=0}^{N} \sum_{k=0}^{n} |a_k|\,|b_{n-k}| = \sum_{\substack{i,j\ge 0 \\ i+j\le N}} |a_i|\,|b_j|,$$

the last equality being the substitution $i = k$, $j = n - k$, which runs over exactly the pairs (i, j) of nonnegative integers with $i + j \le N$. Every such pair satisfies $i \le N$ and $j \le N$, so this triangle of indices sits inside the full square $\{(i, j) : 0 \le i, j \le N\}$; since all summands are nonnegative, enlarging the index set can only increase the sum:

$$\sum_{\substack{i,j\ge 0 \\ i+j\le N}} |a_i|\,|b_j| \le \sum_{i=0}^{N} \sum_{j=0}^{N} |a_i|\,|b_j| = \Big(\sum_{i=0}^{N} |a_i|\Big)\Big(\sum_{j=0}^{N} |b_j|\Big) \le A\,B.$$

Chaining the two displays,

$$\sum_{n=0}^{N} |c_n| \le A\,B \qquad \text{for every } N.$$

The partial sums of $\sum |c_n|$ are thus a nondecreasing sequence (the terms $|c_n| \ge 0$) bounded above by the constant $A\,B$. A nonnegative series whose partial sums are bounded converges (6.6.6)—its partial sums converge by monotone convergence (a bounded monotonic sequence converges; 6.2.2). Hence $\sum |c_n|$ converges, that is, the Cauchy product $\sum c_n$ converges absolutely.

■

REFLECTIONS.

The phrase that fixes everything is *converges absolutely*: the conclusion is not about the delicate value of $\sum c_n$ but only about $\sum |c_n|$, a series of *nonnegative* terms. That single observation steers the whole proof, because a nonnegative series has the friendliest convergence criterion in the notes—its partial sums climb monotonically, so it converges the instant they are bounded above (6.6.6), with no need to exhibit a limit or run an ε argument. The task therefore collapses to one inequality: produce a single constant that caps every partial sum of $\sum |c_n|$.

The word *Cauchy product* supplies the raw material, and reading its shape tells you where the cap comes from. The term $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ collects all products $a_i b_j$ whose indices sum to exactly n; summing $|c_n|$ up to N therefore gathers all products $|a_i|\,|b_j|$ with $i + j \le N$. Picture those pairs (i, j) as lattice points: they fill a triangle below the diagonal line $i + j = N$. The two factors $A = \sum |a_i|$ and $B = \sum |b_j|$, multiplied out, gather instead the points of a *square*—every pair with $i \le N$ and $j \le N$. The triangle sits inside the

square, so the triangle's total cannot exceed the square's, and the square's total is just AB. Believe it on the smallest case: with $N = 1$ the triangle is $\{(0,0),(1,0),(0,1)\}$ and the square adds the corner $(1,1)$, so $|a_0||b_0| + |a_1||b_0| + |a_0||b_1| \le (|a_0| + |a_1|)(|b_0| + |b_1|)$ with the extra term $|a_1||b_1| \ge 0$ accounting for the slack.

The argument is then a short chain, each link resting on the one before:

(1) the triangle inequality turns $|c_n|$ into the dominating sum $\sum_k |a_k|\,|b_{n-k}|$, replacing signed (or complex) products by their nonnegative absolute values—the step that converts the problem into one about a nonnegative series;

(2) reindexing $(k, n-k) \mapsto (i, j)$ recognises $\sum_{n\le N}\sum_k |a_k||b_{n-k}|$ as the sum over the triangle $\{i + j \le N\}$, the bookkeeping that exposes the Cauchy product's diagonal structure;

(3) the triangle lies inside the square $\{i, j \le N\}$, and *because the terms are nonnegative* adding the missing corner pairs only enlarges the sum—this monotonicity of finite sums in their index set is the load-bearing move;

(4) the square sum factors as $\big(\sum_{i\le N} |a_i|\big)\big(\sum_{j\le N} |b_j|\big)$, each factor bounded by the full absolute sum, giving the uniform cap AB;

(5) bounded monotone partial sums converge (6.6.6, by monotone convergence a bounded monotonic sequence converges, 6.2.2), so $\sum |c_n|$ converges.

The rigor turns entirely on nonnegativity, and it is worth seeing exactly where each hypothesis is spent. Both series are assumed *absolutely* convergent precisely so that A and B are finite; mere convergence is not enough, since the rearrangement and the triangle-inside-square comparison are licensed only for nonnegative terms, and the Cauchy product of two conditionally convergent series can genuinely diverge—the classic witness is $a_n = b_n = (-1)^n/\sqrt{n+1}$, where each $|c_n| \ge 1$ and the product fails the term test (6.6.3). The comparison "triangle $\subseteq$ square" would be false for signed terms, where adding the corner could decrease the sum; it is the absolute values that make every step monotone. Nothing here is circular: A and B are finite by hypothesis, the bound AB is built by hand, and convergence is read off the bounded monotone partial sums rather than assumed.

The transferable instinct is to meet "absolutely convergent" by working entirely with nonnegative series, where bounded partial sums *are* convergence (6.6.6)—the same reflex that powers the comparison test (6.6.8) and settles Problem 3.7. The geometric heart, that a sum over the diagonals $i + j = n$ is dominated by the sum over the product grid $i \le N$, $j \le N$, is the discrete shadow of Fubini's theorem: when all terms are nonnegative a double sum may be totalled in any order, and the Cauchy product is exactly the diagonal slicing of the grid $\sum_{i,j} a_i b_j = \big(\sum_i a_i\big)\big(\sum_j b_j\big)$. That is also why the value of the product is AB's signed cousin $\big(\sum a_n\big)\big(\sum b_n\big)$ once convergence is secured—absolute convergence is

what buys the right to reorder.

Problem 3.14 *Rudin, Chapter 3, Exercise 14*

If $\{s_n\}$ is a complex sequence, define its arithmetic means σ_n by

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_n}{n+1} \qquad (n = 0, 1, 2, \ldots).$$

1. If $\lim s_n = s$, prove that $\lim \sigma_n = s$.
2. Construct a sequence $\{s_n\}$ which does not converge, although $\lim \sigma_n = 0$.
3. Can it happen that $s_n > 0$ for all n and that $\limsup s_n = \infty$, although $\lim \sigma_n = 0$?
4. Put $a_n = s_n - s_{n-1}$, for $n \geq 1$. Show that

$$s_n - \sigma_n = \frac{1}{n+1} \sum_{k=1}^{n} k a_k.$$

   Assume that $\lim(na_n) = 0$ and that $\{\sigma_n\}$ converges. Prove that $\{s_n\}$ converges. [This gives a converse of (a), but under the additional assumption that $na_n \to 0$.]

5. Derive the last conclusion from a weaker hypothesis: Assume $M < \infty$, $|na_n| \leq M$ for all n, and $\lim \sigma_n = \sigma$. Prove that $\lim s_n = \sigma$, by completing the following outline:

   If $m < n$, then

$$s_n - \sigma_n = \frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{1}{n-m} \sum_{i=m+1}^{n} (s_n - s_i).$$

   For these i,

$$|s_n - s_i| \leq \frac{(n-i)M}{i+1} \leq \frac{(n-m-1)M}{m+2}.$$

   Fix $\varepsilon > 0$ and associate with each n the integer m that satisfies

$$m \leq \frac{n-\varepsilon}{1+\varepsilon} < m+1.$$

   Then $(m+1)/(n-m) \leq 1/\varepsilon$ and $|s_n - s_i| < M\varepsilon$. Hence

$$\limsup_{n\to\infty} |s_n - \sigma| \leq M\varepsilon.$$

   Since ε was arbitrary, $\lim s_n = \sigma$.

SOLUTION.

Write $\sigma_n = \frac{1}{n+1} \sum_{k=0}^{n} s_k$ throughout.

(a) Averaging preserves the limit. Suppose $s_n \to s$. Fix $\varepsilon > 0$. There is an index N with

$|s_k - s| < \varepsilon/2$ for every $k \geq N$ (5.1.1). For $n \geq N$, split the average about its limit:

$$\sigma_n - s = \frac{1}{n+1}\sum_{k=0}^{n}(s_k - s) = \frac{1}{n+1}\sum_{k=0}^{N-1}(s_k - s) + \frac{1}{n+1}\sum_{k=N}^{n}(s_k - s).$$

Let $C = \sum_{k=0}^{N-1} |s_k - s|$, a fixed number depending only on the frozen head. The tail terms each obey $|s_k - s| < \varepsilon/2$, and there are at most $n+1$ of them, so

$$|\sigma_n - s| \leq \frac{C}{n+1} + \frac{1}{n+1}\sum_{k=N}^{n}|s_k - s| < \frac{C}{n+1} + \frac{n-N+1}{n+1}\cdot\frac{\varepsilon}{2} \leq \frac{C}{n+1} + \frac{\varepsilon}{2}.$$

Since C is fixed, the Archimedean property (Archimedean property) gives an $N' \geq N$ with $C/(n+1) < \varepsilon/2$ for all $n \geq N'$. Then $|\sigma_n - s| < \varepsilon$ for $n \geq N'$, so $\sigma_n \to s$.

(b) Convergence of the means does not force convergence. Take $s_n = (-1)^n$. This sequence diverges, as it has the two distinct subsequential limits $+1$ and -1 (a limit is unique; 5.1.1). Its partial sums are

$$\sum_{k=0}^{n}(-1)^k = \begin{cases} 1, & n \text{ even}, \\ 0, & n \text{ odd}, \end{cases}$$

so $|s_0 + \cdots + s_n| \leq 1$ and $|\sigma_n| \leq \frac{1}{n+1} \to 0$. Hence $\sigma_n \to 0$ while $\{s_n\}$ does not converge.

(c) Yes—even with positive terms and an unbounded lim sup. The means vanish provided the spikes are sparse enough; the right tuning puts them at powers of 2. Define

$$s_n = \begin{cases} k, & n = 2^k \text{ for some integer } k \geq 1, \\ 2^{-n}, & \text{otherwise.} \end{cases}$$

Every term is strictly positive. The spike values $s_{2^k} = k$ form a subsequence tending to $+\infty$, so $\limsup_n s_n = \infty$ (6.4.5). For the means, fix n and split the partial sum into its non-spike part and its spikes. The non-spike terms are at most $\sum_{j=0}^{n} 2^{-j} < \sum_{j=0}^{\infty} 2^{-j} = 2$, a geometric series (6.6.7); the spikes present by index n are those with $2^k \leq n$, namely $k = 1, \ldots, L$ with $L = \lfloor \log_2 n \rfloor$, contributing $\sum_{k=1}^{L} k = \frac{L(L+1)}{2} \leq L^2$. Hence

$$0 < \sigma_n = \frac{1}{n+1}\sum_{k=0}^{n} s_k \leq \frac{2+L^2}{n+1}, \qquad L = \lfloor \log_2 n \rfloor \leq \log_2 n.$$

Now $(\log_2 n)^2$ grows far slower than $n+1$, so $\frac{2+L^2}{n+1} \to 0$ (the logarithm is dwarfed by any positive power, 6.5.1); by the squeeze (6.3.3) $\sigma_n \to 0$. Thus a sequence of strictly positive terms can have $\limsup s_n = \infty$ and yet $\lim \sigma_n = 0$: the spikes are real but so sparse that averaging drowns them.

(d) An identity, and a converse under $na_n \to 0$. With $a_k = s_k - s_{k-1}$ for $k \geq 1$, sum by

parts. Since $\sum_{k=1}^{n} k\,s_k - \sum_{k=1}^{n} k\,s_{k-1} = \sum_{k=1}^{n} k\,s_k - \sum_{j=0}^{n-1}(j+1)\,s_j = n\,s_n - \sum_{j=0}^{n-1} s_j$, we get

$$\sum_{k=1}^{n} k\,a_k = n\,s_n - \sum_{j=0}^{n-1} s_j = (n+1)s_n - \sum_{j=0}^{n} s_j = (n+1)(s_n - \sigma_n),$$

which rearranges to the claimed

$$s_n - \sigma_n = \frac{1}{n+1}\sum_{k=1}^{n} k\,a_k.$$

Now assume $na_n \to 0$ and $\sigma_n \to \sigma$. The numbers $b_k := k\,a_k$ satisfy $b_k \to 0$, so their own arithmetic means tend to 0 as well—this is precisely part (a) applied to $\{b_k\}$. But those means are

$$\frac{1}{n+1}\sum_{k=1}^{n} b_k = \frac{1}{n+1}\sum_{k=1}^{n} k\,a_k = s_n - \sigma_n$$

(the $k = 0$ term b_0 is absent, and dropping one fixed term changes a Cesàro average by $b_0/(n+1) \to 0$, so the conclusion of (a) is unaffected). Hence $s_n - \sigma_n \to 0$. Adding the convergent $\sigma_n \to \sigma$ (limits respect addition; 5.2.1) gives $s_n = (s_n - \sigma_n) + \sigma_n \to 0 + \sigma = \sigma$.

(e) The same conclusion from $|na_n| \le M$. Fix $m < n$. Averaging the identity of (d) over the block $i = m+1, \dots, n$ and isolating s_n yields the stated decomposition

$$s_n - \sigma_n = \frac{m+1}{n-m}\,(\sigma_n - \sigma_m) + \frac{1}{n-m}\sum_{i=m+1}^{n} (s_n - s_i). \tag{$*$}$$

To verify $(*)$, sum $s_i = \sigma_i + (s_i - \sigma_i)$ and use $\sum_{i=0}^{n} s_i = (n+1)\sigma_n$, $\sum_{i=0}^{m} s_i = (m+1)\sigma_m$: then $\sum_{i=m+1}^{n} s_i = (n+1)\sigma_n - (m+1)\sigma_m$, and

$$\frac{1}{n-m}\sum_{i=m+1}^{n} (s_n - s_i) = s_n - \frac{(n+1)\sigma_n - (m+1)\sigma_m}{n-m} = s_n - \sigma_n - \frac{m+1}{n-m}(\sigma_n - \sigma_m),$$

which is $(*)$ rearranged.

The block estimate. For $m < i \le n$, telescoping gives $s_n - s_i = \sum_{j=i+1}^{n} a_j$, and $|a_j| = |ja_j|/j \le M/j \le M/(i+1)$ for $j > i$, so

$$|s_n - s_i| \le \sum_{j=i+1}^{n} \frac{M}{j} \le (n-i)\cdot\frac{M}{i+1} = \frac{(n-i)M}{i+1} \le \frac{(n-m-1)M}{m+2},$$

the final step because $i \ge m+1$ makes the numerator $n - i \le n - m - 1$ no larger and the denominator $i + 1 \ge m + 2$ no smaller.

Choosing m. Fix $\varepsilon > 0$. For each n large enough that $n > \varepsilon$, pick the integer m with

$m \le \frac{n-\varepsilon}{1+\varepsilon} < m+1$; then $0 \le m < n$. From $m \le \frac{n-\varepsilon}{1+\varepsilon}$,

$$(m+1)(1+\varepsilon) \le m(1+\varepsilon) + (1+\varepsilon) \le (n-\varepsilon) + (1+\varepsilon) = n+1,$$

so $(m+1)\varepsilon \le n+1-(m+1) = n-m$, hence $\dfrac{m+1}{n-m} \le \dfrac{1}{\varepsilon}$. From $\frac{n-\varepsilon}{1+\varepsilon} < m+1$ we get $n-\varepsilon < (m+1)(1+\varepsilon)$, i.e. $n-m-1 < \varepsilon(m+2)$, hence $\dfrac{(n-m-1)M}{m+2} < M\varepsilon$. Combining with the block estimate, $|s_n - s_i| < M\varepsilon$ for every i in the block.

Passing to the limit. Apply these to $(*)$:

$$|s_n - \sigma_n| \le \frac{m+1}{n-m}\,|\sigma_n - \sigma_m| + \frac{1}{n-m}\sum_{i=m+1}^{n} |s_n - s_i| \le \frac{1}{\varepsilon}\,|\sigma_n - \sigma_m| + M\varepsilon.$$

As $n \to \infty$ the chosen $m = m(n) \to \infty$ as well (since $m+1 > \frac{n-\varepsilon}{1+\varepsilon} \to \infty$), and $\sigma_n \to \sigma$, $\sigma_m \to \sigma$, so $|\sigma_n - \sigma_m| \to 0$ (a convergent sequence is Cauchy; 5.4.2). Hence

$$\limsup_{n\to\infty} |s_n - \sigma_n| \le M\varepsilon, \qquad \text{so} \qquad \limsup_{n\to\infty} |s_n - \sigma| \le 0 + M\varepsilon = M\varepsilon,$$

where the equality used $s_n - \sigma = (s_n - \sigma_n) + (\sigma_n - \sigma)$ with $\sigma_n - \sigma \to 0$. Since $\varepsilon > 0$ was arbitrary, $\limsup_n |s_n - \sigma| = 0$, i.e. $\lim s_n = \sigma$.

■

Reflections.

The object the problem hands you is the running average σ_n, the *Cesàro mean* of $\{s_n\}$, and the verbs steering each part are different. Part (a) says "if $\lim s_n = s$, prove $\lim \sigma_n = s$"—a convergence claim to establish directly from the definition (5.1.1); part (b) says "construct," part (c) asks "can it happen," both demanding an explicit witness; and parts (d) and (e) ask for a partial *converse*—averages can converge without the sequence converging, by (b), so the converse can be true only with a brake on how fast s_n may move, which is exactly what $na_n \to 0$ and $|na_n| \le M$ supply.

The engine of part (a) is the head–tail split, the single most useful move for any limit attached to a growing average or sum. The reflex it answers is this: $s_k - s$ is small once $k \ge N$, but the first N terms can be anything, so quarantine them. Their total C is a *fixed* number, and a fixed number divided by $n+1$ is crushed by the Archimedean property (Archimedean property)—that is why the head is harmless. The tail, all of it within $\varepsilon/2$ of s, averages to within $\varepsilon/2$. Believe it on the slowest honest example, $s_n = 1/n$: the means are $\frac{1}{n+1}(1 + \frac{1}{2} + \cdots + \frac{1}{n}) \approx \frac{\ln n}{n} \to 0$, the limit preserved even though the averaging visibly drags.

Part (b) is the counterexample that makes the converse interesting, and it reads straight off (a)'s proof: averaging *smooths*, so a bounded oscillation whose partial sums stay bounded

must have means going to 0. The alternating $(-1)^n$ is the cleanest such signal—its partial sums never leave $\{0, 1\}$, so $|\sigma_n| \le 1/(n+1) \to 0$, while the sequence itself has two limits and so has none (a limit is unique). The lesson is that σ_n forgets oscillation.

Part (c) pushes that intuition to its limit: can the means vanish while the terms are positive and shoot to $+\infty$ along a subsequence? They can, provided the spikes are *sparse*, and the honest part of the answer is calibrating that sparsity. Spikes at the squares would still be too dense—heights $k = 1, \dots, L$ at $n = k^2$ (so $L \approx \sqrt{n}$) sum to $\frac{L(L+1)}{2} \approx n/2$, leaving $\sigma_n \to \frac{1}{2}$, a positive limit rather than 0—so spacing them geometrically, at 2^k with height k, makes their running total only $O((\log n)^2)$ against a denominator $n+1$. The squeeze (6.3.3), fed by the fact that a logarithm is dwarfed by any positive power (6.5.1), then sends σ_n to 0, while the geometric baseline 2^{-n} keeps the non-spike mass bounded by a convergent series (6.6.7). The transferable gauge is that a Cesàro mean sees a subsequence only in proportion to how densely its indices sit.

Part (d) is where the converse begins, and it turns on an exact identity proved by summation by parts:

(1) writing $a_k = s_k - s_{k-1}$ and reindexing $\sum k\, s_{k-1}$ collapses $\sum_{k=1}^{n} k\, a_k$ to $n\, s_n - \sum_{j=0}^{n-1} s_j = (n+1)(s_n - \sigma_n)$, which is the stated identity—the discrete analogue of integration by parts, with k playing the role of the slowly varying factor;

(2) reading the identity as $s_n - \sigma_n = \frac{1}{n+1} \sum_{k=1}^{n} b_k$ with $b_k = k\, a_k$ recognises the right side as the Cesàro mean of $\{b_k\}$, so the whole problem feeds on itself: $na_n \to 0$ means $b_k \to 0$, and part (a) (the limit 0 is preserved by averaging) forces $s_n - \sigma_n \to 0$;

(3) adding the assumed $\sigma_n \to \sigma$ back, by the algebra of limits (limits respect the operations; 5.2.1), recovers $s_n = (s_n - \sigma_n) + \sigma_n \to \sigma$.

The hypothesis $na_n \to 0$ is doing precise work: it is the rate condition that says successive terms $a_n = s_n - s_{n-1}$ shrink faster than $1/n$, just fast enough that their k-weighted average dies. Drop it and (b) already refutes the converse, so the condition is not decorative.

Part (e) trades the limit $na_n \to 0$ for the mere bound $|na_n| \le M$, and the gap is bridged by an identity over a *block* rather than the whole head. The decomposition $(*)$ writes $s_n - \sigma_n$ as a controlled multiple of $\sigma_n - \sigma_m$ plus an average of differences $s_n - s_i$ across a window $m < i \le n$. The bound $|na_n| \le M$ converts to $|a_j| \le M/j$, and telescoping over the window gives $|s_n - s_i| \le (n-i)M/(i+1)$: the terms cannot have drifted far if their increments are $O(1/j)$. The clever choice $m \approx \frac{n}{1+\varepsilon}$ tunes the window so that both error pieces are $O(\varepsilon)$ at once—the prefactor $\frac{m+1}{n-m} \le 1/\varepsilon$ keeps the first term in check once $\sigma_n - \sigma_m \to 0$, and the block bound becomes $< M\varepsilon$. Here is the one subtlety to guard: as $n \to \infty$ the chosen $m = m(n)$ also tends to ∞, so σ_m and σ_n chase the *same* limit and their difference vanishes—this is the Cauchy reading of a convergent sequence (convergent

sequences are Cauchy; 5.4.2), and it is where the assumption $\sigma_n \to \sigma$ is genuinely spent. The conclusion is then squeezed through $\limsup$ (6.4.5): $\limsup |s_n - \sigma| \le M\varepsilon$ for every $\varepsilon > 0$ forces the limsup to be 0, hence the limit to be σ.

What ties the five parts into one arc is the principle that averaging is a smoothing, information-losing operation: it always preserves a limit (a), it can manufacture one out of oscillation (b) or sparse blowup (c), and recovering the original sequence from its means is possible only with a quantitative leash on the increments—first $na_n \to 0$ (d), then the weaker $|na_n| \le M$ (e). These are the Cesàro and Tauberian theorems in miniature, and the reusable craftsmanship is the trio of moves: split a long average into a fixed head plus a uniformly small tail, read a weighted sum as a Cesàro mean of an auxiliary sequence, and when a single limit is awkward, control a sliding block whose width you tune to the error you are willing to tolerate.

Problem 3.15 *Rudin, Chapter 3, Exercise 15*

Definition 3.21 can be extended to the case in which the a_n lie in some fixed R^k. Absolute convergence is defined as convergence of $\sum |\mathbf{a}_n|$. Show that Theorems 3.22, 3.23, 3.25(a), 3.33, 3.34, 3.42, 3.45, 3.47, and 3.55 are true in this more general setting. (Only slight modifications are required in any of the proofs.)

Solution.

Throughout, $\mathbf{a}_n \in \mathbb{R}^k$, the partial sums are $\mathbf{s}_n = \sum_{j=0}^{n} \mathbf{a}_j$, and $|\cdot|$ is the Euclidean norm. Two facts about that norm are all the modifications need: it obeys the triangle inequality $|\mathbf{u}+\mathbf{v}| \le |\mathbf{u}| + |\mathbf{v}|$, and $\mathbb{R}^k$ is complete ($\mathbb{R}^k$ is complete; 5.5.3). Each theorem below is the real-line statement with the scalar absolute value read as the $\mathbb{R}^k$ norm; the proof in the notes is quoted and only the place where $|\cdot|$ enters is checked.

Theorem 3.22 (Cauchy criterion). $\sum \mathbf{a}_n$ converges iff for every $\varepsilon > 0$ there is N with $\left|\sum_{j=n}^{m} \mathbf{a}_j\right| \le \varepsilon$ whenever $m \ge n \ge N$. Convergence of the series *is* convergence of the sequence $(\mathbf{s}_n)$ in $\mathbb{R}^k$, and since $\mathbb{R}^k$ is complete that sequence converges iff it is Cauchy (5.4.3; Cauchy sequences); writing $\mathbf{s}_m - \mathbf{s}_{n-1} = \sum_{j=n}^{m} \mathbf{a}_j$ turns the Cauchy condition on $(\mathbf{s}_n)$ into the stated tail condition (6.6.2). Completeness of $\mathbb{R}^k$ is the only ingredient that must be invoked in place of completeness of $\mathbb{R}$.

Theorem 3.23 (terms vanish). If $\sum \mathbf{a}_n$ converges then $\mathbf{a}_n \to \mathbf{0}$. Take $m = n$ in 3.22: for $n \ge N$, $|\mathbf{a}_n| = \left|\sum_{j=n}^{n} \mathbf{a}_j\right| \le \varepsilon$, so $\mathbf{a}_n \to \mathbf{0}$ (6.6.3).

Theorem 3.25(a) (comparison). If $|\mathbf{a}_n| \le c_n$ for all $n \ge N_0$ and $\sum c_n$ converges, then $\sum \mathbf{a}_n$ converges. By the triangle inequality, for $m \ge n \ge N_0$,

$$\Big|\sum_{j=n}^{m} \mathbf{a}_j\Big| \le \sum_{j=n}^{m} |\mathbf{a}_j| \le \sum_{j=n}^{m} c_j.$$

Since $\sum c_n$ converges, its tails are $\le \varepsilon$ for $n \ge N$ by 3.22 applied to the real series, so the left side is $\le \varepsilon$ as well; the criterion 3.22 for the $\mathbb{R}^k$ series is met, and $\sum \mathbf{a}_n$ converges (6.6.8). The triangle inequality $|\sum \mathbf{a}_j| \le \sum |\mathbf{a}_j|$ is the one scalar step ($|\sum a_j| \le \sum |a_j|$) reread in $\mathbb{R}^k$.

Theorems 3.33 and 3.34 (root and ratio tests). If $\limsup_n |\mathbf{a}_n|^{1/n} < 1$, or if $|\mathbf{a}_{n+1}|/|\mathbf{a}_n| \le \beta < 1$ for all large n, then $\sum \mathbf{a}_n$ converges. Both tests are statements about the *scalar* series $\sum |\mathbf{a}_n|$ of nonnegative terms, to which the real root and ratio tests apply verbatim (6.6.9, 6.6.10): each produces a convergent geometric majorant $\sum c_n$ with $|\mathbf{a}_n| \le c_n$, and comparison 3.25(a) above then gives convergence of $\sum \mathbf{a}_n$. Divergence in the root test still follows from 3.23, since $\limsup |\mathbf{a}_n|^{1/n} > 1$ forces $|\mathbf{a}_n| \not\to 0$, hence $\mathbf{a}_n \not\to \mathbf{0}$.

Theorem 3.42 (Dirichlet's test). Let $\mathbf{a}_n \in \mathbb{R}^k$ have bounded partial sums $\mathbf{A}_n$, and let $b_0 \ge b_1 \ge \cdots$ be *real* with $b_n \to 0$; then $\sum b_n \mathbf{a}_n$ converges. Summation by parts is an algebraic identity valid coordinate by coordinate, so with $\mathbf{A}_{-1} = \mathbf{0}$,

$$\sum_{j=n}^{m} b_j \mathbf{a}_j = \sum_{j=n}^{m-1} (b_j - b_{j+1}) \mathbf{A}_j + b_m \mathbf{A}_m - b_n \mathbf{A}_{n-1}.$$

With $|\mathbf{A}_j| \le M$ for all j and $b_j - b_{j+1} \ge 0$, the triangle inequality gives

$$\Big|\sum_{j=n}^{m} b_j \mathbf{a}_j\Big| \le M\Big(\sum_{j=n}^{m-1} (b_j - b_{j+1}) + b_m + b_n\Big) = 2M b_n,$$

the sum telescoping. Since $b_n \to 0$ this is $\le \varepsilon$ for n large, so 3.22 yields convergence. Only the bound on $\mathbf{A}_j$ and the triangle inequality changed; the weights b_j stay real, which is what keeps the scalar multiples $b_j \mathbf{a}_j$ in $\mathbb{R}^k$.

Theorem 3.45 (absolute convergence implies convergence). If $\sum |\mathbf{a}_n|$ converges then $\sum \mathbf{a}_n$ converges. This is the comparison 3.25(a) with $c_n = |\mathbf{a}_n|$, since $|\mathbf{a}_n| \le c_n$ trivially and $\sum c_n = \sum |\mathbf{a}_n|$ converges by hypothesis.

Theorem 3.47 (linearity). If $\sum \mathbf{a}_n = \mathbf{A}$ and $\sum \mathbf{b}_n = \mathbf{B}$, then $\sum(\mathbf{a}_n + \mathbf{b}_n) = \mathbf{A} + \mathbf{B}$ and $\sum c\,\mathbf{a}_n = c\,\mathbf{A}$ for $c \in \mathbb{R}$. The partial sums satisfy $\sum_{j \le n}(\mathbf{a}_j + \mathbf{b}_j) = \mathbf{s}_n + \mathbf{t}_n$ and $\sum_{j \le n} c\,\mathbf{a}_j = c\,\mathbf{s}_n$, and limits in $\mathbb{R}^k$ respect sums and scalar multiples (the algebra of limits; 5.2.1), so these tend to $\mathbf{A} + \mathbf{B}$ and $c\,\mathbf{A}$.

Theorem 3.55 (rearrangement). If $\sum \mathbf{a}_n$ converges absolutely, every rearrangement $\sum \mathbf{a}_{\sigma(n)}$ converges to the same sum. The proof in the notes uses only the Cauchy tail of $\sum |\mathbf{a}_n|$ and the triangle inequality. Given $\varepsilon > 0$, absolute convergence and 3.22 applied to the real series $\sum |\mathbf{a}_n|$ give N with $\sum_{j>N} |\mathbf{a}_j| \le \varepsilon$. Choose p so large that $\{0, \dots, N\} \subseteq \{\sigma(0), \dots, \sigma(p)\}$;

then for the rearranged partial sums $\mathbf{s}'_n$ and any $q \geq p$, the indices $0, \ldots, N$ cancel and

$$|\mathbf{s}'_q - \mathbf{s}_q| \leq \sum_{j>N} |\mathbf{a}_j| \leq \varepsilon,$$

where $\mathbf{A} = \sum \mathbf{a}_n$ exists because absolute convergence forces convergence (3.45 above). Thus $\mathbf{s}'_q - \mathbf{s}_q \to \mathbf{0}$, and since $\mathbf{s}_q \to \mathbf{A}$ this gives $\mathbf{s}'_q \to \mathbf{A}$: the rearrangement converges to the same sum. The only $\mathbb{R}^k$ inputs are the norm reading of $|\mathbf{s}'_q - \mathbf{s}_q| \leq \sum_{j>N} |\mathbf{a}_j|$ and the existence of $\mathbf{A}$, both already secured above.

In every case the only edits are to read the absolute value as the $\mathbb{R}^k$ norm, to invoke completeness of $\mathbb{R}^k$ rather than of $\mathbb{R}$ in the Cauchy criterion, and to use the triangle inequality $\left|\sum \mathbf{a}_j\right| \leq \sum |\mathbf{a}_j|$ in place of its scalar twin. The two tests and absolute convergence reduce outright to the real series $\sum |\mathbf{a}_n|$.

■

Reflections.

The instruction *only slight modifications are required* is the whole exercise stated as a hint: the task is not to rebuild nine proofs but to locate, in each, the single line where the real number a_n was treated as a scalar and ask whether the same line survives when a_n becomes a vector $\mathbf{a}_n \in \mathbb{R}^k$. That reframing tells you to read each theorem looking for two things only, because those are the two places a scalar proof can secretly use that it lives on the line: where it manipulates $|a_n|$ as an absolute value, and where it concludes a Cauchy sequence converges. The first is governed by the triangle inequality, which the Euclidean norm shares; the second by completeness, which $\mathbb{R}^k$ shares ($\mathbb{R}^k$ is complete; 5.5.3). Pin those two down and the rest is transcription.

It is worth believing the principle on the smallest case before trusting it nine times. A series $\sum \mathbf{a}_n$ in $\mathbb{R}^k$ is nothing but its sequence of partial sums $\mathbf{s}_n = \sum_{j \leq n} \mathbf{a}_j$, and that sequence converges iff each of its k coordinate sequences does (coordinatewise convergence; 5.2.2)—so a vector series is, under the hood, k ordinary real series running in parallel. One could prove every part coordinatewise. The reason the norm-and-completeness route is cleaner is that it never has to split into coordinates: the single inequality $\left|\sum \mathbf{a}_j\right| \leq \sum |\mathbf{a}_j|$ does in one stroke what k separate triangle inequalities would do, which is exactly why Rudin can call the modifications slight.

The nine results sort into a small number of mechanisms, and walking them in dependency order shows that only the first two carry any genuine $\mathbb{R}^k$ content:

(1) Theorem 3.22 is the load-bearing one, because convergence of a series is by definition convergence of $(\mathbf{s}_n)$, and the only theorem that lets you certify that convergence without naming the limit is "Cauchy ⇒ convergent" (5.4.3, 6.6.2)—which holds in $\mathbb{R}^k$

precisely because $\mathbb{R}^k$ is complete (Cauchy sequences in a complete space); this is the one spot where completeness of $\mathbb{R}$ is swapped for completeness of $\mathbb{R}^k$;

(2) Theorem 3.25(a) is the second engine, and its only vector step is the triangle inequality $|\sum_{j=n}^{m} \mathbf{a}_j| \leq \sum_{j=n}^{m} |\mathbf{a}_j|$, the norm reading of the scalar $|\sum a_j| \leq \sum |a_j|$; with that in hand the convergent real majorant $\sum c_n$ forces the tail of the vector series small, and 3.22 closes it (6.6.8);

(3) everything else is downstream. Theorem 3.23 is 3.22 with $m = n$; the root and ratio tests (6.6.9, 6.6.10) are theorems about the *nonnegative real* series $\sum |\mathbf{a}_n|$ that build a geometric majorant and then call 3.25(a); Theorem 3.45 is 3.25(a) with $c_n = |\mathbf{a}_n|$; and Theorem 3.55 reuses the absolute-convergence tail of $\sum |\mathbf{a}_n|$ with one triangle inequality on the rearranged difference;

(4) two parts use a different but equally portable tool. Theorem 3.47 is pure algebra of limits on partial sums (the algebra of limits; 5.2.1), which already holds in $\mathbb{R}^k$; and Theorem 3.42 leans on summation by parts, an identity that holds coordinate by coordinate, after which the bounded partial sums $|\mathbf{A}_j| \leq M$ and the triangle inequality telescope the estimate to $2Mb_n \to 0$.

The rigor check is to see which hypotheses are genuinely doing the work and which were never about the line at all. Completeness is not optional: it is exactly the property that fails in $\mathbb{Q}$, where a series of rationals can be Cauchy yet sum to an irrational, so 3.22 would be false—the slight modification "$\mathbb{R} \to \mathbb{R}^k$" is licensed only because $\mathbb{R}^k$ keeps completeness (complete). The triangle inequality is the other non-negotiable, and it is the *only* thing the norm needs to supply; nowhere does any proof multiply two vectors or order them, so the absence of a product or an order on $\mathbb{R}^k$ never bites. Dirichlet's test makes that visible: the weights b_n must stay *real* and monotone, since "$b_0 \geq b_1 \geq \cdots$" has no meaning for vectors and the scalar multiples $b_n\mathbf{a}_n$ are what keep the terms in $\mathbb{R}^k$. And the reduction of the root and ratio tests is honest rather than circular, because $\sum |\mathbf{a}_n|$ is a bona fide real series of nonnegative terms to which the unmodified real theorems apply; the vector series is reached only afterwards, through comparison.

The transferable lesson is what "slight modification" really certifies: a proof generalises exactly as far as the structure it actually uses. These nine arguments use only that the codomain is a complete normed space, so they hold verbatim in $\mathbb{R}^k$—and, for the same reason, in any Banach space, which is the setting they secretly belonged to all along. The two that quietly need more, Dirichlet's test with its monotone weights and any future result that multiplies or compares terms, are precisely the ones that keep a foot on the real line; spotting which is which, by auditing each proof for the properties of $|\cdot|$ it leans on, is the skill the exercise is training.

Problem 3.16 *Rudin, Chapter 3, Exercise 16*

Fix a positive number α. Choose $x_1 > \sqrt{\alpha}$, and define $x_2, x_3, x_4, \ldots$, by the recursion formula

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{\alpha}{x_n}\right).$$

1. Prove that $\{x_n\}$ decreases monotonically and that $\lim x_n = \sqrt{\alpha}$.

2. Put $\varepsilon_n = x_n - \sqrt{\alpha}$, and show that

$$\varepsilon_{n+1} = \frac{\varepsilon_n^2}{2x_n} < \frac{\varepsilon_n^2}{2\sqrt{\alpha}}$$

so that, setting $\beta = 2\sqrt{\alpha}$,

$$\varepsilon_{n+1} < \beta\left(\frac{\varepsilon_1}{\beta}\right)^{2^n} \qquad (n = 1, 2, 3, \ldots).$$

3. This is a good algorithm for computing square roots, since the recursion formula is simple and the convergence is extremely rapid. For example, if $\alpha = 3$ and $x_1 = 2$, show that $\varepsilon_1/\beta < \frac{1}{10}$ and that therefore

$$\varepsilon_5 < 4 \cdot 10^{-16}, \qquad \varepsilon_6 < 4 \cdot 10^{-32}.$$

Solution.

Throughout, $\sqrt{\alpha} > 0$ is the positive square root of $\alpha > 0$.

Part (a). First, every term is positive: $x_1 > \sqrt{\alpha} > 0$, and if $x_n > 0$ then $x_{n+1} = \frac{1}{2}(x_n + \alpha/x_n)$ is a sum of positive numbers divided by 2, hence positive. So $x_n > 0$ for all n by induction (0.6.2), and the recursion never divides by zero.

Next, $x_n > \sqrt{\alpha}$ for every n. This holds at $n = 1$ by hypothesis. The identity

$$x_{n+1} - \sqrt{\alpha} = \frac{1}{2}\Big(x_n + \frac{\alpha}{x_n}\Big) - \sqrt{\alpha} = \frac{x_n^2 - 2\sqrt{\alpha}\,x_n + \alpha}{2x_n} = \frac{(x_n - \sqrt{\alpha})^2}{2x_n}$$

shows the right-hand side is ≥ 0, with equality only when $x_n = \sqrt{\alpha}$. So if $x_n > \sqrt{\alpha}$ then $(x_n - \sqrt{\alpha})^2 > 0$ and $x_{n+1} > \sqrt{\alpha}$; by induction $x_n > \sqrt{\alpha}$ for all n.

The sequence decreases. From $x_n > \sqrt{\alpha}$ we get $x_n^2 > \alpha$, so

$$x_n - x_{n+1} = x_n - \frac{1}{2}\Big(x_n + \frac{\alpha}{x_n}\Big) = \frac{x_n^2 - \alpha}{2x_n} > 0,$$

i.e. $x_{n+1} < x_n$. Thus $\{x_n\}$ is monotonically decreasing (6.2.1) and bounded below by $\sqrt{\alpha}$, so it converges (a bounded monotone sequence converges; 6.2.2); call the limit L. Since

$x_n > \sqrt{\alpha}$ for all n, the order survives in the limit (6.3.1), giving $L \geq \sqrt{\alpha} > 0$. Passing to the limit in $x_{n+1} = \frac{1}{2}(x_n + \alpha/x_n)$ through the algebra of limits (limits respect the algebraic operations; 5.2.1)—legitimate because $L \neq 0$—yields $L = \frac{1}{2}(L + \alpha/L)$, whence $2L^2 = L^2 + \alpha$, so $L^2 = \alpha$. As $L > 0$, $L = \sqrt{\alpha}$. Hence $\lim x_n = \sqrt{\alpha}$.

Part (b). Put $\varepsilon_n = x_n - \sqrt{\alpha}$, so $\varepsilon_n > 0$ by part (a). The identity computed above is exactly

$$\varepsilon_{n+1} = x_{n+1} - \sqrt{\alpha} = \frac{(x_n - \sqrt{\alpha})^2}{2x_n} = \frac{\varepsilon_n^2}{2x_n}.$$

Since $x_n > \sqrt{\alpha}$, the denominator obeys $2x_n > 2\sqrt{\alpha}$, and as $\varepsilon_n^2 > 0$ this gives

$$\varepsilon_{n+1} = \frac{\varepsilon_n^2}{2x_n} < \frac{\varepsilon_n^2}{2\sqrt{\alpha}}.$$

Write $\beta = 2\sqrt{\alpha} > 0$. Dividing the last inequality by β rewrites it as

$$\frac{\varepsilon_{n+1}}{\beta} < \Big(\frac{\varepsilon_n}{\beta}\Big)^2,$$

since $\varepsilon_n^2/(2\sqrt{\alpha}) = \beta\,(\varepsilon_n/\beta)^2$. Set $t_n = \varepsilon_n/\beta > 0$, so $t_{n+1} < t_n^2$. We claim $t_{n+1} < t_1^{2^n}$ for $n \geq 1$, by induction (1.1.3). The base case $n = 1$ is $t_2 < t_1^2 = t_1^{2^1}$. If $t_{n+1} < t_1^{2^n}$ for some $n \geq 1$, then, both sides being positive, squaring preserves the inequality, and

$$t_{n+2} < t_{n+1}^2 < \big(t_1^{2^n}\big)^2 = t_1^{2^{n+1}}.$$

This is the claim at $n + 1$. Multiplying $t_{n+1} < t_1^{2^n}$ by β restores

$$\varepsilon_{n+1} < \beta\Big(\frac{\varepsilon_1}{\beta}\Big)^{2^n} \qquad (n = 1, 2, 3, \dots).$$

Part (c). Take $\alpha = 3$ and $x_1 = 2$, so $\beta = 2\sqrt{3}$ and $\varepsilon_1 = x_1 - \sqrt{3} = 2 - \sqrt{3}$. The inequality $\varepsilon_1/\beta < \frac{1}{10}$ is, after clearing the positive denominator $\beta = 2\sqrt{3}$,

$$10(2 - \sqrt{3}) < 2\sqrt{3} \iff 20 < 12\sqrt{3} \iff \sqrt{3} > \tfrac{5}{3} \iff 3 > \tfrac{25}{9},$$

and $3 = \frac{27}{9} > \frac{25}{9}$ holds, so $\varepsilon_1/\beta < \frac{1}{10}$.

Now use part (b). Because $0 < \varepsilon_1/\beta < \frac{1}{10}$ and $2^n \geq 1$, raising to the power 2^n preserves the inequality, $(\varepsilon_1/\beta)^{2^n} < (\frac{1}{10})^{2^n} = 10^{-2^n}$, so

$$\varepsilon_{n+1} < \beta\Big(\frac{\varepsilon_1}{\beta}\Big)^{2^n} < \beta \cdot 10^{-2^n} = 2\sqrt{3} \cdot 10^{-2^n}.$$

Since $2\sqrt{3} = \sqrt{12} < \sqrt{16} = 4$, we have $\beta < 4$. Taking $n = 4$ gives $2^4 = 16$ and

$$\varepsilon_5 < \beta \cdot 10^{-16} < 4 \cdot 10^{-16};$$

taking $n = 5$ gives $2^5 = 32$ and

$$\varepsilon_6 < \beta \cdot 10^{-32} < 4 \cdot 10^{-32}.$$

■

Reflections.

The decisive word in part (a) is *decreases monotonically*: a sequence asked to be monotone and to have a limit is the exact shape the monotone convergence theorem was built for (a bounded monotone sequence converges; 6.2.2), so before touching the algebra one should aim to show the terms fall while staying above a floor. That theorem is the only tool in our notes that delivers a limit from a recursion you cannot solve in closed form, because it never asks you to name the limit in advance—it certifies existence from shape alone (6.2.1). What makes the floor obvious is the form of the recursion: $\frac{1}{2}(x+\alpha/x)$ is an average of x and α/x, two numbers whose product is α, so the average can never dip below their geometric mean $\sqrt{\alpha}$—the statement quietly hands you the lower bound by choosing $x_1 > \sqrt{\alpha}$.

It is worth believing this on the intended case before proving it. With $\alpha = 3$, $x_1 = 2$ the iteration reads $x_2 = \frac{1}{2}(2 + \frac{3}{2}) = 1.75$, then $x_3 = \frac{1}{2}(1.75 + 3/1.75) \approx 1.732143$, and $\sqrt{3} \approx 1.7320508$; the terms drop toward $\sqrt{3}$ from above and the agreement leaps from one decimal to several in a single step. That last observation—the explosive accuracy—is the whole point of parts (b) and (c).

The proof of (a) runs as a chain of inductions feeding the convergence theorem, each link resting on the one before:

(1) positivity of every x_n comes first by induction (0.6.2), since it is what licenses dividing by x_n and what keeps the later square-root manipulations honest;

(2) the single identity $x_{n+1} - \sqrt{\alpha} = (x_n - \sqrt{\alpha})^2/(2x_n)$, got by completing the square in the numerator, does double duty: a square over a positive number is ≥ 0, so $x_n > \sqrt{\alpha}$ propagates—this is the lower bound, proved by induction rather than quoted;

(3) monotonicity is then a one-line sign check, $x_n - x_{n+1} = (x_n^2 - \alpha)/(2x_n) > 0$, which is positive precisely because step (2) put x_n above $\sqrt{\alpha}$ (6.2.1);

(4) bounded below and decreasing, the sequence converges to some L (monotone convergence; 6.2.2), and the order $x_n > \sqrt{\alpha}$ passes to $L \geq \sqrt{\alpha} > 0$ in the limit (6.3.1);

(5) identifying L uses the standard fixed-point move: the limit of x_{n+1} equals the limit of x_n, so applying the algebra of limits (limits respect the operations; 5.2.1) to both sides of the recursion gives $L = \frac{1}{2}(L + \alpha/L)$, hence $L^2 = \alpha$ and $L = \sqrt{\alpha}$.

Step (5) is where rigour is easy to lose, and two guards are load-bearing. The algebra of limits may be applied to α/x_n only because $L \neq 0$, which is why step (4) bothered to record $L \geq \sqrt{\alpha} > 0$ rather than the weaker $L \geq \sqrt{\alpha}$; and the equation $L^2 = \alpha$ has two roots, so the sign $L > 0$ is needed to select $\sqrt{\alpha}$ over $-\sqrt{\alpha}$. The hypothesis $x_1 > \sqrt{\alpha}$ also earns its keep beyond seeding the bound: starting *at* $\sqrt{\alpha}$ would make the sequence constant, and starting below it (but positive) lands the very next term above $\sqrt{\alpha}$, so the strict inequality is what guarantees a genuine, strictly decreasing approach rather than a degenerate one.

Part (b) is the reason the method is celebrated, and its clue is the substitution $\varepsilon_n = x_n - \sqrt{\alpha}$: naming the error turns the recursion into a statement about how the error shrinks, and the identity from step (2) is already exactly $\varepsilon_{n+1} = \varepsilon_n^2/(2x_n)$. The error at each stage is the *square* of the previous error, scaled—this is quadratic convergence, and once the ratio ε_n/β drops below 1 the exponents 2^n make it plummet. The displayed bound is then a clean induction (1.1.3) on $t_n = \varepsilon_n/\beta$, where $t_{n+1} < t_n^2$ unwinds to $t_{n+1} < t_1^{2^n}$; squaring is order-preserving only because every $t_n > 0$, which is again the positivity of the errors from part (a).

Part (c) is purely a matter of feeding numbers through that bound, and the only subtlety is doing the two estimates without a calculator: $\varepsilon_1/\beta < \frac{1}{10}$ reduces by clearing denominators to $3 > \frac{25}{9}$, and $\beta = 2\sqrt{3} = \sqrt{12} < 4$ since $12 < 16$. With $\varepsilon_1/\beta < \frac{1}{10}$ the bound gives $\varepsilon_{n+1} < \beta \cdot 10^{-2^n}$, and reading off $n = 4$ and $n = 5$ produces the double-exponential decay 10^{-16} then 10^{-32}—each step roughly doubling the count of correct digits, the hallmark Rudin is advertising. The transferable move is the whole arc: when a recursion has a fixed point you cannot solve for, prove monotonicity and a bound to get convergence for free (monotone convergence), then pass to the limit in the recursion to identify it, and finally track the *error* rather than the term to read off the speed—an error that recurs as its own square is the signature of the Newton-type iteration this problem is.

Problem 3.17 *Rudin, Chapter 3, Exercise 17*

Fix $\alpha > 1$. Take $x_1 > \sqrt{\alpha}$, and define

$$x_{n+1} = \frac{\alpha + x_n}{1 + x_n} = x_n + \frac{\alpha - x_n^2}{1 + x_n}.$$

(a) Prove that $x_1 > x_3 > x_5 > \cdots$.

(b) Prove that $x_2 < x_4 < x_6 < \cdots$.

(c) Prove that $\lim x_n = \sqrt{\alpha}$.

(d) Compare the rapidity of convergence of this process with the one described in Exercise 16.

Solution.

Write $s = \sqrt{\alpha}$, so $s > 1$ and $s^2 = \alpha$. Two facts about the iteration come first.

Every term is positive. The start $x_1 > s > 0$ is positive, and if $x_n > 0$ then both $\alpha + x_n > 0$ and $1 + x_n > 0$, so $x_{n+1} = (\alpha + x_n)/(1 + x_n) > 0$. By induction $x_n > 0$ for all n.

The error obeys a single linear recursion. Put $\varepsilon_n = x_n - s$. Subtracting s from the recursion and clearing the denominator,

$$\varepsilon_{n+1} = x_{n+1} - s = \frac{\alpha + x_n}{1 + x_n} - s = \frac{(\alpha - s) + (1 - s)x_n}{1 + x_n} = \frac{s(s-1) - (s-1)x_n}{1 + x_n} = -\frac{s-1}{1 + x_n}\,\varepsilon_n,$$

where the last step used $\alpha - s = s(s-1)$ and factored out $-(s-1)$. Set

$$c_n = \frac{s-1}{1 + x_n}, \qquad \text{so} \qquad \varepsilon_{n+1} = -\,c_n\,\varepsilon_n.$$

Since $s > 1$ and $x_n > 0$, each $c_n > 0$. The sign and the size of the error are read straight off $\varepsilon_{n+1} = -c_n\varepsilon_n$:

$$\operatorname{sign}\varepsilon_{n+1} = -\operatorname{sign}\varepsilon_n, \qquad |\varepsilon_{n+1}| = c_n\,|\varepsilon_n|.$$

Because $\varepsilon_1 = x_1 - s > 0$, the signs alternate: $\varepsilon_n > 0$ for odd n and $\varepsilon_n < 0$ for even n. Equivalently, $x_n > s$ for odd n and $x_n < s$ for even n.

A clean two-step contraction governs the magnitudes. Multiplying out the denominators,

$$(1 + x_n)(1 + x_{n+1}) = (1 + x_n) + (\alpha + x_n) = 1 + \alpha + 2x_n,$$

using $1 + x_{n+1} = 1 + \dfrac{\alpha + x_n}{1 + x_n} = \dfrac{(1 + x_n) + (\alpha + x_n)}{1 + x_n}$. Hence

$$|\varepsilon_{n+2}| = c_{n+1}c_n\,|\varepsilon_n|, \qquad c_{n+1}c_n = \frac{(s-1)^2}{(1 + x_n)(1 + x_{n+1})} = \frac{(s-1)^2}{1 + \alpha + 2x_n}.$$

Since $x_n > 0$, the denominator exceeds $1 + \alpha = 1 + s^2$, and $(s-1)^2 = s^2 - 2s + 1 < s^2 + 1$ because $s > 0$; therefore

$$c_{n+1}c_n \;<\; \frac{(s-1)^2}{1 + s^2} \;=:\; r \;<\; 1,$$

a constant independent of n. Each two consecutive steps shrink the error by the fixed factor $r < 1$.

(a) and (b). For odd n, both ε_n and ε_{n+2} are positive while $|\varepsilon_{n+2}| = c_{n+1}c_n|\varepsilon_n| < |\varepsilon_n|$, so $0 < \varepsilon_{n+2} < \varepsilon_n$, i.e. $x_{n+2} < x_n$; this is $x_1 > x_3 > x_5 > \cdots$. For even n, both errors are

negative with $|\varepsilon_{n+2}| < |\varepsilon_n|$, so $\varepsilon_n < \varepsilon_{n+2} < 0$, i.e. $x_n < x_{n+2}$; this is $x_2 < x_4 < x_6 < \cdots$.

(c) Iterating $|\varepsilon_{n+2}| < r\,|\varepsilon_n|$ down to the first term of each parity gives, for $k \geq 1$,

$$|\varepsilon_{2k+1}| < r^k\,|\varepsilon_1|, \qquad |\varepsilon_{2k}| < r^{k-1}\,|\varepsilon_2|.$$

With $0 < r < 1$ one has $r^k \to 0$ (a standard limit, 6.5.1), so $|\varepsilon_n| \to 0$ along both parities, hence along the whole sequence: given any $\delta > 0$, all but finitely many odd terms and all but finitely many even terms satisfy $|\varepsilon_n| < \delta$, so all but finitely many terms do. Thus $\varepsilon_n = x_n - s \to 0$, that is $\lim x_n = s = \sqrt{\alpha}$.

The same conclusion follows from monotone convergence, which is worth recording. By (a) and (b) the odd-indexed subsequence decreases and is bounded below by s, while the even-indexed subsequence increases and is bounded above by s; each is bounded and monotone, so each converges (a bounded monotone sequence converges; 6.2.2). A limit L of either satisfies $L = (\alpha + L)/(1 + L)$, i.e. $L^2 = \alpha$, and $L \geq 0$ as a limit of positive terms (6.3.1), so $L = s$; both subsequences reach the same value s, and together they exhaust the indices.

(d) The errors here contract *linearly* (geometrically): from $\varepsilon_{n+1} = -c_n\varepsilon_n$ with $c_n = \dfrac{s-1}{1+x_n} \to \dfrac{s-1}{1+s}$ as $x_n \to s$, the ratio $|\varepsilon_{n+1}|/|\varepsilon_n|$ tends to the fixed constant $\dfrac{s-1}{1+s} \in (0,1)$, so each step multiplies the error by roughly this factor and the number of correct digits grows by a constant amount per step. The process of Problem 3.16 (Exercise 16) instead has $\varepsilon_{n+1} = \varepsilon_n^2/(2x_n) \sim \varepsilon_n^2/(2s)$, a *quadratic* contraction in which the error is squared at each step and the number of correct digits roughly doubles. Squaring beats any fixed multiplier once the error is small, so that process converges far more rapidly; the present iteration is the slower, geometric one.

■

REFLECTIONS.

The statement hands you a recursively defined real sequence and asks three things about it—two subsequences are monotone, and the whole sequence converges to $\sqrt{\alpha}$—so the words to seize on are $x_1 > x_3 > \cdots$, $x_2 < x_4 < \cdots$, and $\lim x_n = \sqrt{\alpha}$. The first two are monotone-subsequence claims, which name two bounded monotone subsequences and so summon monotone convergence (bounded monotone sequences converge; 6.2.2) as one honest route to the third; the value $\sqrt{\alpha}$ is exactly the positive solution of $x = (\alpha+x)/(1+x)$, the equation a limit would have to satisfy, so the target is the recursion's fixed point. That a single sequence splits into a decreasing odd part *and* an increasing even part is itself the loudest clue: the terms are not marching one way but *oscillating* across $\sqrt{\alpha}$, jumping to the far side each step while landing closer. The object that records oscillation cleanly is the error $\varepsilon_n = x_n - \sqrt{\alpha}$, and the whole solution is the discovery that this error satisfies one tidy linear rule.

Believe the oscillation before proving it. With $\alpha = 3$ and $x_1 = 2$ the iterates run $2,\ \frac{5}{3},\ \frac{7}{4},\ \frac{19}{11}, \ldots$, i.e. $2.000,\ 1.667,\ 1.750,\ 1.727, \ldots$, bouncing above and below $\sqrt{3} \approx 1.732$ and squeezing in from both sides—odd terms stepping down, even terms stepping up, both funnelling to $\sqrt{3}$. That picture of two pincers closing on the fixed point is the entire content of (a)–(c).

The engine is the identity $\varepsilon_{n+1} = -c_n \varepsilon_n$ with $c_n = (\sqrt{\alpha} - 1)/(1 + x_n) > 0$, and everything follows by reading off its sign and its size:

(1) subtracting $\sqrt{\alpha}$ from the recursion and using $\alpha - \sqrt{\alpha} = \sqrt{\alpha}(\sqrt{\alpha} - 1)$ collapses the difference into $\varepsilon_{n+1} = -\frac{\sqrt{\alpha}-1}{1+x_n}\varepsilon_n$; this algebraic factorisation is the crux—without it the three parts are unrelated, with it they are one statement;

(2) the minus sign forces the sign of ε_n to alternate, and since $\varepsilon_1 > 0$ (from $x_1 > \sqrt{\alpha}$) the odd errors are positive and the even ones negative, which is just "$x_n > \sqrt{\alpha}$ for odd n, $x_n < \sqrt{\alpha}$ for even n"—the bracketing the two monotone claims need;

(3) the factor c_n is positive, and multiplying out the denominators gives the clean two-step identity $c_{n+1}c_n = (\sqrt{\alpha}-1)^2/(1+\alpha+2x_n)$, bounded by the fixed $r = (\sqrt{\alpha}-1)^2/(1+\alpha) < 1$; so $|\varepsilon_{n+2}| < r|\varepsilon_n|$ and within each parity the magnitudes strictly shrink, which with the fixed sign is $x_1 > x_3 > \cdots$ and $x_2 < x_4 < \cdots$, parts (a) and (b);

(4) for (c) iterating $|\varepsilon_{n+2}| < r|\varepsilon_n|$ down each parity bounds $|\varepsilon_n|$ by a constant times $r^{\lfloor n/2 \rfloor}$, and $r^k \to 0$ since $0 < r < 1$ (a standard limit, 6.5.1), so the error vanishes along both parities and hence along the whole sequence—this is the quantitative route; equally, each subsequence is monotone (by (a),(b)) and bounded (odd terms above $\sqrt{\alpha}$, even below it), so monotone convergence (6.2.2) hands each a limit L with $L^2 = \alpha$ and $L \geq 0$ by order-preservation (6.3.1), forcing $L = \sqrt{\alpha}$;

(5) both parities reach the same $\sqrt{\alpha}$ and together exhaust the indices, so the full sequence converges to $\sqrt{\alpha}$.

The rigor turns on three points worth naming. Positivity of every x_n is not cosmetic: it keeps $1 + x_n > 0$, so $c_n > 0$ and the sign argument is valid, and it is what makes the limit nonnegative so that $L^2 = \alpha$ resolves to $+\sqrt{\alpha}$ rather than $-\sqrt{\alpha}$. The bracketing $x_n < \sqrt{\alpha}$ on even indices and $x_n > \sqrt{\alpha}$ on odd indices is what supplies the bounds the monotone-convergence theorem demands—monotone alone is not enough, an unbounded monotone sequence diverges. And the last step genuinely needs *both* subsequences to reach the same limit: knowing only the odd terms converge would leave the even terms unaccounted for, so the step that the odd and even indices exhaust $\mathbb{N}$ and share the limit is load-bearing, not bookkeeping. Nothing is circular—the value $\sqrt{\alpha}$ is forced out of the fixed-point equation, never assumed.

Part (d) is the payoff and the reason the error identity, not just a convergence verdict, was worth extracting. Here the error is multiplied by $c_n \to (\sqrt{\alpha}-1)/(1+\sqrt{\alpha})$, a fixed constant strictly between 0 and 1: convergence is *geometric*, adding a constant number of correct digits per step. The Newton-style iteration of Problem 3.16 squares the error each step, doubling the correct digits—*quadratic* convergence—which is incomparably faster once the error is small, since a square crushes a quantity below 1 far harder than any fixed multiplier. Reading the rate off the recursion for the error is the transferable move: for a fixed-point iteration, linearise near the fixed point and look at the multiplier on ε_n. A nonzero constant multiplier means linear (geometric) speed; a multiplier that itself vanishes with ε_n—an ε_n^2 law—means quadratic speed. The same $\varepsilon_{n+1} = -c_n\varepsilon_n$ that proved the convergence in (a)–(c) is precisely what measures its rapidity in (d).

Problem 3.18 *Rudin, Chapter 3, Exercise 18*

Replace the recursion formula of Exercise 16 by

$$x_{n+1} = \frac{p-1}{p}x_n + \frac{\alpha}{p}x_n^{-p+1}$$

where p is a fixed positive integer, and describe the behavior of the resulting sequences $\{x_n\}$.

SOLUTION.

Fix $\alpha > 0$ and a positive integer p, and let $g(x) = \dfrac{p-1}{p}\,x + \dfrac{\alpha}{p}\,x^{1-p}$ be the map driving the recursion, defined for $x > 0$. Write $L = \alpha^{1/p} > 0$ for the positive pth root of α, so that $L^p = \alpha$. The verdict: for $p \geq 2$ and any starting value $x_1 > 0$, the sequence satisfies $x_n \geq L$ for all $n \geq 2$, decreases monotonically from the second term on, and converges to $L = \alpha^{1/p}$; the degenerate case $p = 1$ is constant.

The case $p = 1$ is immediate: the formula reads $x_{n+1} = 0 \cdot x_n + \alpha\, x_n^0 = \alpha$, so $x_n = \alpha = \alpha^{1/1}$ for every $n \geq 2$, a constant sequence at its root. Assume $p \geq 2$ from here on.

First, a lower bound: $g(x) \geq L$ for every $x > 0$, with equality only at $x = L$. Substitute $x = tL$ with $t > 0$; using $\alpha = L^p$,

$$g(tL) = \frac{p-1}{p}\,tL + \frac{L^p}{p}\,(tL)^{1-p} = \frac{L}{p}\big[(p-1)t + t^{1-p}\big],$$

so $g(x) \geq L$ is equivalent to $(p-1)t + t^{1-p} \geq p$. Multiplying this by $p\,t^{p-1} > 0$ turns it into the polynomial inequality $(p-1)t^p - p\,t^{p-1} + 1 \geq 0$, and a direct expansion gives the factorization

$$(p-1)t^p - p\,t^{p-1} + 1 = (t-1)^2 \sum_{k=1}^{p-1} k\,t^{\,k-1}.$$

The second factor has only nonnegative coefficients, so it is positive for $t > 0$, while $(t-1)^2 \geq 0$; hence the product is ≥ 0, vanishing exactly when $t = 1$, i.e. $x = L$. Therefore

$g(x) \geq L$ for all $x > 0$, and the map sends $(0, \infty)$ into $[L, \infty)$.

Second, a decrease above the root: $g(x) \leq x$ whenever $x \geq L$. The difference is

$$x - g(x) = \frac{1}{p}\, x - \frac{\alpha}{p}\, x^{1-p} = \frac{1}{p}\, x^{1-p}(x^p - \alpha),$$

and for $x \geq L > 0$ one has $x^p \geq L^p = \alpha$, so $x - g(x) \geq 0$.

Now describe the sequence. Whatever $x_1 > 0$ is chosen, $x_2 = g(x_1) \geq L$ by the lower bound, and the same bound applied at each step keeps $x_n = g(x_{n-1}) \geq L$ for every $n \geq 2$. For $n \geq 2$ the term $x_n \geq L$ then feeds the decrease estimate, giving $x_{n+1} = g(x_n) \leq x_n$. So $\{x_n\}_{n \geq 2}$ is monotonically decreasing and bounded below by L, hence convergent by monotone convergence (a bounded monotone sequence converges; 6.2.2). Call the limit ℓ; since every $x_n \geq L$ for $n \geq 2$, order is preserved in the limit (6.3.1) and $\ell \geq L > 0$.

It remains to identify ℓ. Because $\ell > 0$, the algebra of limits (limits respect the algebraic operations; 5.2.1) applies to the recursion: $x_n^{p-1} \to \ell^{p-1}$ as a finite product of convergent sequences, and as $\ell^{p-1} \neq 0$ the reciprocals give $x_n^{1-p} \to \ell^{1-p}$. The shifted sequence $\{x_{n+1}\}$ has the same limit ℓ, so letting $n \to \infty$ in $x_{n+1} = \dfrac{p-1}{p} x_n + \dfrac{\alpha}{p} x_n^{1-p}$ yields

$$\ell = \frac{p-1}{p}\, \ell + \frac{\alpha}{p}\, \ell^{1-p}.$$

Multiplying through by $\dfrac{p}{\ell^{1-p}} = p\, \ell^{p-1}$ and simplifying gives $\ell^p = \alpha$, so $\ell = \alpha^{1/p} = L$. Thus the recursion is Newton's method for the root of $x^p - \alpha$: from any positive seed it converges monotonically (after the first step) to $\alpha^{1/p}$, recovering Exercise 16 at $p = 2$.

■

Reflections.

The statement is one of a family—Exercises 16, 17, 18—each handing you a recursion $x_{n+1} = g(x_n)$ and asking for its long-run behavior, and the phrase *describe the behavior of the resulting sequences* is the tell. A sequence defined by iterating a single map has no closed form to manipulate, so the only handle our notes offer is monotone convergence (6.2.2): show the sequence is monotone and bounded, harvest a limit, then pin the limit down. That two-move plan—first that the limit *exists*, separately what it *is*—is the spine of every iteration problem, and recognizing the recursion as the cue for the monotone convergence theorem is the first thing to read off.

Believe the answer before proving it. The formula is Newton's method applied to $f(x) = x^p - \alpha$: the update $x_{n+1} = x_n - f(x_n)/f'(x_n)$ works out to exactly $\frac{p-1}{p} x_n + \frac{\alpha}{p} x_n^{1-p}$, and a fixed point $\ell = g(\ell)$ forces $\ell^p = \alpha$. So the target must be $\alpha^{1/p}$, the pth-root generalization of the square-root algorithm in Problem 3.16, which is the case $p = 2$. Naming $L = \alpha^{1/p}$

up front converts every vague "it settles down" into a sharp inequality against a fixed number.

The argument then assembles from four moves, each resting on a specific fact:

(1) the lower bound $g(x) \geq L$ for all $x > 0$, which after the substitution $x = tL$ collapses to the single inequality $(p-1)t + t^{1-p} \geq p$, proved with no calculus by clearing denominators and factoring $(p-1)t^p - p\,t^{p-1} + 1 = (t-1)^2 \sum_{k=1}^{p-1} k\,t^{k-1}$—a perfect square times a polynomial with nonnegative coefficients, hence ≥ 0 on $(0, \infty)$;

(2) the decrease $g(x) \leq x$ for $x \geq L$, which is even cleaner, since $x - g(x) = \frac{1}{p}x^{1-p}(x^p - \alpha)$ is nonnegative precisely when $x^p \geq \alpha$, i.e. when $x \geq L$;

(3) combining the two, every x_n with $n \geq 2$ lies in $[L, \infty)$ and the sequence decreases there, so it is bounded-below and monotone and converges by monotone convergence (6.2.2), with limit $\ell \geq L > 0$ by order-preservation in the limit (6.3.1);

(4) passing to the limit in the recursion via the algebra of limits (the algebra of limits; 5.2.1)—legitimate because $\ell \neq 0$, so $x_n^{1-p} \to \ell^{1-p}$—gives $\ell^p = \alpha$, hence $\ell = \alpha^{1/p}$.

Reading where each hypothesis is spent shows why the description is shaped as it is. The lower bound is what guarantees the sequence enters $[L, \infty)$ *after one step regardless of the seed*, which is why the monotonicity is asserted only from $n \geq 2$: a seed $x_1 < L$ can overshoot upward on the first iterate, and the numerics bear this out, yet $x_2 \geq L$ is forced and the descent begins. This is a genuine difference from Exercise 16's framing, which assumes $x_1 > \sqrt{\alpha}$ to make the sequence decrease from the very start; here the lower bound does that work automatically. The decrease estimate in turn needs $x \geq L$ to make $x^p - \alpha \geq 0$—drop it and the sequence need not be monotone below the root. And the final limit step needs $\ell > 0$ to divide by ℓ^{1-p}; this is exactly what order-preservation supplies, since $\ell \geq L > 0$. Nothing is circular: existence of the limit is settled by monotone convergence before its value is computed, never the reverse.

The transferable move is the whole skeleton of an iteration problem. To analyze $x_{n+1} = g(x_n)$, locate the fixed point $\ell = g(\ell)$ first—it is the only candidate limit—then prove the iterates are trapped on one side of ℓ and monotone toward it, so monotone convergence hands you a limit that the algebra of limits forces to equal ℓ. The same template runs Exercise 16 (Problem 3.16) and, with a two-subsequence twist for the oscillation, Exercise 17 (Problem 3.17); the only craft that changes problem to problem is the inequality certifying boundedness, and here that craft was the telescoping factorization $(t-1)^2 \sum_k k\,t^{k-1}$, the algebraic heart of the weighted arithmetic–geometric mean inequality dressed for a single variable.

Problem 3.19 *Rudin, Chapter 3, Exercise 19*

Associate to each sequence $a = \{\alpha_n\}$, in which α_n is 0 or 2, the real number

$$x(a) = \sum_{n=1}^{\infty} \frac{\alpha_n}{3^n}.$$

Prove that the set of all $x(a)$ is precisely the Cantor set described in Sec. 2.44.

Solution.

The construction of Sec. 2.44 runs as follows. Set $E_0 = [0,1]$, and obtain E_m from E_{m-1} by deleting the open middle third of each of its intervals; then E_m is a union of 2^m closed intervals of length 3^{-m}, the sets are nested $E_0 \supseteq E_1 \supseteq \cdots$, and the Cantor set is

$$P = \bigcap_{m=0}^{\infty} E_m.$$

Write $C = \{x(a) : \alpha_n \in \{0,2\} \text{ for all } n\}$ for the set of values in question. The series defining $x(a)$ converges by comparison with the geometric series $\sum 2 \cdot 3^{-n}$ (each term lies in $[0, 2 \cdot 3^{-n}]$; 6.6.8, 6.6.7), so $x(a) \in [0,1]$ is well defined. We prove $C = P$ by two inclusions.

The key is to read off, from the intervals of E_m, which digits a point may carry. A direct induction shows that E_m is exactly the set of $x \in [0,1]$ admitting a representation

$$x = \sum_{n=1}^{m} \frac{\beta_n}{3^n} + r, \qquad \beta_n \in \{0,2\}, \quad 0 \le r \le 3^{-m}, \tag{$*$}$$

that is, a left endpoint $\sum_{n \le m} \beta_n 3^{-n}$ with all digits 0 or 2, plus a remainder no larger than one interval length. At $m = 0$ this reads $0 \le x \le 1$, which is E_0. Assume it for m. A stage-m interval $\left[t,\, t + 3^{-m}\right]$ with $t = \sum_{n \le m} \beta_n 3^{-n}$ has its open middle third $\left(t + 3^{-m-1},\, t + 2 \cdot 3^{-m-1}\right)$ deleted, leaving the two closed thirds

$$\left[t,\ t + 3^{-m-1}\right] \quad \text{and} \quad \left[t + 2 \cdot 3^{-m-1},\ t + 3^{-m}\right],$$

i.e. $\left[t + \frac{0}{3^{m+1}},\, t + \frac{0}{3^{m+1}} + 3^{-m-1}\right]$ and $\left[t + \frac{2}{3^{m+1}},\, t + \frac{2}{3^{m+1}} + 3^{-m-1}\right]$. So the stage-$(m+1)$ intervals are exactly those with left endpoint $t + \beta_{m+1} 3^{-m-1}$, $\beta_{m+1} \in \{0,2\}$, and length 3^{-m-1}, which is $(*)$ at $m+1$.

For the inclusion $C \subseteq P$, fix a with all $\alpha_n \in \{0,2\}$ and any $m \ge 0$, and split the series at m,

$$x(a) = \underbrace{\sum_{n=1}^{m} \frac{\alpha_n}{3^n}}_{=:t} + \underbrace{\sum_{n=m+1}^{\infty} \frac{\alpha_n}{3^n}}_{=:r},$$

the head t has digits in $\{0,2\}$, and the tail is bounded by the geometric estimate

$$0 \le r \le \sum_{n=m+1}^{\infty} \frac{2}{3^n} = \frac{2 \cdot 3^{-(m+1)}}{1 - 1/3} = 3^{-m}$$

(6.6.7). By $(*)$, $x(a) \in E_m$. As m was arbitrary, $x(a) \in \bigcap_m E_m = P$.

For the reverse inclusion $P \subseteq C$, fix $x \in P$, so $x \in E_m$ for every m. For each m choose, by $(*)$, digits $\beta_1^{(m)}, \ldots, \beta_m^{(m)} \in \{0,2\}$ with $x - \sum_{n \le m} \beta_n^{(m)} 3^{-n} \in [0, 3^{-m}]$. The stage-$m$ interval containing x is unique (distinct stage-m intervals are disjoint, being separated by the deleted middle thirds), so its left endpoint, hence each digit $\beta_n^{(m)}$, is determined by x; and a stage-$(m+1)$ interval lies inside a stage-m interval, so passing from m to $m+1$ keeps the earlier digits and appends one. The digits are therefore consistent: there is a single sequence $\alpha_n \in \{0,2\}$ with $\beta_n^{(m)} = \alpha_n$ for all $n \le m$. For each m,

$$\Big| x - \sum_{n=1}^{m} \frac{\alpha_n}{3^n} \Big| \le 3^{-m},$$

and $3^{-m} \to 0$ (Archimedean; 6.5.1), so the partial sums of $\sum \alpha_n 3^{-n}$ converge to x. By uniqueness of limits (a limit is unique) the sum equals x, that is $x = x(a) \in C$.

The two inclusions give $C = P$.

■

REFLECTIONS.

The statement hands you an infinite series and asks you to identify its range of values with a set built by a completely different recipe—repeatedly deleting middle thirds. The clue is the base: every denominator is a power of 3 and every numerator is a ternary digit (0 or 2, the two digits that survive when the middle third, digit 1, is thrown away). That is the tell that $x(a)$ is nothing but the base-3 expansion $0.\alpha_1\alpha_2\alpha_3\ldots$ with the digit 1 forbidden, and the Cantor set is exactly the points whose ternary expansion can avoid the digit 1. So the problem is the bridge between a *geometric* object (nested intervals) and an *arithmetic* one (digit strings), and the whole task is to make "$x \in E_m$" and "x's first m ternary digits avoid 1" literally the same statement.

Before any inclusion, the series itself must be shown to name a real number, and that is the first place the sequence machinery enters: each term lies in $[0, 2 \cdot 3^{-n}]$, so the series is dominated by the geometric series $\sum 2 \cdot 3^{-n}$ and converges by comparison (6.6.8, 6.6.7). The same geometric sum, $\sum_{n>m} 2 \cdot 3^{-n} = 3^{-m}$, is the quantitative heart of everything that follows: it says the entire tail of the expansion past digit m can shift the value by at most one stage-m interval length. Believe the correspondence on the smallest case first—$\alpha_1 = 0$ keeps you in $[0, \frac{1}{3}]$ and $\alpha_1 = 2$ in $[\frac{2}{3}, 1]$, which are exactly the two intervals of E_1, while $\alpha_1 = 1$ would land you in the forbidden middle $(\frac{1}{3}, \frac{2}{3})$.

The engine is the single inductive description $(*)$: E_m is the set of x that can be written as a length-m ternary head with digits in $\{0,2\}$ plus a remainder in $[0,3^{-m}]$. The induction is where the geometry becomes arithmetic, and it runs in clean steps:

(1) the base case is $E_0 = [0,1]$, which is $(*)$ with an empty head and remainder $r = x$;

(2) the inductive step takes one stage-m interval $[t, t+3^{-m}]$ with t a $\{0,2\}$-head, deletes its open middle third, and observes that the two surviving closed thirds have left endpoints $t + 0 \cdot 3^{-m-1}$ and $t + 2 \cdot 3^{-m-1}$—appending precisely a new digit 0 or 2, which is $(*)$ one level deeper;

(3) with $(*)$ in hand, $C \subseteq P$ is the forward reading: split $x(a)$ into its first m digits plus a tail, bound the tail by the geometric 3^{-m} (6.6.7), and $(*)$ places $x(a)$ in every E_m, hence in the intersection P;

(4) $P \subseteq C$ is the reverse reading: a point of every E_m chooses a digit at each stage, those choices are forced and consistent because the stage-$(m+1)$ interval sits inside the stage-m interval, and the resulting partial sums approximate x to within 3^{-m}, so they converge to x by the standard limit $3^{-m} \to 0$ (6.5.1).

The rigor turns on two points that are easy to wave past. First, in $P \subseteq C$ the digits must be *consistent* across stages, not merely available at each: this is what nestedness of the intervals buys you—because each stage-$(m+1)$ interval lies in a unique stage-m interval, the choice at level $m+1$ cannot overwrite the earlier digits, so the per-stage digits assemble into one genuine sequence $\{\alpha_n\}$. Drop nestedness and you would have a different finite digit string at each stage with no single limit. Second, the passage "partial sums within 3^{-m} of x, and $3^{-m} \to 0$, therefore the sum is x" is exactly convergence of the series to x (5.1.1), made legitimate by the standard limit $3^{-m} \to 0$ (Archimedean, 6.5.1) and pinned down by uniqueness of the limit (a limit is unique)—the series cannot converge to its own value and also to something else. Nothing is circular: $(*)$ is proved by an induction that never mentions C or the series, and only afterward is it read in both directions.

One subtlety the digit-1 ban quietly handles is the ambiguity of expansions. A few reals carry two ternary names—$\frac{1}{3} = 0.1000\cdots = 0.0222\ldots$—and the Cantor set keeps such a point precisely because it admits the all-$\{0,2\}$ name $0.0222\ldots$; forbidding the digit 1 is what selects the surviving representation. This is the ternary twin of the decimal $0.4999\cdots = 0.5000\ldots$ tail-dodging that made Cantor's diagonal argument honest in HW2 Problem 4, and it is why the endpoints of every deleted interval stay in P.

The transferable move is to recognise a digit expansion hiding inside a geometric construction: whenever a set is built by keeping a fixed subset of positions at each scale—ternary digits in $\{0,2\}$ here, the digits 4 and 7 in the digits-4-and-7 set of HW2 Problem 17—encode membership as a constraint on base-b digits, and a nested-interval geometry collapses into

an arithmetic statement about sequences. Once the correspondence is set up, the convergence of the defining series and the standard limit $b^{-m} \to 0$ do all the analytic work, and the topology of the Cantor set (it is uncountable, perfect, and nowhere dense) becomes a statement about $\{0, 2\}$-valued sequences, which is the cleanest way to see why P is as large as $\mathbb{R}$ yet contains no interval.

Problem 3.20 *Rudin, Chapter 3, Exercise 20*

Suppose $\{p_n\}$ is a Cauchy sequence in a metric space X, and some subsequence $\{p_{n_i}\}$ converges to a point $p \in X$. Prove that the full sequence $\{p_n\}$ converges to p.

SOLUTION.

Write d for the metric on X. A sequence is Cauchy when its terms eventually crowd together (5.4.1), and $\{p_n\}$ converges to p when every tolerance is met from some index on (5.1.1); the goal is to produce, for each $\varepsilon > 0$, an index past which $d(p_n, p) < \varepsilon$.

Fix $\varepsilon > 0$. Since $\{p_n\}$ is Cauchy, there is an N with

$$d(p_m, p_n) < \tfrac{\varepsilon}{2} \qquad \text{for all } m, n \geq N.$$

Since the subsequence $\{p_{n_i}\}$ converges to p, there is an I with $d(p_{n_i}, p) < \frac{\varepsilon}{2}$ for all $i \geq I$. The indices of a subsequence strictly increase, so $n_i \geq i$ for every i (5.3.1); choose a single index i large enough that both $i \geq I$ and $n_i \geq N$, for instance $i = \max\{I, N\}$. Then $q := p_{n_i}$ satisfies $d(q, p) < \frac{\varepsilon}{2}$, and $n_i \geq N$.

Now take any $n \geq N$. Both p_n and $q = p_{n_i}$ carry indices at least N, so the Cauchy estimate applies to the pair, and the triangle inequality routes p_n to p through q:

$$d(p_n, p) \leq d(p_n, p_{n_i}) + d(p_{n_i}, p) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Thus $d(p_n, p) < \varepsilon$ for every $n \geq N$, and since $\varepsilon > 0$ was arbitrary, $p_n \to p$ (5.1.1).

■

REFLECTIONS.

The statement hands you two pieces of information about the *same* sequence and asks you to merge them. The first, *Cauchy*, says the terms cluster together without naming a target (5.4.1); the second, that a *subsequence converges to p*, names the only candidate target in sight (5.3.1, 5.3.3). Convergence of $\{p_n\}$ needs both a destination and a guarantee the whole tail arrives there (5.1.1): the subsequence supplies the destination p, and the Cauchy condition supplies the discipline that drags the rest of the tail along. So the proof is a meeting of the two hypotheses, and the natural meeting place is the triangle inequality, the one tool that lets a fact about p_n near q and a fact about q near p combine into a fact about p_n near p. This is exactly the quotable packaging "if a subsequence of a Cauchy

sequence converges to p, so does the sequence"; here that fact is not invoked but proved.

Believe it first on the picture. A Cauchy sequence is a swarm that contracts to a speck: past some index every term sits within $\varepsilon/2$ of every other. The convergent subsequence pins where that speck is—it is sitting on p. A swarm contracting onto one of its own members that happens to rest at p cannot settle anywhere but p; the only thing missing is a single subsequential term deep enough in the tail to act as the anchor, and that is what choosing i large achieves.

The argument is short, and each move is load-bearing:

(1) from the Cauchy property pull an N with $d(p_m, p_n) < \varepsilon/2$ for all $m, n \geq N$ (5.4.1)—the half-ε is chosen now so the two contributions sum to ε later;

(2) from the subsequence's convergence pull an I with $d(p_{n_i}, p) < \varepsilon/2$ for $i \geq I$ (5.1.1, all but finitely many terms lie in each ball);

(3) secure one anchor term $q = p_{n_i}$ that is *both* close to p and deep in the tail, using $n_i \geq i$ (subsequence indices strictly increase, 5.3.1) to force $n_i \geq N$ by taking $i \geq \max\{I, N\}$—this is the step that fuses the two indices into one usable witness;

(4) for any $n \geq N$, the triangle inequality $d(p_n, p) \leq d(p_n, q) + d(q, p) < \varepsilon/2 + \varepsilon/2$ closes the gap, the first term controlled because both p_n and q have indices $\geq N$, the second because q is a far-enough subsequential term.

Each hypothesis earns its place, which is the way to see the result is sharp. Drop "Cauchy" and a convergent subsequence says nothing about the rest: $p_n = (-1)^n$ has the even subsequence converging to 1, yet $\{p_n\}$ diverges, because there is no contraction to spread that limit to the odd terms. Drop the converging subsequence and a Cauchy sequence in an incomplete space may have no limit at all to inherit—this is precisely why a Cauchy sequence *with* a convergent subsequence is the hypothesis that closes the gap, the mechanism behind completeness of $\mathbb{R}^k$ and of compact spaces (5.5.3). The reasoning is not circular: nowhere is $p_n \to p$ assumed; p is read off the subsequence, and the limit of the whole sequence is built from the half-ε estimates, not presupposed.

The transferable move is the triangle-inequality relay with a single well-chosen waypoint: to show p_n is near p when you only know p_n is near *its neighbours* and that *one* neighbour is near p, route through that neighbour and split the tolerance in half. The one subtlety worth banking is step (3)—a subsequence reaches arbitrarily far out ($n_i \geq i$), so it can always hand you an anchor lying as deep in the tail as the Cauchy index demands. This same fusion of "Cauchy" with "one convergent subsequence" is what powers the completeness theorems: a compact space forces a convergent subsequence out of any sequence (every sequence in a compact space has a convergent subsequence, 5.3.3), and this exercise then

upgrades a Cauchy sequence to a convergent one, so the space is complete (5.5.3).

Problem 3.21 *Rudin, Chapter 3, Exercise 21*

Prove the following analogue of Theorem 3.10(b): If $\{E_n\}$ is a sequence of closed nonempty and bounded sets in a complete metric space X, if $E_n \supset E_{n+1}$, and if

$$\lim_{n\to\infty} \operatorname{diam} E_n = 0,$$

then $\bigcap_1^\infty E_n$ consists of exactly one point.

SOLUTION.

Work in the metric space (X, d), and write $\operatorname{diam} E = \sup\{d(x, y) : x, y \in E\}$ for the diameter (5.5.1). The claim has two halves: the intersection $E = \bigcap_{n=1}^\infty E_n$ is nonempty, and it has at most one point.

First, E has at most one point. If $x, y \in E$, then $x, y \in E_n$ for every n, so $0 \le d(x, y) \le \operatorname{diam} E_n$ for every n. Letting $n \to \infty$ forces $d(x, y) \le \lim_n \operatorname{diam} E_n = 0$, hence $d(x, y) = 0$ and $x = y$—and this is the only place the diameter hypothesis is used.

It remains to show E is nonempty. Each E_n is nonempty, so choose a point $x_n \in E_n$ for every n. The sequence $\{x_n\}$ is Cauchy. Given $\varepsilon > 0$, the hypothesis $\operatorname{diam} E_n \to 0$ supplies an N with $\operatorname{diam} E_N < \varepsilon$. For any $m, n \ge N$ the nesting $E_n \subseteq E_N$ and $E_m \subseteq E_N$ (which follows from $E_k \supseteq E_{k+1}$) places both x_m and x_n in E_N, so

$$d(x_m, x_n) \le \operatorname{diam} E_N < \varepsilon.$$

Thus $\{x_n\}$ is Cauchy (5.4.1). Since X is complete (5.4.3), the sequence converges to some $x \in X$.

This limit lies in every E_m. Fix m. For each $n \ge m$ the nesting gives $x_n \in E_n \subseteq E_m$, so the tail $\{x_n\}_{n\ge m}$ lies entirely in E_m and still converges to x. A convergent sequence drawn from a set has its limit in that set's closure; as E_m is closed (3.2.4), the limit satisfies $x \in E_m$. Concretely, were $x \notin E_m$, then x would be neither a point nor a cluster point of E_m, so some ball $B_r(x)$ would miss E_m (3.2.5)—yet $x_n \to x$ puts $x_n \in B_r(x)$ for all large n (all but finitely many terms enter every ball about the limit), and these $x_n \in E_m$, a contradiction. Hence $x \in E_m$.

Since m was arbitrary, $x \in \bigcap_{n=1}^\infty E_n = E$, so E is nonempty. Combined with the first half, E consists of exactly one point.

■

REFLECTIONS.

The statement announces itself as an "analogue of Theorem 3.10(b)," and the first move is to recall what that theorem says: a nested sequence of nonempty *compact* sets has nonempty intersection (our nested nonempty compacts meet, 4.2.2). The present problem makes two changes at once, and reading each is the whole orientation. "Compact" has been weakened to "closed and bounded in a complete space"—and over a general metric space those are not the same (closed-and-bounded need not be compact), so the proof cannot lean on compactness and must instead earn nonemptiness from *completeness* (5.4.3). In exchange, the new diameter hypothesis $\operatorname{diam} E_n \to 0$ is handed to us, and a hypothesis that strong is exactly what upgrades the conclusion from "nonempty" to "exactly one point." So the word *complete* points at the machinery for producing the point, and the phrase *exactly one* splits the task cleanly in two.

It helps to believe the model case first. On the real line take $E_n = [0, 1/n]$: each is closed, nonempty, bounded, nested downward, with $\operatorname{diam} E_n = 1/n \to 0$, and the intersection is the single point $\{0\}$. That is the nested-interval theorem (4.2.3) in miniature, and the present exercise is its abstraction—the same picture of sets clamping down on one point, but with completeness standing in for the special structure of $\mathbb{R}$.

How does completeness produce a point when there is no ambient compactness to call on? By the one tool that conjures a limit out of internal data alone: the Cauchy criterion (5.4.1). The recipe is to manufacture a candidate sequence and prove it Cauchy, and the proof runs in these steps:

(1) picking one x_n from each E_n is legitimate precisely because every E_n is nonempty—the nonemptiness hypothesis is spent here, on the very existence of the sequence;

(2) the sequence is Cauchy because, once $\operatorname{diam} E_N < \varepsilon$, every later index $m, n \geq N$ has both terms trapped inside the single set E_N (this is where the nesting $E_n \subseteq E_N$ does its work), so $d(x_m, x_n) \leq \operatorname{diam} E_N < \varepsilon$;

(3) completeness (5.4.3; Cauchy sequences converge in a complete space) converts "Cauchy" into an actual limit $x \in X$—the step that has no analogue in an incomplete space;

(4) the limit lands in each E_m because, from index m onward, the whole tail of the sequence sits in the closed set E_m, and a closed set contains the limits of its convergent sequences (3.2.4, via the tail of a convergent sequence enters every ball about its limit);

(5) uniqueness is the diameter hypothesis again: two points of the intersection lie in every E_n, so their distance is bounded by $\operatorname{diam} E_n \to 0$, forcing distance zero.

Watching where each hypothesis is consumed shows the statement is sharp: drop any one and it fails. Without nonemptiness there is no sequence to build (and the intersection could

be empty outright); without nesting, terms from different E_n need not huddle together and the Cauchy estimate collapses ($E_n = \{0\}$ and $\{1\}$ alternately keep diameters small yet share no point); without closedness the limit can leak out the boundary—$E_n = (0, 1/n)$ in $\mathbb{R}$ has empty intersection though everything else holds; without completeness the Cauchy sequence may have nowhere to converge—inside $\mathbb{Q}$, the nested closed sets $E_n = \{q \in \mathbb{Q} : |q - \sqrt{2}| \leq 1/n\}$ shrink onto a hole and meet in nothing; and without $\operatorname{diam} E_n \to 0$ the intersection can be large, as $E_n = [0, 1] \subseteq \mathbb{R}$ shows, where it is the whole interval. The argument is not circular: the point is built from the Cauchy sequence, never assumed to exist in advance, and uniqueness is checked separately from existence.

The transferable lesson is the division of labour between the two new conditions. Completeness is the existence engine—it is the abstract replacement for compactness in any "nested sets must meet" theorem, accessed through the Cauchy criterion rather than through subsequences. The shrinking diameter is the uniqueness engine, the same device that pins a real number between nested intervals. Together they recover, in any complete space, the full force of the nested-interval theorem (4.2.3); and the diameter-of-nested-sets viewpoint here is exactly the one the notes use to prove that compact spaces and $\mathbb{R}^k$ are complete (5.5.2, 5.5.3), so the present exercise is that proof's mirror image—there completeness is the conclusion, here it is the tool.

Problem 3.22 *Rudin, Chapter 3, Exercise 22*

Suppose X is a nonempty complete metric space, and $\{G_n\}$ is a sequence of dense open subsets of X. Prove Baire's theorem, namely, that $\bigcap_1^\infty G_n$ is not empty. (In fact, it is dense in X.) *Hint:* Find a shrinking sequence of neighborhoods E_n such that $\overline{E}_n \subset G_n$, and apply Exercise 21.

SOLUTION.

Write d for the metric, $B(p, \rho) = \{x : d(p, x) < \rho\}$ for the open ball, and $\overline{B}(p, \rho) = \{x : d(p, x) \leq \rho\}$. Density of a set G means G meets every nonempty open set (4.4.1). It is enough to prove the stronger claim that $\bigcap_{n\geq 1} G_n$ is dense, for then, X being a nonempty open set, the intersection meets X and so is nonempty.

So fix a nonempty open $W \subseteq X$; the goal is a point of $\bigcap_{n\geq 1} G_n$ lying in W. One small shrinking lemma is used repeatedly: if $B(p, \rho) \subseteq V$ with V open, then $\overline{B}(p, \rho/2) \subseteq B(p, \rho) \subseteq V$, so any nonempty open set contains the closure of a ball about any of its points, of as small a radius as we please.

Build closed balls $\overline{E}_n = \overline{B}(x_n, r_n)$ by recursion. Since G_1 is dense and W is nonempty open, $W \cap G_1$ is nonempty; it is open, being the intersection of two open sets. Choose a point in it and, by the shrinking lemma, a radius with

$$\overline{E}_1 = \overline{B}(x_1, r_1) \subseteq W \cap G_1, \qquad 0 < r_1 < 1.$$

Suppose $\overline{E}_n = \overline{B}(x_n, r_n)$ has been chosen. The open ball $B(x_n, r_n)$ is nonempty and open (2.3.1), and G_{n+1} is dense and open, so $B(x_n, r_n) \cap G_{n+1}$ is nonempty and open. Applying the shrinking lemma inside it, choose

$$\overline{E}_{n+1} = \overline{B}(x_{n+1}, r_{n+1}) \subseteq B(x_n, r_n) \cap G_{n+1}, \qquad 0 < r_{n+1} < \tfrac{1}{n+1}.$$

Three properties of the family $\{\overline{E}_n\}$ follow at once. Each $\overline{E}_n$ is a closed ball, hence nonempty (it contains x_n), closed, and bounded, with $\operatorname{diam} \overline{E}_n \leq 2r_n$ (5.5.1). They are nested, since $\overline{E}_{n+1} \subseteq B(x_n, r_n) \subseteq \overline{B}(x_n, r_n) = \overline{E}_n$, so $\overline{E}_n \supseteq \overline{E}_{n+1}$ for every n. And the diameters vanish: from $r_n < \frac{1}{n}$ (with $r_1 < 1 = \frac{1}{1}$),

$$0 \leq \operatorname{diam} \overline{E}_n \leq 2r_n < \tfrac{2}{n} \xrightarrow[n\to\infty]{} 0.$$

The family $\{\overline{E}_n\}$ is therefore a nested sequence of nonempty closed bounded sets in the complete space X with $\operatorname{diam} \overline{E}_n \to 0$. By Problem 3.21, the intersection $\bigcap_{n\geq 1} \overline{E}_n$ consists of exactly one point; call it x.

This x lands where it should. For every n the inclusion $\overline{E}_{n+1} \subseteq B(x_n, r_n) \cap G_{n+1} \subseteq G_{n+1}$ gives $x \in \overline{E}_{n+1} \subseteq G_{n+1}$, and $x \in \overline{E}_1 \subseteq G_1$, so $x \in G_n$ for every $n \geq 1$; and $x \in \overline{E}_1 \subseteq W$. Hence

$$x \in W \cap \bigcap_{n=1}^{\infty} G_n.$$

Thus $\bigcap_{n\geq 1} G_n$ meets the arbitrary nonempty open set W, so it is dense in X (4.4.1), and being dense in the nonempty space X it is nonempty.

■

Reflections.

The phrase to read first is *complete metric space*: completeness is the one hypothesis without which the conclusion is false, so the proof must spend it somewhere, and the explicit hint says where—through Problem 3.21, the analogue of nested closed intervals that turns a tower of shrinking closed sets into a single common point. That points the whole strategy at a construction: a point lying in *all* of the G_n cannot be exhibited by a formula, so it must be *trapped*, caught as the unique limit of a nested family. The other two words in the hypotheses say what the family must dodge at each stage—*open* is what lets a ball be seated inside a set at all (2.3.3), and *dense* is what guarantees the next G_n reaches into whatever ball has been built so far (4.4.1). And "in fact dense" tells you not to aim at a bare point of $\bigcap G_n$ but to make the construction start inside an arbitrary nonempty open W, since meeting every such W is exactly what density means (4.4.1) and nonemptiness then comes for free from $W = X$.

It is worth seeing the difficulty before the construction. To sit in G_1 is easy; to sit in $G_1 \cap G_2$

is still easy; the trouble is the *infinite* intersection, where no finite stage commits you to a point and the limit could slip out through a gap. Completeness is precisely the promise that a controlled limit cannot slip out, provided the closed sets shrink to nothing—which is why the radii are forced to zero rather than merely kept positive.

The construction is a short recursive loop, each step resting on the one before:

(1) density of G_1 makes $W \cap G_1$ nonempty and intersection-of-opens keeps it open, so the shrinking lemma seats a closed ball $\overline{E}_1 \subseteq W \cap G_1$ of radius < 1—the base of the tower, planted inside W so the final point will land in W;

(2) at stage n the *open* ball $B(x_n, r_n)$ is a legitimate nonempty open target (2.3.1), so density of G_{n+1} lets it reach in and openness of $B(x_n, r_n) \cap G_{n+1}$ lets the shrinking lemma seat the next closed ball $\overline{E}_{n+1}$ inside it, with radius $< \frac{1}{n+1}$; one shrinks into the *open* ball $B(x_n, r_n)$, not the closed ball $\overline{E}_n$, exactly because a ball can only be seated inside an open set, and $B(x_n, r_n) \cap G_{n+1}$ is open while $\overline{E}_n \cap G_{n+1}$ need not be;

(3) nesting $\overline{E}_{n+1} \subseteq \overline{E}_n$ then comes for free from $B(x_n, r_n) \subseteq \overline{B}(x_n, r_n) = \overline{E}_n$, and the radius cap forces $\operatorname{diam} \overline{E}_n \leq 2r_n < \frac{2}{n} \to 0$ (5.5.1);

(4) the family is now nested, nonempty, closed, bounded, with diameters shrinking to zero, so Problem 3.21 hands back a single point $x \in \bigcap_n \overline{E}_n$, and $\overline{E}_n \subseteq G_n$ together with $\overline{E}_1 \subseteq W$ places it in $W \cap \bigcap_n G_n$.

The rigor turns on three things being genuinely present. The diameters must *vanish*, not merely the sets be nested: that is the whole hypothesis of Problem 3.21 and the reason it returns *one* point rather than possibly none—a nested sequence of nonempty closed sets can have empty intersection in a complete space (the closed sets $\{n, n+1, \dots\} \subseteq \mathbb{R}$ are a witness) unless their size is squeezed to zero. Completeness is load-bearing in the same breath, since Problem 3.21 is exactly where it is consumed; drop it and Baire fails outright, as $\mathbb{Q}$ shows—it is the countable union of its closed, empty-interior singletons, and the matching dense open sets $\mathbb{Q} \setminus \{q\}$ have empty intersection. And openness is what licenses every seating of a ball, while density is what keeps each G_{n+1} from missing the current ball; remove either and the recursion stalls at its first step.

The decisive contrast is with the $\mathbb{R}^k$ version, Problem 30, where the same tower is pinned down instead by nested nonempty *compact* sets (nested compacts meet), and the radii were pure decoration. That shortcut is unavailable here: in a general metric space a closed ball need not be compact, so there is no nested-compact property to fall back on, and the only substitute is completeness packaged as Problem 3.21—which is precisely why the diameters that were decorative there become essential now (completeness of X; 5.4.3). The transferable move is the one the hint is really teaching: to produce a point lying in infinitely many sets at once, never seek it directly—nest a tower of closed sets so that membership in

> the n-th set is built into the n-th ball, shrink the diameters to zero, and let completeness deliver the unique common point. This is the skeleton of every Baire-category argument.

Problem 3.23 *Rudin, Chapter 3, Exercise 23*

Suppose $\{p_n\}$ and $\{q_n\}$ are Cauchy sequences in a metric space X. Show that the sequence $\{d(p_n, q_n)\}$ converges. *Hint:* For any m, n,

$$d(p_n, q_n) \leq d(p_n, p_m) + d(p_m, q_m) + d(q_m, q_n);$$

it follows that

$$|d(p_n, q_n) - d(p_m, q_m)|$$

is small if m and n are large.

Solution.

Write $a_n = d(p_n, q_n)$; this is a sequence of real numbers, and the claim is that it converges in $\mathbb{R}$. We show $\{a_n\}$ is Cauchy and invoke the completeness of $\mathbb{R}$.

The key estimate compares a_n with a_m by routing each through the other pair of indices. The triangle inequality (2.2.1), applied twice, gives the hint's bound

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

which rearranges to

$$a_n - a_m = d(p_n, q_n) - d(p_m, q_m) \leq d(p_n, p_m) + d(q_m, q_n).$$

The roles of m and n are symmetric, so swapping them yields $a_m - a_n \leq d(p_m, p_n) + d(q_n, q_m)$. Since a metric is symmetric, the two right-hand sides agree, and combining the inequalities,

$$|a_n - a_m| = |d(p_n, q_n) - d(p_m, q_m)| \leq d(p_n, p_m) + d(q_n, q_m).$$

Now fix $\varepsilon > 0$. Because $\{p_n\}$ is Cauchy (5.4.1), there is an N_1 with $d(p_n, p_m) < \varepsilon/2$ whenever $m, n \geq N_1$; because $\{q_n\}$ is Cauchy, there is an N_2 with $d(q_n, q_m) < \varepsilon/2$ whenever $m, n \geq N_2$. Put $N = \max\{N_1, N_2\}$. For all $m, n \geq N$ the displayed bound gives

$$|a_n - a_m| \leq d(p_n, p_m) + d(q_n, q_m) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

Thus $\{a_n\}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete (every Cauchy sequence in $\mathbb{R}$ converges; 5.5.3), the sequence $\{a_n\} = \{d(p_n, q_n)\}$ converges to a real number.

■

REFLECTIONS.

Two clues in the statement decide the whole approach. The hypotheses hand you *Cauchy* sequences (5.4.1) but never a limit, and the goal is to show a derived sequence *converges* without naming what it converges to—and a request to prove convergence with no candidate limit in sight is the standing signal to test the Cauchy condition instead (6.1.3). The second clue is where the target sequence lives: $d(p_n, q_n)$ is a real number, so $\{a_n\} = \{d(p_n, q_n)\}$ is a sequence in $\mathbb{R}$, and $\mathbb{R}$ is the space where "Cauchy" is allowed to mean "convergent." Those two readings fit together into a single plan—show $\{a_n\}$ is Cauchy, then let completeness ($\mathbb{R}$ is complete; 5.5.3) supply the limit for free.

It is worth seeing why the limit cannot simply be written down. The space X need not be complete, so $\{p_n\}$ and $\{q_n\}$ may have no limits in X at all—imagine $X = \mathbb{Q}$ with $p_n \to \sqrt{2}$ and $q_n \to 0$ "from outside." There is then no point $p = \lim p_n$ to feed into $d(\cdot, \cdot)$, so any argument that first locates the limits of the two sequences is doomed. What survives is that the *distances* between the moving points settle down even when the points themselves have nowhere to land, and that settling-down is exactly Cauchyness of $\{a_n\}$.

The hint is really the proof, and unpacking it is the exercise:

(1) the triangle inequality (2.2.1), applied along the path $p_n \rightsquigarrow p_m \rightsquigarrow q_m \rightsquigarrow q_n$, gives the hint's three-term bound and so controls $a_n - a_m$ by $d(p_n, p_m) + d(q_m, q_n)$;

(2) swapping m and n controls $a_m - a_n$ by the same quantity (the metric is symmetric, so the bound is unchanged), and the two one-sided estimates combine into the two-sided $|a_n - a_m| \le d(p_n, p_m) + d(q_n, q_m)$;

(3) feeding in the two Cauchy conditions with tolerance $\varepsilon/2$ apiece, and taking the larger of the two thresholds, forces $|a_n - a_m| < \varepsilon$ for all large m, n—so $\{a_n\}$ is Cauchy in $\mathbb{R}$;

(4) completeness of $\mathbb{R}$ (5.5.3) converts Cauchy into convergent, finishing the proof.

The rigor turns on three quiet points, each load-bearing. The two-sided absolute value in step (2) is what a Cauchy estimate requires, and getting it needs *both* directions of the triangle inequality together with symmetry of d; with only the one-sided bound from the hint one would have controlled $a_n - a_m$ but not $|a_n - a_m|$. The split of ε into halves and the $N = \max\{N_1, N_2\}$ in step (3) are the standard device for honouring "for all m, n large" from two separate Cauchy hypotheses at once; drop either sequence's Cauchyness and the bound has an uncontrolled term, after which $\{d(p_n, q_n)\}$ can oscillate—take $p_n = 0$ and $q_n = 1 + (-1)^n$ in $\mathbb{R}$, where $\{p_n\}$ is Cauchy but $\{q_n\}$ is not and $d(p_n, q_n)$ runs $0, 2, 0, 2, \ldots$ without settling. Completeness in step (4) is not optional decoration: it is invoked for $\mathbb{R}$, the codomain of the metric, precisely because X itself may fail to be complete, which is why the conclusion is stated for the real sequence and not for any limit inside X.

The transferable move is to recognise when "prove it converges" should be answered by "prove it is Cauchy." Whenever the candidate limit is unavailable—unnamed, or living in a space that may not contain it—the Cauchy criterion (5.4.1) lets you certify convergence from the terms alone, provided you finish in a complete space. Here the trick is to notice that although the points may wander out of X, their pairwise distance is a real number, and $\mathbb{R}$ is always complete: the quotable relationship between convergent and Cauchy sequences run in reverse, pushed through the one inequality that every metric guarantees.

Problem 3.24 *Rudin, Chapter 3, Exercise 24*

Let X be a metric space.

1. Call two Cauchy sequences $\{p_n\}$, $\{q_n\}$ in X *equivalent* if

$$\lim_{n\to\infty} d(p_n, q_n) = 0.$$

   Prove that this is an equivalence relation.

2. Let X^* be the set of all equivalence classes so obtained. If $P \in X^*$, $Q \in X^*$, $\{p_n\} \in P$, $\{q_n\} \in Q$, define

$$\Delta(P, Q) = \lim_{n\to\infty} d(p_n, q_n);$$

   by Exercise 23, this limit exists. Show that the number $\Delta(P, Q)$ is unchanged if $\{p_n\}$ and $\{q_n\}$ are replaced by equivalent sequences, and hence that Δ is a distance function in X^*.

3. Prove that the resulting metric space X^* is complete.

4. For each $p \in X$, there is a Cauchy sequence all of whose terms are p; let P_p be the element of X^* which contains this sequence. Prove that

$$\Delta(P_p, P_q) = d(p, q)$$

   for all $p, q \in X$. In other words, the mapping φ defined by $\varphi(p) = P_p$ is an isometry (i.e., a distance-preserving mapping) of X into X^*.

5. Prove that $\varphi(X)$ is dense in X^*, and that $\varphi(X) = X^*$ if X is complete. By (d), we may identify X and $\varphi(X)$ and thus regard X as embedded in the complete metric space X^*. We call X^* the *completion* of X.

Solution.

Write $\mathcal{C}$ for the set of Cauchy sequences in (X, d), and for $\{p_n\}, \{q_n\} \in \mathcal{C}$ write $\{p_n\} \sim \{q_n\}$ when $\lim_n d(p_n, q_n) = 0$. Each such limit exists by Problem 3.23, whose estimate $|d(p_n, q_n) - d(p_m, q_m)| \le d(p_n, p_m) + d(q_n, q_m)$ makes $\{d(p_n, q_n)\}$ a Cauchy, hence convergent, sequence of reals.

(a) An equivalence relation. Reflexivity is $\lim_n d(p_n, p_n) = \lim_n 0 = 0$. Symmetry is the symmetry of the metric: $d(p_n, q_n) = d(q_n, p_n)$, so the two limits agree. For transitivity, suppose $\{p_n\} \sim \{q_n\}$ and $\{q_n\} \sim \{r_n\}$. The triangle inequality gives $0 \le d(p_n, r_n) \le d(p_n, q_n) + d(q_n, r_n)$, and the right-hand side tends to $0 + 0 = 0$, so the squeeze (6.3.3) forces $d(p_n, r_n) \to 0$, i.e. $\{p_n\} \sim \{r_n\}$. Thus $\sim$ is reflexive, symmetric, and transitive (0.4.16).

(b) Δ is well defined and a metric. Let $\{p_n\} \sim \{p'_n\}$ and $\{q_n\} \sim \{q'_n\}$. Applying the triangle inequality twice,

$$d(p_n, q_n) \le d(p_n, p'_n) + d(p'_n, q'_n) + d(q'_n, q_n),$$

and the symmetric estimate with the roles reversed, gives

$$\big|\, d(p_n, q_n) - d(p'_n, q'_n) \,\big| \le d(p_n, p'_n) + d(q_n, q'_n).$$

The right-hand side tends to 0 by the two equivalences, so $\lim_n d(p_n, q_n) = \lim_n d(p'_n, q'_n)$: the value depends only on the classes P, Q, not on the chosen representatives. Hence $\Delta(P, Q) = \lim_n d(p_n, q_n)$ is a well-defined number, and it inherits the metric axioms (2.2.1) from d in the limit. Each $\Delta(P, Q) = \lim_n d(p_n, q_n) \ge 0$; it is symmetric because each $d(p_n, q_n)$ is; the triangle inequality $d(p_n, r_n) \le d(p_n, q_n) + d(q_n, r_n)$ passes to the limit by the order properties of limits (6.3.1), giving $\Delta(P, R) \le \Delta(P, Q) + \Delta(Q, R)$. Finally $\Delta(P, Q) = 0$ means $\lim_n d(p_n, q_n) = 0$, which is exactly $\{p_n\} \sim \{q_n\}$, i.e. $P = Q$. So Δ is a distance function on X^*.

(c) X^* is complete. Let $\{P^{(k)}\}_{k \ge 0}$ be a Cauchy sequence in (X^*, Δ); the goal is a class $P \in X^*$ with $\Delta(P^{(k)}, P) \to 0$. For each k pick a representative Cauchy sequence $\{p_n^{(k)}\}_n \in P^{(k)}$. Since $\{p_n^{(k)}\}_n$ is Cauchy in X, choose an index N_k with

$$d(p_m^{(k)}, p_n^{(k)}) < \tfrac{1}{k+1} \qquad \text{for all } m, n \ge N_k,$$

and set $r_k := p_{N_k}^{(k)}$, the "N_k-th term" of the k-th representative. We show the diagonal sequence $\{r_k\}_k$ is Cauchy in X and that its class is the limit of $\{P^{(k)}\}$.

First, $\Delta\big(\varphi(r_k), P^{(k)}\big) \le \frac{1}{k+1}$, where $\varphi(r_k)$ is the class of the constant sequence at r_k. The distance is $\lim_n d\big(r_k, p_n^{(k)}\big)$, and for $n \ge N_k$ the choice of N_k gives $d\big(r_k, p_n^{(k)}\big) = d\big(p_{N_k}^{(k)}, p_n^{(k)}\big) < \frac{1}{k+1}$; passing to the limit in n (6.3.1) yields the bound. Now

$$\begin{aligned} d(r_j, r_k) &= \Delta\big(\varphi(r_j), \varphi(r_k)\big) \\ &\le \Delta\big(\varphi(r_j), P^{(j)}\big) + \Delta\big(P^{(j)}, P^{(k)}\big) + \Delta\big(P^{(k)}, \varphi(r_k)\big) \\ &\le \tfrac{1}{j+1} + \Delta\big(P^{(j)}, P^{(k)}\big) + \tfrac{1}{k+1}. \end{aligned}$$

using that φ preserves distance (part (d) below). Given $\varepsilon > 0$, take K with $\frac{1}{K+1} < \frac{\varepsilon}{3}$ and, by the Cauchyness of $\{P^{(k)}\}$, an index M with $\Delta(P^{(j)}, P^{(k)}) < \frac{\varepsilon}{3}$ for $j, k \geq M$; then $d(r_j, r_k) < \varepsilon$ for all $j, k \geq \max(K, M)$. So $\{r_k\}$ is Cauchy in X (5.4.1), and we let $P \in X^*$ be its equivalence class.

Finally $\Delta(P^{(k)}, P) \to 0$. By the triangle inequality in X^*,

$$\Delta\big(P^{(k)}, P\big) \leq \Delta\big(P^{(k)}, \varphi(r_k)\big) + \Delta\big(\varphi(r_k), P\big) \leq \tfrac{1}{k+1} + \Delta\big(\varphi(r_k), P\big).$$

The second term is $\lim_j d(r_k, r_j)$, and the Cauchy estimate for $\{r_k\}$ makes this small: given $\varepsilon > 0$, for k large enough $d(r_k, r_j) < \varepsilon$ for all large j, so $\Delta(\varphi(r_k), P) \leq \varepsilon$. Hence $\Delta(P^{(k)}, P) \to 0$, every Cauchy sequence in X^* converges, and X^* is complete (5.4.3; a complete space is one in which Cauchy sequences converge).

(d) φ is an isometry. For $p \in X$ the constant sequence $(p, p, p, \dots)$ is Cauchy, and $P_p = \varphi(p)$ is its class. For $p, q \in X$ the representatives are the constant sequences, so

$$\Delta(P_p, P_q) = \lim_n d(p, q) = d(p, q).$$

Thus φ preserves distance; in particular $\varphi(p) = \varphi(q)$ forces $d(p, q) = 0$, i.e. $p = q$, so φ is injective and is an isometry of X into X^*.

(e) $\varphi(X)$ is dense, and equals X^* when X is complete. Let $P \in X^*$ with representative $\{p_n\}$, and fix $\varepsilon > 0$. Being Cauchy, $\{p_n\}$ has an index N with $d(p_m, p_n) < \varepsilon$ for all $m, n \geq N$. Then $\varphi(p_N)$ approximates P:

$$\Delta\big(P, \varphi(p_N)\big) = \lim_n d(p_n, p_N) \leq \varepsilon,$$

since $d(p_n, p_N) < \varepsilon$ for all $n \geq N$ and the bound survives the limit (6.3.1). Every ball about P thus meets $\varphi(X)$, so $\varphi(X)$ is dense in X^*.

Suppose now X is complete and let $P \in X^*$ have representative $\{p_n\}$. As a Cauchy sequence in a complete space, $\{p_n\}$ converges to some $p \in X$ (Cauchy sequences converge in a complete space). Then $\{p_n\} \sim (p, p, \dots)$, because $d(p_n, p) \to 0$ by convergence (5.1.1); so P is the class of the constant sequence at p, that is $P = \varphi(p) \in \varphi(X)$. Hence $\varphi(X) = X^*$ when X is complete.

■

Reflections.

The exercise is, line for line, the construction our notes record as the *Cauchy completion* (5.4.4)—and reading it as a construction rather than a list of five drills is the whole orientation. The driving idea is the one from 6.1.3: a Cauchy sequence is a packet of data that "ought to" name a limit even when the space has no point to receive it, so if

X is missing points, the cleanest way to supply them is to let the Cauchy sequences *be* the points. Two Cauchy sequences that close in on each other should name the same new point, which is precisely why the problem opens by quotienting Cauchy sequences under $\lim_n d(p_n, q_n) = 0$. The words *equivalence relation*, *equivalence classes*, *distance function*, *complete*, and *isometry* are a checklist: build the set of points, put a metric on it, prove no Cauchy sequence escapes it, and show the original space sits inside undistorted.

It helps to believe the target before building it. The model case is $X = \mathbb{Q}$, where the missing points are the irrationals: the Cauchy sequence $1, 1.4, 1.41, 1.414, \ldots$ has no rational limit, yet it deserves to name $\sqrt{2}$, and any other rational sequence creeping toward the same gap is declared equivalent to it. Quotienting by "their distance tends to 0" glues exactly the sequences pointing at one hole into a single new point, and the completion $\mathbb{Q}^*$ is a copy of $\mathbb{R}$. Every abstract step below is this picture in disguise.

The five parts divide the labour, and each rests on a specific tool:

(1) the relation is reflexive and symmetric from the metric axioms, and transitive by the squeeze $0 \le d(p_n, r_n) \le d(p_n, q_n) + d(q_n, r_n) \to 0$ (6.3.3)—the triangle inequality is what lets "close to a common sequence" chain into "close to each other," so $\sim$ is genuinely an equivalence relation (0.4.16);

(2) before Δ can be a metric it must be *well defined on classes*, the recurring worry whenever a rule is stated through representatives (0.4.20); the four-term triangle estimate $|d(p_n, q_n) - d(p'_n, q'_n)| \le d(p_n, p'_n) + d(q_n, q'_n)$ shows swapping to equivalent representatives moves the limit by 0, and the metric axioms (2.2.1) then pass from d to Δ in the limit by the order properties of limits (6.3.1), with $\Delta(P, Q) = 0$ decoding to $P = Q$ exactly because $\sim$ was defined by vanishing distance;

(3) completeness is the heart, and the move is a diagonal one: from the k-th representative pull a single term $r_k = p^{(k)}_{N_k}$ chosen past the point where that sequence has settled to within $\frac{1}{k+1}$, so that $\varphi(r_k)$ sits within $\frac{1}{k+1}$ of $P^{(k)}$; the diagonal $\{r_k\}$ is then Cauchy in X (5.4.1), its class P is a point of X^*, and $P^{(k)} \to P$ because $P^{(k)}$ is squeezed between itself and P through the nearby $\varphi(r_k)$;

(4) the isometry is immediate—constant representatives give $\Delta(P_p, P_q) = \lim_n d(p, q) = d(p, q)$ directly—and distance preservation forces injectivity, which is the one place φ being an embedding (not merely a map) is used;

(5) density is the same approximation as (3) read backwards: a representative $\{p_n\}$ of P is Cauchy, so a late term p_N has $\Delta(P, \varphi(p_N)) = \lim_n d(p_n, p_N) \le \varepsilon$, putting a point of $\varphi(X)$ in every ball about P; and if X is already complete the representative *converges* in X (Cauchy sequences converge in a complete space), so its class is already a constant class and nothing new is added.

The argument earns the name "completion" only if no step quietly assumes what it is proving, and the delicate point is (3). The limit class P is not posited; it is built from the diagonal sequence $\{r_k\}$, which lives entirely inside the *original* X, so its Cauchyness is a statement about d, not about the new metric Δ—there is no circular appeal to a completeness of X^* we are trying to establish. The chosen indices N_k are exactly what converts each abstract class $P^{(k)}$ into a concrete nearby point $r_k \in X$ with controlled error $\frac{1}{k+1} \to 0$ (Archimedean property), and the triangle inequality, used through the isometry $d(r_j, r_k) = \Delta(\varphi(r_j), \varphi(r_k))$, transfers the Cauchyness of $\{P^{(k)}\}$ to that of $\{r_k\}$. Each hypothesis is load-bearing: drop "$\{p_n\}$ Cauchy" from the definition and the representative-distance of Problem 3.23 need not converge, so Δ would not even be a number; the constant sequences are Cauchy automatically, which is why φ is defined on all of X.

The transferable move is the completion recipe itself, and it is worth carrying whole: to repair a space that fails completeness, take its Cauchy sequences as raw points, glue together the ones whose distance vanishes, transport the metric by a limit, and recover the old space as the constant sequences. This is the abstract engine behind building $\mathbb{R}$ from $\mathbb{Q}$ (1.6.3), now run in any metric space; the density in (e) is what makes X "almost all" of its completion, and the isometry in (d) is what lets us stop distinguishing X from its image and simply regard X as living inside the complete space X^*.

Problem 3.25 *Rudin, Chapter 3, Exercise 25*

Let X be the metric space whose points are the rational numbers, with the metric $d(x, y) = |x-y|$. What is the completion of this space? (Compare Exercise 24.)

Solution.

The completion of $(\mathbb{Q}, |\cdot|)$ is $(\mathbb{R}, |\cdot|)$.

A completion of (X, d) is a complete metric space into which X embeds isometrically as a dense subspace, and it is unique up to isometry (5.4.4). So three things must be checked for the candidate $\mathbb{R}$: that it is complete, that the inclusion $\mathbb{Q} \hookrightarrow \mathbb{R}$ is an isometry, and that $\mathbb{Q}$ is dense in $\mathbb{R}$.

The metric on $\mathbb{R}$ is the restriction of $|\cdot|$, so for $x, y \in \mathbb{Q}$ the distance $|x - y|$ is the same whether computed in X or in $\mathbb{R}$; the inclusion map is therefore an isometry, and $\mathbb{Q}$ sits inside $\mathbb{R}$ as a subspace.

$\mathbb{R}$ is complete: every Cauchy sequence of real numbers converges, since $\mathbb{R} = \mathbb{R}^1$ is complete ($\mathbb{R}^k$ is complete; 5.5.3).

$\mathbb{Q}$ is dense in $\mathbb{R}$: every real number is a limit of rationals. Given $x \in \mathbb{R}$, density of $\mathbb{Q}$ ($\mathbb{Q}$ is dense in $\mathbb{R}$; 1.7.2) supplies, for each $n \geq 1$, a rational q_n with $|q_n - x| < \frac{1}{n}$, so every ball $B_{1/n}(x)$ meets $\mathbb{Q}$; hence $q_n \to x$ and $x \in \overline{\mathbb{Q}}$, and since x was arbitrary $\overline{\mathbb{Q}} = \mathbb{R}$.

Thus $\mathbb{R}$ is a complete metric space containing $\mathbb{Q}$ isometrically as a dense subspace, which

is exactly a completion of X. By the essential uniqueness in 5.4.4, any completion of $\mathbb{Q}$ is isometric to $\mathbb{R}$, so the completion is $\mathbb{R}$ with the metric $|x-y|$.

The abstract completion of Exercise 24 produces the same space in disguise. There $\widehat{X}$ is the set of Cauchy sequences of rationals modulo $\{p_n\} \sim \{q_n\} \iff |p_n - q_n| \to 0$, with $\widehat{d}([p_n],[q_n]) = \lim_n |p_n - q_n|$, and each $q \in \mathbb{Q}$ embeds as the class of the constant sequence. The map $[p_n] \mapsto \lim_n p_n \in \mathbb{R}$ is a well-defined bijection onto $\mathbb{R}$: the limit exists because $\{p_n\}$ is Cauchy in the complete space $\mathbb{R}$ (Cauchy sequences in $\mathbb{R}$ converge); it is independent of the representative because $|p_n - q_n| \to 0$ forces equal limits; it preserves distance, since $\widehat{d}([p_n],[q_n]) = \lim_n |p_n - q_n| = |\lim_n p_n - \lim_n q_n|$; and it is onto because every real is the limit of some rational sequence, by the density just shown. So Exercise 24's $\widehat{X}$ *is* $\mathbb{R}$.

■

REFLECTIONS.

The word that fixes everything is *completion*, and our notes give it a single precise meaning: a complete space into which X embeds isometrically as a *dense* subspace, unique up to isometry (5.4.4). Reading that definition is already most of the work, because it tells you that to *identify* a completion you do not have to build one—you only have to recognise a familiar space that meets the three conditions, and uniqueness does the rest. The parenthetical "compare Exercise 24" is the second clue: that exercise constructs the completion abstractly out of Cauchy sequences, so Exercise 25 is asking you to put a name to the object Exercise 24 manufactures.

The reason $\mathbb{R}$ is the right name is the relationship our notes record between $\mathbb{Q}$ and $\mathbb{R}$: $\mathbb{R}$ is the complete ordered field in which $\mathbb{Q}$ is dense (1.6.2), and $\mathbb{Q}$ is famously *not* complete—a sequence of rationals approaching $\sqrt{2}$ is Cauchy with no rational limit (6.1.2). That single failure is the whole motive for completing $\mathbb{Q}$ in the first place: completion is the operation that supplies the missing limits, and $\mathbb{R}$ is precisely $\mathbb{Q}$ with those limits filled in (1.6.1).

Believe it on the canonical hole first. The decimal truncations 1, 1.4, 1.41, 1.414, ... are rationals, they bunch together (Cauchy), yet inside $\mathbb{Q}$ they converge to nothing. Pass to $\mathbb{R}$ and the same sequence acquires the limit $\sqrt{2}$. The completion is exactly the act of adjoining one new point for each such rational Cauchy sequence that had nowhere to land.

The verification then walks the three clauses of the definition in turn, each a short citation:

(1) the inclusion $\mathbb{Q} \hookrightarrow \mathbb{R}$ is an isometry, because the metric on $\mathbb{R}$ restricts to the very same $|x-y|$ on rationals—there is nothing to compute, only to notice that the distance does not change when the ambient space grows;

(2) $\mathbb{R}$ is complete ($\mathbb{R}^k$ is complete; 5.5.3), which is the one substantive theorem in play and the reason a candidate exists at all—an incomplete target could never be a completion;

(3) $\mathbb{Q}$ is dense in $\mathbb{R}$ ($\mathbb{Q}$ is dense; 1.7.2): every real x has rationals $q_n \to x$ with $|q_n - x| < \frac{1}{n}$, so $x \in \overline{\mathbb{Q}}$ and the closure is all of $\mathbb{R}$.

With the three conditions met, the essential-uniqueness half of 5.4.4 converts "$\mathbb{R}$ *is a* completion" into "$\mathbb{R}$ *is the* completion"—the step that makes the answer a clean identification rather than merely an example.

The rigor check is to confirm all three clauses are genuinely load-bearing, since dropping any one breaks the identification. Without density, $\mathbb{R}^2$ (with $\mathbb{Q}$ on the first axis) would be a complete space containing $\mathbb{Q}$ isometrically, but it is far too large to be *the* completion; density is what forbids spurious extra points. Without completeness, $\mathbb{Q}$ itself trivially contains $\mathbb{Q}$ densely, but it is the very space we set out to repair. And the isometry clause is what pins the metric, not just the set: the completion of $\mathbb{Q}$ depends on which metric it carries—this is the force of the cross-reference to Exercise 24 and to other metrics on $\mathbb{Q}$ (under a p-adic metric the same set $\mathbb{Q}$ completes to a different space entirely). Matching the abstract construction of Exercise 24 to $\mathbb{R}$ by the map $[p_n] \mapsto \lim_n p_n$ closes the loop without circularity: that limit exists precisely because $\mathbb{R}$ is already known complete (Cauchy sequences in $\mathbb{R}$ converge), so the concrete answer is consistent with the abstract one rather than assumed from it.

The transferable move is to treat "what is the completion / closure / compactification of X?" as a *recognition* problem governed by a uniqueness theorem: find any concrete space satisfying the defining clauses—here complete, isometric, dense—and uniqueness promotes that example to the answer. It is also the cleanest way to see what $\mathbb{R}$ *is*: not a list of axioms but the metric completion of $\mathbb{Q}$, the smallest complete space holding the rationals densely, which is the construction 5.4.4 says is hiding inside the passage from $\mathbb{Q}$ to $\mathbb{R}$ in Lecture 1.

Rudin Chapter 4 – Continuity

Problem 4.1 *Rudin, Chapter 4, Exercise 1*

Suppose f is a real function defined on R^1 which satisfies

$$\lim_{h\to 0}[f(x+h)-f(x-h)]=0$$

for every $x \in R^1$. Does this imply that f is continuous?

SOLUTION.

No. Define f on R^1 by $f(0)=1$ and $f(x)=0$ for $x \neq 0$. Fix x and take $h \neq 0$ small: if $x \neq 0$ then for $|h|<|x|$ both $x+h$ and $x-h$ avoid 0, while if $x=0$ then $x \pm h = \pm h \neq 0$. In either case $f(x+h)=f(x-h)=0$, so $f(x+h)-f(x-h)=0$ and the limit is 0 for every x. Yet f is not continuous at 0: $f(x)\to 0$ as $x\to 0$, whereas $f(0)=1$ (R 4.5; 7.2.1). So the condition does not imply continuity.

■

REFLECTIONS.

The expression $f(x+h)-f(x-h)$ only compares f at points placed symmetrically about x, and it never involves $f(x)$ itself. So it measures whether f is becoming balanced around x, not whether it is approaching the right value there. A lone spike — a function equal to 0 everywhere except for one raised point — is perfectly balanced around every x, since once h is small both sample points $x \pm h$ sit at height 0, so the symmetric limit vanishes everywhere even though the spike breaks continuity at that one point. The condition is genuinely weaker than continuity: it is a statement about symmetry, while continuity is the statement that $f(x)$ matches the values nearby.

Problem 4.2 *Rudin, Chapter 4, Exercise 2*

If f is a continuous mapping of a metric space X into a metric space Y, prove that

$$f(\overline{E}) \subset \overline{f(E)}$$

for every set $E \subset X$. ($\overline{E}$ denotes the closure of E.) Show, by an example, that $f(\overline{E})$ can be a proper subset of $\overline{f(E)}$.

SOLUTION.

Let $y \in f(\overline{E})$, say $y=f(x)$ with $x \in \overline{E}$. There is a sequence $\{x_n\}$ in E with $x_n \to x$ (take $x_n = x$ if $x \in E$, or a sequence in E approaching x if x is a limit point of E). By continuity $f(x_n)\to f(x)=y$ (R 4.2; R 4.6), and each $f(x_n) \in f(E)$, so y is a limit of points of $f(E)$, i.e. $y \in \overline{f(E)}$. Hence $f(\overline{E}) \subset \overline{f(E)}$.

For a proper inclusion, take $X=Y=R^1$, $f(x)=1/(1+x^2)$, and $E=\{1,2,3,\dots\}$. Then

E is closed, so $\overline{E} = E$ and $f(\overline{E}) = f(E) = \{1/(1+n^2) : n \geq 1\}$. Since $1/(1+n^2) \to 0$, the point 0 lies in $\overline{f(E)}$ but not in $f(E)$; thus $f(\overline{E})$ is a proper subset of $\overline{f(E)}$.

■

Reflections.

Closure adds the points reachable as limits from a set, and continuity is exactly the property that carries limits across the map: $x_n \to x$ forces $f(x_n) \to f(x)$. So a value $f(x)$ with x in the closure of E is automatically a limit of values coming from E — that is the whole inclusion. It can be *proper* because $\overline{f(E)}$ may acquire limit points that no point of $\overline{E}$ produces: in the example the outputs $1/(1+n^2)$ drift down toward 0, so 0 sits in the closure of the image, yet nothing in E (whose closure is E again, as E is closed) maps to 0. The map is continuous, but "0 at infinity" is a limit of outputs without being an output.

Problem 4.3 *Rudin, Chapter 4, Exercise 3*

Let f be a continuous real function on a metric space X. Let $Z(f)$ (the *zero set* of f) be the set of all $p \in X$ at which $f(p) = 0$. Prove that $Z(f)$ is closed.

Solution.

Let p be a limit point of $Z(f)$ and take $p_n \in Z(f)$ with $p_n \to p$. By continuity, $f(p_n) \to f(p)$ (R 4.6); but $f(p_n) = 0$ for all n, so $f(p) = 0$, i.e. $p \in Z(f)$. Thus $Z(f)$ contains all of its limit points and is closed (R 2.18). Equivalently, $Z(f) = f^{-1}(\{0\})$ is the preimage of the closed set $\{0\} \subset R^1$ under a continuous map, hence closed (R 4.8).

■

Reflections.

A zero set is a level set, $Z(f) = f^{-1}(\{0\})$, and the clean way to see it is closed is the general principle that continuous preimages of closed sets are closed. Concretely: if the points where f vanishes pile up at some p, continuity forces the value at p to be the limit of those zeros, namely 0, so p is itself a zero and the set cannot shed a boundary point. This is the same pattern by which closedness and openness of target sets pull back through continuous maps, which is the content of the open-set characterization of continuity (7.4.3).

Problem 4.4 *Rudin, Chapter 4, Exercise 4*

Let f and g be continuous mappings of a metric space X into a metric space Y, and let E be a dense subset of X. Prove that $f(E)$ is dense in $f(X)$. If $g(p) = f(p)$ for all $p \in E$, prove that $g(p) = f(p)$ for all $p \in X$. (In other words, a continuous mapping is determined by its values on a dense subset of its domain.)

Solution.

First, $f(E)$ is dense in $f(X)$. Let $y \in f(X)$, say $y = f(x)$. Since E is dense, there is a sequence $\{e_n\}$ in E with $e_n \to x$, and by continuity $f(e_n) \to f(x) = y$ (R 4.6; 7.2.1). Thus

y is a limit of points of $f(E)$, so every point of $f(X)$ lies in $\overline{f(E)}$, i.e. $f(E)$ is dense in $f(X)$.

Now suppose $g(p) = f(p)$ for all $p \in E$. Given $x \in X$, choose $e_n \in E$ with $e_n \to x$. Then $f(e_n) = g(e_n)$ for every n, while $f(e_n) \to f(x)$ and $g(e_n) \to g(x)$ by continuity. Limits in a metric space are unique (R 3.2(b)), so $f(x) = g(x)$. Hence $f = g$ on all of X.

■

REFLECTIONS.

A continuous map is completely determined by its values on a dense set, for the simple reason that every point of the space is a limit of dense-set points and continuity transports that limit. To compare two continuous maps that agree on a dense E, run both along a sequence from E approaching an arbitrary x: the two maps give identical values all along the sequence, each value-sequence converges to the map's value at x, and uniqueness of limits forces those two values to agree. The density-of-image statement is the same machinery read forward — the outputs over E already crowd against every output of the map. This is why a continuous function on R^1 is fixed once you know it on the rationals.

Problem 4.5 *Rudin, Chapter 4, Exercise 5*

If f is a real continuous function defined on a closed set $E \subset R^1$, prove that there exist continuous real functions g on R^1 such that $g(x) = f(x)$ for all $x \in E$. (Such functions g are called *continuous extensions* of f from E to R^1.) Show that the result becomes false if the word "closed" is omitted. Extend the result to vector-valued functions. *Hint:* Let the graph of g be a straight line on each of the segments which constitute the complement of E (compare Exercise 29, Chap. 2). The result remains true if R^1 is replaced by any metric space, but the proof is not so simple.

SOLUTION.

Since E is closed, $R^1 \setminus E$ is open, hence a countable union of disjoint open intervals (R 2.29). Define $g = f$ on E; on each bounded complementary interval (a, b) — whose endpoints a, b lie in E — let g be the line joining $(a, f(a))$ to $(b, f(b))$,

$$g(x) = f(a) + \frac{f(b) - f(a)}{b - a}\,(x - a) \qquad (a \le x \le b);$$

on an unbounded complementary interval $(-\infty, b)$ or (a, ∞) set $g \equiv f(b)$ or $g \equiv f(a)$, respectively.

Then g is real-valued on R^1 and equals f on E. It is continuous on $R^1 \setminus E$ (linear or constant near each such point) and at every interior point of E (where $g = f$ near the point). The one case to check is a point $p \in E$ that is an endpoint of complementary intervals. Given $\varepsilon > 0$, continuity of f at p gives $\delta > 0$ with $|f(q) - f(p)| < \varepsilon$ for $q \in E$ with $|q - p| < \delta$. If $|x - p| < \delta$, then x lies in E or in a complementary interval whose

endpoints are within δ of p; in the latter case $g(x)$ lies between the two endpoint values $f(\cdot)$, each within ε of $f(p)$. Either way $|g(x) - f(p)| < \varepsilon$, so g is continuous at p. Thus g is a continuous extension of f.

If "closed" is omitted the result fails: on the non-closed set $E = (0, 1)$ the function $f(x) = 1/x$ is continuous but unbounded near 0, so no function continuous on R^1 — being locally bounded — can agree with it on E.

For vector-valued $f = (f_1, \ldots, f_k) : E \to R^k$, extend each component f_i as above to g_i; then $g = (g_1, \ldots, g_k)$ is continuous and extends f.

■

Reflections.

A closed subset of the line is the whole line with a family of open "gaps" cut out, and those gaps are disjoint intervals. The extension simply connects the dots: across each finite gap draw the straight segment between the two known endpoint heights, and beyond an infinite gap hold the value flat. Nothing can go wrong inside a gap, where g is a line, or well inside E, where g is just f; the only sensitive place is a point of E at the mouth of a gap, and there the interpolated values are trapped between nearby values of f, so f's own continuity pulls them to the right limit. Closedness is precisely what forces the gap endpoints to belong to E, so f provides heights to interpolate between — drop it and the function can escape to infinity at a missing endpoint, as $1/x$ does on $(0, 1)$, leaving nothing to extend. The argument is one-dimensional, resting on the gap structure of open sets in R^1, which is why the general metric-space version takes more work.

Problem 4.6 *Rudin, Chapter 4, Exercise 6*

If f is defined on E, the *graph* of f is the set of points $(x, f(x))$, for $x \in E$. In particular, if E is a set of real numbers, and f is real-valued, the graph of f is a subset of the plane.
Suppose E is compact, and prove that f is continuous on E if and only if its graph is compact.

Solution.

Suppose first that f is continuous. The map $\Phi : E \to E \times f(E)$, $\Phi(x) = (x, f(x))$, has continuous coordinate functions and so is continuous, and the graph is $\Phi(E)$. As the continuous image of the compact set E, the graph is compact (R 4.14; 7.5.1).

Conversely, suppose the graph $G = \{(x, f(x)) : x \in E\}$ is compact, and let $x_n \to x$ in E; we show $f(x_n) \to f(x)$. If not, then $d_Y(f(x_{n_k}), f(x)) \geq \varepsilon$ for some $\varepsilon > 0$ and some subsequence. The points $(x_{n_k}, f(x_{n_k}))$ lie in the compact set G, so a further subsequence converges to a point $(a, b) \in G$, where $b = f(a)$ (R 3.6(a)). Its first coordinate is $\lim x_{n_k} = x$, so $a = x$ and $b = f(x)$; but then $f(x_{n_{k_j}}) \to b = f(x)$, contradicting $d_Y(f(x_{n_k}), f(x)) \geq \varepsilon$. Hence $f(x_n) \to f(x)$, and f is continuous (R 4.6).

■

REFLECTIONS.

The graph is a relabeled copy of the domain sitting up in the product $X \times Y$. One direction is the packaging fact that a map into a product is continuous exactly when its coordinates are, so $x \mapsto (x, f(x))$ is continuous and sends the compact domain to a compact graph. The other direction is where compactness earns its keep: if f tried to jump, you could follow a sequence on the graph whose heights stay ε away from the target, yet compactness of the graph forces a subsequence to settle onto an actual graph point; since the base coordinates head to x, that limit can only be $(x, f(x))$, which drags the heights back to $f(x)$ and contradicts the jump. The compactness of E is essential: it is what supplies the convergent subsequence that traps the limit on the graph.

Problem 4.7 *Rudin, Chapter 4, Exercise 7*

If $E \subset X$ and if f is a function defined on X, the *restriction* of f to E is the function g whose domain of definition is E, such that $g(p) = f(p)$ for $p \in E$. Define f and g on R^2 by: $f(0,0) = g(0,0) = 0$, $f(x,y) = xy^2/(x^2+y^4)$, $g(x,y) = xy^2/(x^2+y^6)$ if $(x,y) \neq (0,0)$. Prove that f is bounded on R^2, that g is unbounded in every neighborhood of $(0,0)$, and that f is not continuous at $(0,0)$; nevertheless, the restrictions of both f and g to every straight line in R^2 are continuous!

SOLUTION.

f is bounded: if $x = 0$ or $y = 0$ (in particular at the origin) then $f(x,y) = 0$. Otherwise $|x|\,y^2 > 0$, and $(|x| - y^2)^2 \geq 0$ gives $x^2 + y^4 \geq 2|x|y^2$, so

$$|f(x,y)| = \frac{|x|\,y^2}{x^2+y^4} \leq \frac{|x|\,y^2}{2|x|\,y^2} = \frac{1}{2}.$$

Hence $|f| \leq \frac{1}{2}$ on R^2.

g is unbounded in every neighborhood of the origin: along the curve $x = y^3$ $(y \neq 0)$,

$$g(y^3, y) = \frac{y^3 \cdot y^2}{y^6 + y^6} = \frac{y^5}{2y^6} = \frac{1}{2y} \xrightarrow[y \to 0]{} \infty,$$

and such points occur in every neighborhood of $(0,0)$.

f is not continuous at the origin: along $x = y^2$ $(y \neq 0)$, $f(y^2, y) = y^4/(y^4+y^4) = \frac{1}{2}$, so $f \to \frac{1}{2}$ along this approach while $f(0,0) = 0$; the limit does not equal the value (R 4.5; 7.3.6).

Finally, the restriction to any straight line is continuous. Away from the origin both f and g are continuous, being quotients with nonvanishing denominators, so only a line through the origin needs checking. Parametrize it as $(x,y) = (at, bt)$, $t \in R$, with $(a,b) \neq (0,0)$.

For $t \neq 0$,

$$f(at, bt) = \frac{ab^2 t}{a^2 + b^4 t^2}, \qquad g(at, bt) = \frac{ab^2 t}{a^2 + b^6 t^4},$$

and if $a = 0$ the line is the y-axis, on which $f = g = 0$. In every case the value $\to 0 = f(0,0) = g(0,0)$ as $t \to 0$, so each restriction is continuous.

■

Reflections.

The surprise is that being continuous along every straight line through a point is much weaker than being continuous there. Both functions are built so that any *linear* approach to the origin kills the numerator faster than the denominator — on a line $x = at$, $y = bt$ the ratio behaves like a constant times t, which runs to 0 — so every line sees a continuous function taking the value 0 at the origin. But continuity at a point demands that *all* approaches agree, and the denominators $x^2 + y^4$ and $x^2 + y^6$ are tuned to special curved approaches: along the parabola $x = y^2$ the two terms of f's denominator balance and f holds at $\frac{1}{2}$, while along $x = y^3$ the function g even blows up. So testing continuity in several variables means controlling every path into the point, not just the straight ones; checking lines can confirm a limit only once you already know the limit exists.

Problem 4.8 *Rudin, Chapter 4, Exercise 8*

Let f be a real uniformly continuous function on the bounded set E in R^1. Prove that f is bounded on E.

Show that the conclusion is false if boundedness of E is omitted from the hypothesis.

Solution.

By uniform continuity, for $\varepsilon = 1$ there is $\delta > 0$ with $|f(x) - f(y)| < 1$ whenever $x, y \in E$ and $|x - y| < \delta$ (R 4.18). Since E is bounded, $E \subset [a, b]$ for some $a < b$. Partition $[a, b]$ into finitely many subintervals $I_1, \ldots, I_N$ each of length $< \delta$, and for every k with $E \cap I_k \neq \varnothing$ fix a point $e_k \in E \cap I_k$. Any $x \in E$ lies in some such I_k, where $|x - e_k| < \delta$, so $|f(x)| \leq |f(e_k)| + 1 \leq \max_k |f(e_k)| + 1$. The right-hand side is a finite bound independent of x, so f is bounded on E.

If boundedness of E is dropped, the conclusion fails: $f(x) = x$ on $E = R^1$ is uniformly continuous (take $\delta = \varepsilon$) yet unbounded.

■

Reflections.

Uniform continuity hands you one δ that works at every point at once, and that uniformity is exactly what converts a bounded domain into a bound on f. Tile the bounded set with finitely many pieces, each narrower than δ; across any single piece f can change by less than a fixed amount, so starting from the finitely many sample values it can climb only a bounded distance and has no room to run to infinity. Finiteness of the tiling is where

boundedness of the domain is used; on an unbounded domain there are infinitely many pieces and nothing stops a uniformly continuous function from growing without end, as $f(x) = x$ shows. Ordinary continuity would not be enough either, since its δ could shrink from piece to piece and the tiling argument would fall apart.

Problem 4.9 *Rudin, Chapter 4, Exercise 9*

Show that the requirement in the definition of uniform continuity can be rephrased as follows, in terms of diameters of sets: To every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\operatorname{diam} f(E) < \varepsilon$ for all $E \subset X$ with $\operatorname{diam} E < \delta$.

SOLUTION.

By definition, f is uniformly continuous if for every $\varepsilon > 0$ there is $\delta > 0$ with $d_Y(f(p), f(q)) < \varepsilon$ whenever $d_X(p, q) < \delta$ (R 4.18), and $\operatorname{diam} A = \sup_{p,q \in A} d(p, q)$ (R 3.9).

Suppose f is uniformly continuous, and let $\varepsilon > 0$. Choose $\delta > 0$ so that $d_X(p, q) < \delta$ implies $d_Y(f(p), f(q)) < \varepsilon/2$. If $A \subset X$ has $\operatorname{diam} A < \delta$, then any $p, q \in A$ satisfy $d_X(p, q) < \delta$, hence $d_Y(f(p), f(q)) < \varepsilon/2$; taking the supremum over $p, q \in A$ gives $\operatorname{diam} f(A) \leq \varepsilon/2 < \varepsilon$.

Conversely, suppose the diameter condition holds, and let $\varepsilon > 0$; take the corresponding δ. For any p, q with $d_X(p, q) < \delta$, the two-point set $A = \{p, q\}$ has $\operatorname{diam} A = d_X(p, q) < \delta$, so $d_Y(f(p), f(q)) = \operatorname{diam} f(A) < \varepsilon$. Hence f is uniformly continuous, and the two formulations coincide.

■

REFLECTIONS.

This is the same definition in different dress: uniform continuity says small input distances force small output distances, with a threshold δ that does not depend on where you are. Stating it through diameters makes that location-independence vivid — "every set small enough, of diameter $< \delta$, has a small image, of diameter $< \varepsilon$" speaks about all sets at once, with no basepoint singled out. To recover the ordinary form, feed in the smallest interesting set, the pair $\{p, q\}$, whose diameter is exactly $d(p, q)$. The set-level phrasing is useful precisely because it talks about whole sets rather than points, which is what streamlines arguments like the extension theorem, where one follows the images of shrinking neighborhoods.

Problem 4.10 *Rudin, Chapter 4, Exercise 10*

Complete the details of the following alternative proof of Theorem 4.19: If f is not uniformly continuous, then for some $\varepsilon > 0$ there are sequences $\{p_n\}$, $\{q_n\}$ in X such that $d_X(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) > \varepsilon$. Use Theorem 2.37 to obtain a contradiction.

SOLUTION.

Suppose f is not uniformly continuous. Negating the definition, there is $\varepsilon > 0$ such that for every $\delta > 0$ some pair $p, q \in X$ has $d_X(p, q) < \delta$ yet $d_Y(f(p), f(q)) > \varepsilon$. Taking $\delta = 1/n$ produces sequences $\{p_n\}, \{q_n\}$ in X with $d_X(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) > \varepsilon$ for all n.

Since X is compact, $\{p_n\}$ has a subsequence $p_{n_k} \to p$ for some $p \in X$ (R 3.6(a), which rests on R 2.37). As $d_X(p_n, q_n) \to 0$, also $q_{n_k} \to p$, because $d_X(q_{n_k}, p) \le d_X(q_{n_k}, p_{n_k}) + d_X(p_{n_k}, p) \to 0$. By continuity of f at p, both $f(p_{n_k}) \to f(p)$ and $f(q_{n_k}) \to f(p)$ (R 4.6), so

$$d_Y(f(p_{n_k}), f(q_{n_k})) \le d_Y(f(p_{n_k}), f(p)) + d_Y(f(p), f(q_{n_k})) \to 0,$$

contradicting $d_Y(f(p_{n_k}), f(q_{n_k})) > \varepsilon$. Hence f is uniformly continuous (R 4.19).

■

Reflections.

This is the sequential route to the fact that continuity becomes uniform on a compact space. The failure of uniform continuity is really a statement about two moving points: however close you require the inputs to be, the outputs still refuse to come within ε. Record that refusal as two sequences p_n, q_n that close in on each other while their images stay ε apart. Compactness then supplies a single point p that a subsequence of the p_n approaches, and the q_n are dragged to the same p since they shadow the p_n; continuity at that one point pulls both image-streams to $f(p)$, so the images cannot stay apart, and the contradiction closes. Compactness does the essential work here by guaranteeing the common limit point that ties the two sequences together — on a noncompact space the construction survives and uniform continuity can genuinely fail.

Problem 4.11 *Rudin, Chapter 4, Exercise 11*

Suppose f is a uniformly continuous mapping of a metric space X into a metric space Y and prove that $\{f(x_n)\}$ is a Cauchy sequence in Y for every Cauchy sequence $\{x_n\}$ in X. Use this result to give an alternative proof of the theorem stated in Exercise 13.

Solution.

Let $\{x_n\}$ be a Cauchy sequence in X. Given $\varepsilon > 0$, uniform continuity provides $\delta > 0$ with $d_Y(f(x), f(y)) < \varepsilon$ whenever $d_X(x, y) < \delta$ (R 4.18). Since $\{x_n\}$ is Cauchy, there is N with $d_X(x_n, x_m) < \delta$ for all $n, m \ge N$, and then $d_Y(f(x_n), f(x_m)) < \varepsilon$. Thus $\{f(x_n)\}$ is Cauchy in Y.

This yields the extension theorem of Exercise 13 whenever Y is complete: if E is dense in X and $f : E \to Y$ is uniformly continuous, then for $p \in X$ pick $e_n \in E$ with $e_n \to p$. The sequence $\{e_n\}$ is Cauchy, so $\{f(e_n)\}$ is Cauchy and hence converges in the complete space Y; set $g(p) = \lim_n f(e_n)$. This limit does not depend on the approximating sequence (interleaving two sequences tending to p shows their images share a limit), g agrees with

f on E, and g is continuous — so g is the desired extension, with the estimates carried out in detail in Problem 13.

■

REFLECTIONS.

Uniform continuity is exactly the strength needed to carry Cauchy sequences to Cauchy sequences, and ordinary continuity is not: $f(x) = 1/x$ is continuous on $(0, 1)$ and sends the Cauchy sequence $1/n$ to n, which is not Cauchy. What makes uniform continuity succeed is its single location-free δ — once the inputs are eventually within δ of one another, the outputs are within ε, with no dependence on where the sequence is heading. This is the property that lets a uniformly continuous function be extended from a dense set to the whole space: a point outside the set is the limit of a Cauchy sequence from inside, its images form a Cauchy sequence, and in a complete range that sequence has a limit to serve as the extended value.

Problem 4.12 *Rudin, Chapter 4, Exercise 12*

A uniformly continuous function of a uniformly continuous function is uniformly continuous. State this more precisely and prove it.

SOLUTION.

Precisely: if $f : X \to Y$ and $g : Y \to Z$ are uniformly continuous mappings of metric spaces, then $g \circ f : X \to Z$ is uniformly continuous.

Let $\varepsilon > 0$. By uniform continuity of g there is $\eta > 0$ with $d_Z(g(y), g(y')) < \varepsilon$ whenever $d_Y(y, y') < \eta$. By uniform continuity of f there is $\delta > 0$ with $d_Y(f(x), f(x')) < \eta$ whenever $d_X(x, x') < \delta$ (R 4.18). Then $d_X(x, x') < \delta$ gives $d_Y(f(x), f(x')) < \eta$, which gives $d_Z(g(f(x)), g(f(x'))) < \varepsilon$. Hence $g \circ f$ is uniformly continuous.

■

REFLECTIONS.

The proof is just chaining the two moduli of continuity in the right order: begin with the target tolerance ε, let g turn it into a tolerance η on its inputs, then let f turn η into a tolerance δ on its inputs. What makes the conclusion *uniform* is that at each link the tolerance depends only on the previous tolerance and never on the location, so the final δ depends on ε alone. The identical chaining proves that a composition of merely continuous functions is continuous (R 4.7); there the tolerances may depend on the point, and the argument is simply run one point at a time.

Problem 4.13 *Rudin, Chapter 4, Exercise 13*

Let E be a dense subset of a metric space X, and let f be a uniformly continuous *real* function defined on E. Prove that f has a continuous extension from E to X (see Exercise 5 for terminology). (Uniqueness follows from Exercise 4.) *Hint:* For each $p \in X$ and each positive integer n,

let $V_n(p)$ be the set of all $q \in E$ with $d(p,q) < 1/n$. Use Exercise 9 to show that the intersection of the closures of the sets $f(V_1(p))$, $f(V_2(p))$, ..., consists of a single point, say $g(p)$, of R^1. Prove that the function g so defined on X is the desired extension of f.
Could the range space R^1 be replaced by R^k? By any compact metric space? By any complete metric space? By any metric space?

Solution.

Let $f : E \to R^1$ be uniformly continuous on the dense set $E \subset X$. For $p \in X$, density gives $e_n \in E$ with $e_n \to p$; being convergent, $\{e_n\}$ is Cauchy, so by Problem 11 $\{f(e_n)\}$ is Cauchy in R^1 and therefore converges (R 3.11). Define $g(p) = \lim_n f(e_n)$.

This is independent of the sequence: if $e_n \to p$ and $e'_n \to p$, the interleaving $e_1, e'_1, e_2, e'_2, \ldots$ also converges to p, hence is Cauchy, so its image is Cauchy and convergent, forcing $\lim f(e_n) = \lim f(e'_n)$. For $p \in E$ the constant sequence gives $g(p) = f(p)$, so g extends f.

To see g is uniformly continuous, let $\varepsilon > 0$ and choose $\delta > 0$ from the uniform continuity of f so that $a, b \in E$ with $d(a,b) < \delta$ give $|f(a) - f(b)| \le \varepsilon/2$. Suppose $p, p' \in X$ with $d(p,p') < \delta/3$, and take $e_n \to p$, $e'_n \to p'$. For large n, $d(e_n, p) < \delta/3$ and $d(e'_n, p') < \delta/3$, so $d(e_n, e'_n) < \delta$ and $|f(e_n) - f(e'_n)| \le \varepsilon/2$; letting $n \to \infty$ gives $|g(p) - g(p')| \le \varepsilon/2 < \varepsilon$. Thus g is uniformly continuous, and uniqueness follows from Problem 4, since two continuous extensions agree on the dense set E.

The only property of the range used is that Cauchy sequences converge there. So R^1 may be replaced by R^k, by any compact metric space, or by any complete metric space — each is complete. It cannot be replaced by an arbitrary metric space: with $X = R^1$, $E = \mathbb{Q}$, and $f : \mathbb{Q} \to \mathbb{Q}$ the identity (uniformly continuous), there is no continuous extension $g : R^1 \to \mathbb{Q}$, for a continuous map from the connected space R^1 into $\mathbb{Q}$ has connected image (R 4.22), hence is constant, while g must restrict to the identity on $\mathbb{Q}$.

■

Reflections.

The extension is forced on you: a point p outside E is approached by points e_n of the dense set, and if g is to be continuous it has no choice but to take the value $\lim f(e_n)$. The only question is whether that limit exists, and uniform continuity guarantees it — the e_n form a Cauchy sequence, uniform continuity (through Problem 11) makes $f(e_n)$ a Cauchy sequence, and in a complete range that sequence converges. This is why the range must be complete: completeness is the promise that the manufactured Cauchy sequence of values has a limit to become the new value. R^k, compact spaces, and complete spaces all keep that promise; the rationals do not, and the identity on $\mathbb{Q}$ has nowhere to send the irrational limits. Uniform rather than ordinary continuity is essential, since ordinary continuity need not preserve Cauchyness and the extension can genuinely fail to exist.

Problem 4.14 *Rudin, Chapter 4, Exercise 14*

Let $I = [0, 1]$ be the closed unit interval. Suppose f is a continuous mapping of I into I. Prove that $f(x) = x$ for at least one $x \in I$.

SOLUTION.

Define $h : I \to R^1$ by $h(x) = f(x) - x$; it is continuous as the difference of continuous functions (R 4.9). Since f maps I into $I = [0, 1]$, we have $h(0) = f(0) \geq 0$ and $h(1) = f(1) - 1 \leq 0$. If $h(0) = 0$ then 0 is a fixed point, and if $h(1) = 0$ then 1 is; otherwise $h(0) > 0 > h(1)$, and by the intermediate value theorem (R 4.23; 7.5.5) there is $x \in (0, 1)$ with $h(x) = 0$. In every case $f(x) = x$ for some $x \in I$.

■

REFLECTIONS.

This is the one-dimensional Brouwer fixed-point theorem, and the move that does the real work is to stop hunting for a fixed point and instead hunt for a zero of $h(x) = f(x) - x$. A fixed point of f is exactly a root of h, and h is easy to read at the two ends: because f cannot leave $[0, 1]$, at the left end $f(0) \geq 0$ puts h at or above zero, and at the right end $f(1) \leq 1$ puts h at or below zero. A continuous function that is nonnegative at one end and nonpositive at the other must vanish somewhere between, which is the intermediate value theorem. Geometrically, the graph of f sits in the unit square and has to cross the diagonal $y = x$, since it begins on or above the diagonal and finishes on or below it.

Problem 4.15 *Rudin, Chapter 4, Exercise 15*

Call a mapping of X into Y *open* if $f(V)$ is an open set in Y whenever V is an open set in X. Prove that every continuous open mapping of R^1 into R^1 is monotonic.

SOLUTION.

We first show f is injective. Suppose $f(x_1) = f(x_2)$ with $x_1 < x_2$. By the extreme value theorem (R 4.16), f attains a maximum and a minimum on $[x_1, x_2]$. If either is attained at an interior point $p \in (x_1, x_2)$ with $f(p) \neq f(x_1)$, then $f(p)$ is the largest (or smallest) value of f on the open interval (x_1, x_2), so $f((x_1, x_2))$ contains $f(p)$ but no value beyond it, and $f(p)$ is not an interior point of that image — contradicting that f carries the open set (x_1, x_2) to an open set. Otherwise both extrema equal $f(x_1)$, so f is constant on $[x_1, x_2]$ and $f((x_1, x_2))$ is a single point, again not open. Hence f is injective.

A continuous injective function on R^1 is strictly monotonic. Indeed, were it not, there would be $a < b < c$ with $f(b)$ not between $f(a)$ and $f(c)$; as the three values are distinct, $f(b)$ either exceeds both or lies below both. Say $f(a) < f(c) < f(b)$ (the remaining cases are symmetric). Choosing y with $f(c) < y < f(b)$, the intermediate value theorem (R 4.23; 7.5.5) yields $s \in (a, b)$ and $t \in (b, c)$ with $f(s) = f(t) = y$, and $s \neq t$ contradicts injectivity. Therefore f is monotonic.

■

Reflections.

The picture is that an open map cannot fold the line back on itself. A fold would place a local maximum or minimum at some interior point, and there the function turns around, so just past the peak the values stop climbing — the image acquires a top edge it never crosses, and a set with a top edge is not open. So openness forbids interior extrema, which is the same as saying f never repeats a value, since a repeat would force an interior extremum in between; that is, f is injective. The final step is the familiar fact that a continuous injection on an interval must be strictly monotonic, because any out-of-order triple of values would let the intermediate value theorem manufacture two inputs sharing one output. The extreme value theorem and the intermediate value theorem are the two workhorses here, one forbidding folds and the other detecting repeats.

Problem 4.16 *Rudin, Chapter 4, Exercise 16*

Let $[x]$ denote the largest integer contained in x, that is, $[x]$ is the integer such that $x-1 < [x] \leq x$; and let $(x) = x - [x]$ denote the fractional part of x. What discontinuities do the functions $[x]$ and (x) have?

Solution.

Both functions are continuous away from the integers — on each interval $[n, n+1)$ the floor $[x] = n$ is constant and $(x) = x - n$ is linear — and each has a simple discontinuity (one of the first kind) at every integer n (R 4.26).

For the floor: as $x \to n^-$ one has $[x] = n-1$, so $[\,n^-] = n-1$, while $[\,n^+] = [n] = n$. Both one-sided limits exist and differ, a jump of 1 at each integer.

For the fractional part $(x) = x - [x]$: $(n) = 0$, and as $x \to n^-$ the relation $[x] = n-1$ gives $(x) = x - (n-1) \to 1$, so $(n^-) = 1$, whereas $(n^+) = 0 = (n)$. Again both one-sided limits exist, now with a jump from 1 down to 0 at each integer. Neither function has any discontinuity other than these.

■

Reflections.

Both functions are driven by the integer count that $[x]$ performs, so they can misbehave only where that count ticks over — at the integers — and between consecutive integers each is as tame as possible, constant or a straight line of slope 1. At an integer the floor jumps up by 1, and the fractional part, being what is left after subtracting the floor, drops from 1 back down to 0: the familiar sawtooth reset. Each of these is a discontinuity of the first kind, meaning both one-sided limits exist and are finite and the function fails to be continuous only because the two sides, or the assigned value, disagree. There are no oscillatory or infinite discontinuities anywhere.

Problem 4.17 *Rudin, Chapter 4, Exercise 17*

Let f be a real function defined on (a, b). Prove that the set of points at which f has a simple discontinuity is at most countable. *Hint:* Let E be the set on which $f(x-) < f(x+)$. With each point x of E, associate a triple (p, q, r) of rational numbers such that

1. $f(x-) < p < f(x+)$,
2. $a < q < t < x$ implies $f(t) < p$,
3. $x < t < r < b$ implies $f(t) > p$.

The set of all such triples is countable. Show that each triple is associated with at most one point of E. Deal similarly with the other possible types of simple discontinuities.

SOLUTION.

At a simple discontinuity the one-sided limits $f(x-)$ and $f(x+)$ both exist (R 4.26); the possibilities are $f(x-) < f(x+)$, $f(x-) > f(x+)$, or $f(x-) = f(x+) \neq f(x)$. It is enough to show each type forms an at most countable set.

Take first the set E where $f(x-) < f(x+)$. For each $x \in E$ choose rationals p, q, r with

$$f(x-) < p < f(x+), \qquad a < q < x < r < b,$$

such that $f(t) < p$ for $q < t < x$ and $f(t) > p$ for $x < t < r$. These exist: p is any rational between the two one-sided limits, and since $f(t) \to f(x-) < p$ as $t \to x^-$ and $f(t) \to f(x+) > p$ as $t \to x^+$, the inequalities $f(t) < p$ and $f(t) > p$ hold on suitable one-sided neighborhoods, in which we place rationals q and r. The assignment $x \mapsto (p, q, r)$ is injective on E: if $x < x'$ shared the same triple, choose t with $x < t < x'$; then $t < r$ forces $f(t) > p$ (from x), while $q < t$ forces $f(t) < p$ (from x'), a contradiction. So E injects into $\mathbb{Q}^3$ and is at most countable.

The type $f(x-) > f(x+)$ is identical with the inequalities reversed. For $f(x-) = f(x+) = L \neq f(x)$, say $L < f(x)$ (the case $L > f(x)$ is symmetric), choose rationals p, q, r with $L < p < f(x)$, $a < q < x < r < b$, and $f(t) < p$ for all $t \in (q, r) \setminus \{x\}$ (possible since $f(t) \to L < p$ from both sides); if $x < x'$ shared this triple, then $x' \in (q, r) \setminus \{x\}$ would force $f(x') < p$, contradicting $f(x') > p$. Hence each type is at most countable, and the set of all simple discontinuities, a union of finitely many such sets, is at most countable.

■

REFLECTIONS.

The strategy is to give every simple discontinuity its own rational "address" and then count the addresses. At a jump with $f(x-) < f(x+)$ there is a real gap between the two one-sided limits; drop a rational p into that gap and bracket x by rationals q, r tight enough that f stays below p just to the left of x and above p just to the right. No second

discontinuity can reuse that triple, since two of them would demand both $f > p$ and $f < p$ at a shared point between them. As there are only countably many rational triples, there can be only countably many such points. The existence of the one-sided limits is exactly what opens the gap for the rational p — which is why the count works for simple discontinuities and says nothing about wilder ones, and why a monotonic function, whose discontinuities are all jumps, can have at most countably many.

Problem 4.18 *Rudin, Chapter 4, Exercise 18*

Every rational x can be written in the form $x = m/n$, where $n > 0$, and m and n are integers without any common divisors. When $x = 0$, we take $n = 1$. Consider the function f defined on R^1 by

$$f(x) = \begin{cases} 0 & (x \text{ irrational}), \\ \dfrac{1}{n} & \left(x = \dfrac{m}{n}\right). \end{cases}$$

Prove that f is continuous at every irrational point, and that f has a simple discontinuity at every rational point.

Solution.

We show $\lim_{x \to x_0} f(x) = 0$ for every $x_0 \in R^1$. Fix x_0 and $\varepsilon > 0$, and choose an integer N with $1/N < \varepsilon$. The interval $(x_0 - 1, x_0 + 1)$ contains only finitely many rationals m/n in lowest terms with $n \leq N$, since each such denominator permits only finitely many numerators in a bounded range. Let $\delta \in (0, 1)$ be less than the distance from x_0 to each of those rationals other than x_0 itself. Then for $0 < |x - x_0| < \delta$: if x is irrational, $f(x) = 0 < \varepsilon$; if $x = m/n$ in lowest terms, then x is none of the excluded rationals, so $n > N$ and $f(x) = 1/n < 1/N < \varepsilon$. Hence $\lim_{x \to x_0} f(x) = 0$.

If x_0 is irrational, then $f(x_0) = 0 = \lim_{x \to x_0} f(x)$, so f is continuous at x_0 (R 4.6). If $x_0 = m/n$ is rational, then $f(x_0) = 1/n > 0$ while both one-sided limits equal $\lim_{x \to x_0} f(x) = 0 \neq f(x_0)$, so f has a simple discontinuity (of the first kind) at x_0 (R 4.26).

■

Reflections.

This is Thomae's function, and its whole behavior rests on one counting fact: the rationals with a *small* denominator — exactly the ones where f takes a *large* value — are sparse, only finitely many in any bounded stretch. So to make f small near a point you only have to dodge those finitely many spikes, which a small enough δ does, while everything else (the irrationals, and the rationals with large denominators) already has f tiny. That forces the limit to be 0 at every point, regardless of whether the point is rational. The rational-versus-irrational split then decides continuity by the value alone: at an irrational the value is 0 and matches the limit, so f is continuous; at a rational the value $1/n$ pokes up above the limit 0, a clean removable-type discontinuity. It is the standard example

of a function continuous at every irrational and discontinuous at every rational — and a theorem of Baire says no function can do the reverse.

Problem 4.19 *Rudin, Chapter 4, Exercise 19*

Suppose f is a real function with domain R^1 which has the intermediate value property: If $f(a) < c < f(b)$, then $f(x) = c$ for some x between a and b.
Suppose also, for every rational r, that the set of all x with $f(x) = r$ is closed.
Prove that f is continuous.
Hint: If $x_n \to x_0$ but $f(x_n) > r > f(x_0)$ for some r and all n, then $f(t_n) = r$ for some t_n between x_0 and x_n; thus $t_n \to x_0$. Find a contradiction. (N. J. Fine, *Amer. Math. Monthly*, vol. 73, 1966, p. 782.)

SOLUTION.

Suppose, for contradiction, that f is discontinuous at some x_0. Then there are $\varepsilon > 0$ and a sequence $x_n \to x_0$ with $|f(x_n) - f(x_0)| \geq \varepsilon$ for all n. Passing to a subsequence we may assume $f(x_n) > f(x_0) + \varepsilon$ for all n, or else $f(x_n) < f(x_0) - \varepsilon$ for all n; take the first case (the second is symmetric).

Choose a rational r with $f(x_0) < r < f(x_0) + \varepsilon$, so that $f(x_0) < r < f(x_n)$ for every n. By the intermediate value property there is a point t_n between x_0 and x_n with $f(t_n) = r$; since t_n lies between x_0 and x_n and $x_n \to x_0$, we get $t_n \to x_0$. The points t_n all belong to $\{x : f(x) = r\}$, which is closed by hypothesis, so its limit x_0 belongs to it too (R 2.18); that is, $f(x_0) = r$. But $r > f(x_0)$, a contradiction. Hence f is continuous at every point.

■

REFLECTIONS.

Neither hypothesis alone forces continuity — the intermediate value property permits badly behaved functions, and closed level sets are easy to arrange — yet together they pin f down. The idea is to assume a jump at x_0 and watch what the intermediate value property does with it. If f leaps above $f(x_0)$ along a sequence approaching x_0, then for any rational level r just above $f(x_0)$ the function must, between x_0 and each x_n, pass through r, producing solutions of $f = r$ that march right up to x_0. Closedness of the level set $\{f = r\}$ then traps x_0 inside it, forcing $f(x_0) = r$, impossible since r sits strictly above $f(x_0)$. The rationals appear only so that countably many closed-set hypotheses are enough; the real mechanism is that the intermediate value property turns a jump into nearby solutions of $f = r$ that the closed level set refuses to release.

Problem 4.20 *Rudin, Chapter 4, Exercise 20*

If E is a nonempty subset of a metric space X, define the distance from $x \in X$ to E by

$$\rho_E(x) = \inf_{z \in E} d(x, z).$$

1. Prove that $\rho_E(x) = 0$ if and only if $x \in \overline{E}$.

2. Prove that ρ_E is a uniformly continuous function on X, by showing that

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y)$$

for all $x \in X$, $y \in X$. *Hint:* $\rho_E(x) \leq d(x, z) \leq d(x, y) + d(y, z)$, so that

$$\rho_E(x) \leq d(x, y) + \rho_E(y).$$

Solution.

1. If $\rho_E(x) = 0$, then for each n there is $z_n \in E$ with $d(x, z_n) < 1/n$, so $z_n \to x$ and $x \in \overline{E}$. Conversely, if $x \in \overline{E}$ then either $x \in E$, whence $\rho_E(x) \leq d(x, x) = 0$, or x is a limit point of E, so every ball about x meets E and the infimum $\rho_E(x) = 0$. Thus $\rho_E(x) = 0$ if and only if $x \in \overline{E}$ (R 2.18).

2. For any $z \in E$ and any $x, y \in X$, the triangle inequality gives $\rho_E(x) \leq d(x, z) \leq d(x, y) + d(y, z)$. Taking the infimum over $z \in E$ on the right,

$$\rho_E(x) \leq d(x, y) + \rho_E(y), \qquad \text{that is,} \qquad \rho_E(x) - \rho_E(y) \leq d(x, y).$$

By symmetry $\rho_E(y) - \rho_E(x) \leq d(x, y)$, so $|\rho_E(x) - \rho_E(y)| \leq d(x, y)$. Given $\varepsilon > 0$, taking $\delta = \varepsilon$ then shows ρ_E is uniformly continuous (R 4.18; 7.7.1).

■

Reflections.

The function ρ_E measures how far a point sits from the set E, and the two parts confirm the two things you would hope for. Part (a): being at distance zero from E means there are points of E arbitrarily close, which is exactly what it means to lie in the closure — so ρ_E vanishes precisely on $\overline{E}$, not merely on E. Part (b): moving from x to y can change the distance-to-E by no more than the distance you moved, because any route from x to a point of E may detour through y. That single inequality, $|\rho_E(x) - \rho_E(y)| \leq d(x, y)$, is Lipschitz continuity with constant 1, which is far stronger than plain continuity and is automatically uniform. This little function is the standard device for manufacturing continuous functions tailored to a set — it is what lets one separate disjoint closed sets and prove that a metric space is normal in the problems that follow.

Problem 4.21 *Rudin, Chapter 4, Exercise 21*

Suppose K and F are disjoint sets in a metric space X, K is compact, F is closed. Prove that there exists $\delta > 0$ such that $d(p, q) > \delta$ if $p \in K$, $q \in F$. *Hint:* ρ_F is a continuous positive function on K.

Show that the conclusion may fail for two disjoint closed sets if neither is compact.

SOLUTION.

Consider $\rho_F(p) = \inf_{q\in F} d(p,q)$ on K. By Problem 20, ρ_F is continuous and $\rho_F(p) = 0$ exactly when $p \in \overline{F} = F$. Since $K \cap F = \varnothing$, we have $\rho_F(p) > 0$ for every $p \in K$. As ρ_F is continuous on the compact set K, it attains a minimum (R 4.16): there is $p_0 \in K$ with $\rho_F(p_0) = \min_{p\in K} \rho_F(p) =: m$, and $m = \rho_F(p_0) > 0$. Put $\delta = m/2 > 0$. For any $p \in K$ and $q \in F$, $d(p,q) \geq \rho_F(p) \geq m > \delta$.

The conclusion can fail for two disjoint closed sets when neither is compact. In R^2 let $K = \{(x,0) : x \in R\}$ (the x-axis) and $F = \{(x, 1/x) : x > 0\}$. Both are closed and disjoint, and neither is compact (both are unbounded), yet $d((x,0),(x,1/x)) = 1/x \to 0$, so no $\delta > 0$ separates them.

■

REFLECTIONS.

The distance from a point to a closed set is positive unless the point lies in the set, and ρ_F records that distance continuously. On the compact set K this positive function attains its minimum — it cannot merely approach zero without reaching it — so the minimum is a genuine positive number, and that number is a uniform gap between K and F. Compactness is doing the essential work: it upgrades "positive at every point" to "bounded below by a positive constant." Without it the gap can shrink to zero in the limit even though the sets never meet, as the x-axis and the graph of $1/x$ show — they stay disjoint while drawing arbitrarily close out toward infinity.

Problem 4.22 *Rudin, Chapter 4, Exercise 22*

Let A and B be disjoint nonempty closed sets in a metric space X, and define

$$f(p) = \frac{\rho_A(p)}{\rho_A(p) + \rho_B(p)} \qquad (p \in X).$$

Show that f is a continuous function on X whose range lies in $[0,1]$, that $f(p) = 0$ precisely on A and $f(p) = 1$ precisely on B. This establishes a converse of Exercise 3: Every closed set $A \subset X$ is $Z(f)$ for some continuous real f on X. Setting

$$V = f^{-1}\left(\left[0, \tfrac{1}{2}\right)\right), \qquad W = f^{-1}\left(\left(\tfrac{1}{2}, 1\right]\right),$$

show that V and W are open and disjoint, and that $A \subset V$, $B \subset W$. (Thus pairs of disjoint closed sets in a metric space can be covered by pairs of disjoint open sets. This property of metric spaces is called *normality*.)

SOLUTION.

First, $\rho_A(p)+\rho_B(p)>0$ for every p: if it vanished then $\rho_A(p)=\rho_B(p)=0$, so $p\in\overline{A}=A$ and $p\in\overline{B}=B$ (Problem 20), contradicting $A\cap B=\varnothing$. Hence f is defined everywhere, and as a quotient of the continuous functions ρ_A and $\rho_A+\rho_B$ with nonvanishing denominator it is continuous (R 4.9). Since $0\le\rho_A\le\rho_A+\rho_B$, the range of f lies in $[0,1]$. Moreover $f(p)=0 \iff \rho_A(p)=0 \iff p\in A$, and $f(p)=1 \iff \rho_B(p)=0 \iff p\in B$. In particular $A=\{p: f(p)=0\}=Z(f)$, so every closed set is a zero set of a continuous function — the converse of Problem 3 (indeed ρ_A alone already has zero set A).

Finally, $V=f^{-1}([0,\frac12))=f^{-1}((-\infty,\frac12))$ and $W=f^{-1}((\frac12,1])=f^{-1}((\frac12,\infty))$ are preimages of open sets under the continuous f, hence open (R 4.8). They are disjoint, and $A\subset V$ (since $f=0<\frac12$ on A) while $B\subset W$ (since $f=1>\frac12$ on B). Thus disjoint closed sets are separated by disjoint open sets, i.e. X is normal.

■

Reflections.

This builds, by hand, a continuous function that is 0 on A, 1 on B, and slides between them in between — a metric-space version of Urysohn's lemma. The construction is the natural ratio: weigh the distance to A against the total distance to both sets, so a point of A scores 0, a point of B scores 1, and the denominator never vanishes precisely because no point can touch both disjoint closed sets at once. Once such a function exists, normality is immediate — split the values at $\frac12$ and pull back the two open halves — and so is the converse to "zero sets are closed," since A is exactly where f vanishes. The distance function ρ_A from the previous problems is the single ingredient that makes all of this run.

Problem 4.23 *Rudin, Chapter 4, Exercise 23*

A real-valued function f defined in (a,b) is said to be *convex* if

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$$

whenever $a<x<b$, $a<y<b$, $0<\lambda<1$. Prove that every convex function is continuous. Prove that every increasing convex function of a convex function is convex. (For example, if f is convex, so is e^f.)

If f is convex in (a,b) and if $a<s<t<u<b$, show that

$$\frac{f(t)-f(s)}{t-s}\le\frac{f(u)-f(s)}{u-s}\le\frac{f(u)-f(t)}{u-t}.$$

Solution.

We prove the slope inequality first, since continuity follows from it. Fix $a<s<t<u<b$

and write $t = \lambda s + (1-\lambda)u$ with $\lambda = \frac{u-t}{u-s} \in (0,1)$, so $1-\lambda = \frac{t-s}{u-s}$. Convexity gives

$$f(t) \le \lambda f(s) + (1-\lambda) f(u) = \frac{(u-t)f(s) + (t-s)f(u)}{u-s}.$$

Subtracting $f(s)$ and dividing by $t-s>0$ gives $\frac{f(t)-f(s)}{t-s} \le \frac{f(u)-f(s)}{u-s}$; subtracting the displayed bound from $f(u)$ and dividing by $u-t>0$ gives $\frac{f(u)-f(s)}{u-s} \le \frac{f(u)-f(t)}{u-t}$. Together these are the asserted chain.

Continuity. Fix $x_0 \in (a,b)$ and choose s,u with $a<s<x_0<u<b$. Applying the chain to $s<x_0<x$ and to $x_0<x<u$, for any $x \in (x_0,u)$,

$$\frac{f(x_0)-f(s)}{x_0-s} \le \frac{f(x)-f(x_0)}{x-x_0} \le \frac{f(u)-f(x_0)}{u-x_0}.$$

The difference quotient is thus bounded by constants independent of x, so $|f(x)-f(x_0)| \le C\,(x-x_0) \to 0$; the same bound holds for $x \in (s,x_0)$. Hence f is continuous at x_0 (R 4.5; 7.2.1).

Increasing convex of convex. Let φ be increasing and convex on an interval containing the range of f. For x,y and $\lambda \in (0,1)$, convexity of f gives $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$; applying φ (increasing), then its convexity,

$$\varphi\big(f(\lambda x + (1-\lambda)y)\big) \le \varphi\big(\lambda f(x) + (1-\lambda) f(y)\big) \le \lambda\, \varphi(f(x)) + (1-\lambda)\, \varphi(f(y)),$$

so $\varphi \circ f$ is convex. Taking $\varphi = \exp$ shows e^f is convex whenever f is.

■

REFLECTIONS.

Convexity says the graph lies below each of its chords, and everything here is a way of reading that off the slopes of those chords. The three-term inequality says exactly that chord slopes increase as the chord slides right: the slope from s to t is at most that from s to u, which is at most that from t to u. That monotone-slope picture is the key relation. Continuity falls out because near any interior point the chord slopes are trapped between two fixed values, making f Lipschitz there — and this needs an *open* interval, since a convex function on a closed interval can jump at an endpoint. The composition fact is almost a formality once seen the right way: f convex puts its value below the chord height, an increasing φ preserves that order, and φ's own convexity bounds φ of the chord height by the chord of $\varphi \circ f$.

Problem 4.24 *Rudin, Chapter 4, Exercise 24*

Assume that f is a continuous real function defined in (a, b) such that

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}$$

for all $x, y \in (a, b)$. Prove that f is convex.

Solution.

We prove $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ first for dyadic λ, then for all $\lambda \in [0,1]$ by continuity.

Dyadic λ. Induct on n, proving the inequality for every $\lambda = j/2^n$ with $0 \leq j \leq 2^n$. For $n = 1$ the only nontrivial case is $\lambda = \frac{1}{2}$, which is the hypothesis. Assume the claim for n, and let $\lambda = j/2^{n+1}$ with j odd (if j is even, λ is already a multiple of $1/2^n$). Set $\mu = \frac{j-1}{2^{n+1}}$, $\nu = \frac{j+1}{2^{n+1}}$; as j is odd, μ and ν are multiples of $1/2^n$, and $\lambda = \frac{1}{2}(\mu+\nu)$. With $p = \mu x + (1-\mu)y$ and $q = \nu x + (1-\nu)y$ we have $\lambda x + (1-\lambda)y = \frac{1}{2}(p+q)$, so midpoint convexity gives

$$f(\lambda x + (1-\lambda)y) = f\left(\frac{p+q}{2}\right) \leq \frac{f(p)+f(q)}{2}.$$

The inductive hypothesis bounds $f(p) \leq \mu f(x)+(1-\mu)f(y)$ and $f(q) \leq \nu f(x)+(1-\nu)f(y)$; averaging these and using $\frac{1}{2}(\mu+\nu) = \lambda$,

$$\frac{f(p)+f(q)}{2} \leq \frac{\mu+\nu}{2} f(x) + \Big(1 - \frac{\mu+\nu}{2}\Big) f(y) = \lambda f(x) + (1-\lambda) f(y).$$

All λ. The dyadic rationals are dense in $[0,1]$. Given $\lambda \in [0,1]$, take dyadic $\lambda_k \to \lambda$; then $\lambda_k x + (1-\lambda_k)y \to \lambda x + (1-\lambda)y$, and continuity of f (R 4.6) lets us pass to the limit in $f(\lambda_k x + (1-\lambda_k)y) \leq \lambda_k f(x) + (1-\lambda_k)f(y)$, giving $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$. Hence f is convex.

■

Reflections.

Midpoint convexity is the convexity inequality at the single ratio $\frac{1}{2}$, and the question is how much that one case already forces. Taking midpoints repeatedly reaches all the dyadic ratios — halving generates $\frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \ldots$ — so the inequality spreads from $\frac{1}{2}$ to a dense set of λ by an induction that splits an odd dyadic into the average of its two even neighbors. Density finishes the job only because f is continuous: a continuous function obeying an inequality on a dense set obeys it everywhere. Continuity is essential rather than cosmetic — without it, midpoint-convex functions can be wildly discontinuous (built from a non-measurable additive function) and fail to be convex, so the hypothesis is exactly what excludes those pathologies.

Problem 4.25 *Rudin, Chapter 4, Exercise 25*

If $A \subset R^k$ and $B \subset R^k$, define $A + B$ to be the set of all sums $\mathbf{x} + \mathbf{y}$ with $\mathbf{x} \in A$, $\mathbf{y} \in B$.

1. If K is compact and C is closed in R^k, prove that $K + C$ is closed. *Hint:* Take $\mathbf{z} \notin K + C$, put $F = \mathbf{z} - C$, the set of all $\mathbf{z} - \mathbf{y}$ with $\mathbf{y} \in C$. Then K and F are disjoint. Choose δ as in Exercise 21. Show that the open ball with center $\mathbf{z}$ and radius δ does not intersect $K + C$.

2. Let α be an irrational real number. Let C_1 be the set of all integers, let C_2 be the set of all $n\alpha$ with $n \in C_1$. Show that C_1 and C_2 are closed subsets of R^1 whose sum $C_1 + C_2$ is *not* closed, by showing that $C_1 + C_2$ is a countable dense subset of R^1.

SOLUTION.

1. Let $\mathbf{z} \notin K+C$; we find a ball about $\mathbf{z}$ missing $K+C$. Put $F = \mathbf{z}-C = \{\mathbf{z}-\mathbf{y} : \mathbf{y} \in C\}$, closed since C is closed and $\mathbf{y} \mapsto \mathbf{z} - \mathbf{y}$ is a homeomorphism. Then $K \cap F = \varnothing$: if $\mathbf{w} \in K \cap F$, then $\mathbf{w} = \mathbf{z} - \mathbf{y}$ for some $\mathbf{y} \in C$, giving $\mathbf{z} = \mathbf{w} + \mathbf{y} \in K + C$, a contradiction. By Problem 21 there is $\delta > 0$ with $|\mathbf{p} - \mathbf{q}| > \delta$ for all $\mathbf{p} \in K$, $\mathbf{q} \in F$. If some $\mathbf{w}$ with $|\mathbf{w} - \mathbf{z}| < \delta$ lay in $K + C$, write $\mathbf{w} = \mathbf{k} + \mathbf{c}$ with $\mathbf{k} \in K$, $\mathbf{c} \in C$; then $\mathbf{z} - \mathbf{c} \in F$ and $|\mathbf{k} - (\mathbf{z} - \mathbf{c})| = |\mathbf{w} - \mathbf{z}| < \delta$, contradicting the choice of δ. So the ball of radius δ about $\mathbf{z}$ misses $K + C$, and $K + C$ is closed.

2. $C_1 = \mathbb{Z}$ and $C_2 = \{n\alpha : n \in \mathbb{Z}\}$ are closed: each consists of isolated points (consecutive elements differ by 1 and by $|\alpha|$, respectively), so neither has a limit point. Their sum $C_1 + C_2 = \{m + n\alpha : m, n \in \mathbb{Z}\}$ is countable. It is dense: given $\varepsilon > 0$, choose an integer $N > 1/\varepsilon$ and split $[0, 1)$ into N subintervals of length $1/N$; among the $N + 1$ fractional parts of $0, \alpha, 2\alpha, \ldots, N\alpha$ two lie in one subinterval, so for some $0 \le i < j \le N$ there is an integer m with $0 < |(j - i)\alpha - m| < 1/N < \varepsilon$ (nonzero since α is irrational). Then $\beta = (j - i)\alpha - m \in C_1 + C_2$ satisfies $0 < |\beta| < \varepsilon$, and its integer multiples, all in $C_1 + C_2$, are spaced $|\beta|$ apart, so $C_1 + C_2$ meets every interval of length ε. Being dense, $C_1 + C_2$ has closure all of R^1; but $C_1 + C_2 \neq R^1$ since it is countable, so $C_1 + C_2$ does not contain all its limit points and is not closed.

■

REFLECTIONS.

Part (a) is the Minkowski sum of a compact set with a closed set, and the reason it stays closed is the uniform gap from the previous problem. To show a point $\mathbf{z}$ outside the sum has room around it, reflect the closed set through $\mathbf{z}$ to turn "$\mathbf{z} \notin K + C$" into "K and the reflected set are disjoint"; one is compact and one is closed, so a positive distance δ holds them apart, and that same δ is the radius of a ball about $\mathbf{z}$ the sum cannot reach. Part (b) shows the compactness is not optional — two closed sets can sum to something badly non-closed. The integers together with the integer multiples of an irrational α give $\mathbb{Z} + \mathbb{Z}\alpha$, the countable dense set behind irrational rotations: dense because an irrational

slope forces the fractional parts of $n\alpha$ to crowd into every subinterval, and a countable dense set can never be closed.

Problem 4.26 *Rudin, Chapter 4, Exercise 26*

Suppose X, Y, Z are metric spaces, and Y is compact. Let f map X into Y, let g be a continuous one-to-one mapping of Y into Z, and put $h(x) = g(f(x))$ for $x \in X$.
Prove that f is uniformly continuous if h is uniformly continuous. *Hint:* g^{-1} has compact domain $g(Y)$, and $f(x) = g^{-1}(h(x))$.
Prove also that f is continuous if h is continuous.
Show (by modifying Example 4.21, or by finding a different example) that the compactness of Y cannot be omitted from the hypotheses, even when X and Z are compact.

Solution.

Since Y is compact and $g : Y \to Z$ is continuous and one-to-one, g is a homeomorphism of Y onto $g(Y)$; in particular $g^{-1} : g(Y) \to Y$ is continuous (R 4.17). Also $g(Y)$ is compact (R 4.14), so g^{-1}, being continuous on a compact space, is uniformly continuous (R 4.19). Because $g \circ f = h$, we have $f = g^{-1} \circ h$.

If h is uniformly continuous, then $f = g^{-1} \circ h$ is a composition of uniformly continuous maps (Problem 12), hence uniformly continuous. If h is merely continuous, then $f = g^{-1} \circ h$ is a composition of continuous maps, hence continuous (R 4.7).

Compactness of Y cannot be dropped, even with X and Z compact. Let $X = Z$ be the unit circle (compact) and $Y = [0, 2\pi)$ (not compact), with $g : [0, 2\pi) \to Z$, $g(t) = (\cos t, \sin t)$, the continuous one-to-one map of Example 4.21 (R 4.21). Take $h = \mathrm{id}$ on the circle, which is uniformly continuous. The relation $g \circ f = h$ forces $f = g^{-1}$, the inverse parametrization, discontinuous at $(1, 0)$ (its value jumps from near 2π to 0). So f fails to be continuous although h is uniformly continuous, showing the compactness of Y is essential.

■

Reflections.

The point is that an injective continuous map out of a compact space is automatically a homeomorphism onto its image, so its inverse is not just continuous but uniformly continuous. That lets you write $f = g^{-1} \circ h$ and inherit whatever regularity h has — uniform continuity from uniform continuity, continuity from continuity — because composing with the well-behaved g^{-1} cannot spoil it. Everything hinges on g^{-1} being continuous, and that is precisely where compactness of Y is spent: drop it and g can be a continuous bijection whose inverse jumps, like the wrapping of the half-open interval $[0, 2\pi)$ onto the circle. There h can be as smooth as the identity while $f = g^{-1}$ tears apart at the point where the interval's two ends are glued together.