

4.7 Soundness

Chapter 3 ended on a question it could not answer. Section 3.1 searched for a derivation of $P \rightarrow Q \vdash Q \rightarrow P$, and every route came back to the same gap; but a search that fails proves nothing, since a cleverer derivation might exist. The row that refutes affirming the consequent in Section 4.5, $p = F$ and $q = T$, finds what the search could not: it makes $p \rightarrow q$ true and $q \rightarrow p$ false, and the search of Section 4.6 finds such rows whenever they exist. This section proves the theorem that turns such a row into a proof that **no** derivation exists. It is called soundness: whatever the rules derive is true in every row where the premises are true. The proof is the book's first long proof about proofs, and it is an induction on the lines of a derivation.

4.7.1 What must be proved, and why it is not obvious

Remark. Two levels. This section speaks about the calculus from outside it.

Object level: written in the language	Meta level: claims about it, proved in ordinary mathematics
formulas; derivations, with their lines, boxes, justifications and printed purposes (“asm, for $\rightarrow I$ ”, which the proof reads as the rule whose case applies)	“ $\Gamma \vdash \psi$ ”, “ $\Gamma \models \psi$ ”, “line k is justified by $\vee E$ ”, “ $\Gamma_k \models \varphi_k$ ”

Section 4.2's checklist reads: objects, the lines of one derivation; property, $\Gamma_k \models \varphi_k$ (contexts are defined below); base cases, the premises and assumptions, which cite nothing; steps, one for each rule; principle, Theorem 4.2.2.

🕒 **Second reading** · *Is the proof circular?*

The proof below argues by cases (“if $v(\varphi) = F \dots$ if $v(\varphi) = T$ ”) and by contradiction, in the metalanguage: the very patterns that $\vee E$ and RAA record in the object language. That is not circular. The theorem is not meant to persuade someone who rejects reasoning by cases. It is a theorem of ordinary mathematics about a calculus, and what it rules out is a mismatch between the rules and the tables, of the kind the connective in the fold of Section 4.4 has.

Chapters 1 and 2 checked each rule as it arrived. $\wedge I$ is safe because two true statements make a true conjunction (Section 1.2); $\rightarrow E$ is safe because that is what “if $\dots$ then” asserts (Section 1.3). Section 1.2 said once that Chapter 4 would prove it for every rule of Chapters 1 and 2 at once. Why is that more than collecting the separate checks?

The first attempt at a proof is: go down a derivation, line by line, and show that each line is true in every row in which the premises are true. For a derivation without boxes this works. Each line is a premise, or comes from earlier lines by a rule, and the rule turns true lines into a true line. But it fails at the first box. Here is the derivation of Proposition 2.3.2, written with

atoms:

1	$p \rightarrow q$	
2	$\neg q$	asm, for $\rightarrow$ I
3	p	asm, for $\neg$ I
4	q	$\rightarrow$ E 1, 3
5	$\perp$	$\neg$ E 4, 2
6	$\neg p$	$\neg$ I 3–5
7	$\neg q \rightarrow \neg p$	$\rightarrow$ I 2–6

Line 3 is p , and the row $p = \text{F}, q = \text{F}$ makes the premise $p \rightarrow q$ true and line 3 false. Line 5 is $\perp$, which is false in **every** row. So the lines inside a box are not true in the rows where the premises are true, and the first attempt collapses. Yet the derivation is correct, and its last line $\neg q \rightarrow \neg p$ is true in every row making $p \rightarrow q$ true (Example (An implication and its contrapositive)).

What goes wrong is that a line inside a box is not claimed on the strength of the premises alone. Line 4 says: *if* $p \rightarrow q$, $\neg q$ and p all hold, then q holds. That is a true claim about rows. Every line depends on the premises **and** on the assumptions of the boxes around it, and a correct statement about each line must mention both. Below we give this list a name: the context of the line.

4.7.2 Derivations of formulas

From now on the lines of a derivation carry **formulas** in the sense of Definition 4.1.1, with atoms $p, q, r, \dots$. Nothing in the rules changes. Every rule was stated as a schema, whose letters φ, ψ, χ stand for arbitrary statements (Definition 1.2.1), so it applies to formulas as it stands, and Definition 1.4.2 and Definition 1.2.3 read the same with “formula” for “statement”. As Section 4.1 said, a derivation written with Chapter 2’s letters P, Q is a pattern, and reading its letters as atoms, or as any formulas, gives a derivation of formulas.

Two conventions for the rest of the section. The **primitive rules** are those of Chapters 1 and 2 collected in Section 3.1: $\text{R}, \wedge\text{I}, \wedge\text{E}, \rightarrow\text{I}, \rightarrow\text{E}, \vee\text{I}, \vee\text{E}, \neg\text{I}, \neg\text{E}, \perp\text{E}, \text{RAA}, \leftrightarrow\text{I}$ and $\leftrightarrow\text{E}$. The **box rules** among them are the five that cite a box: $\rightarrow\text{I}, \neg\text{I}, \text{RAA}, \vee\text{E}$ and $\leftrightarrow\text{I}$. The other eight, R included, cite only lines; we call them the **one-step rules**.

A derivation may also use derived rules such as MT and LEM. A line citing one abbreviates a stretch of derivation in rules proved earlier, with the letters replaced by formulas (Definition 2.3.1), and Section 2.3 showed how to write it out: repeat the cited lines, then paste the derivation of the derived rule below them. Every chain of derived rules ends in primitive rules. So whenever $\Gamma \vdash \psi$, there is also a derivation of ψ from premises in Γ that uses **primitive rules only**. That is the only kind of derivation the proof below has to consider.

4.7.3 The context of a line

Let a derivation have lines $1, \dots, n$. We write φ_k for the formula on line k . The **length** of a derivation is its number of lines. The **context** of line k is the list of the premises together with the assumptions of every box containing line k : the statements the line depends on. We write it Γ_k . The context of the last line is just the premises, since by Definition 1.4.2 the last line lies in no box. For lists of formulas Δ and Δ' , we write $\Delta \subseteq \Delta'$ when every formula of Δ occurs in

Δ' .

Exercise C4 of Section 1.4 (Exercise (C4: The assumptions a line depends on, used later)) wrote such lists by hand, as sets of line numbers; its part (c) observed what [Lemma 4.7.1](#) below proves.

The context of a line is everything that line depends on: the premises, and every assumption still open where the line stands.

In the derivation of Proposition 2.3.2 above, the boxes are 2–6 and 3–5, and the contexts are:

Line	Formula	Boxes containing it	Context
1	$p \rightarrow q$	none	$p \rightarrow q$
2	$\neg q$	2–6	$p \rightarrow q, \neg q$
3, 4, 5	$p, q, \perp$	2–6, 3–5	$p \rightarrow q, \neg q, p$
6	$\neg p$	2–6	$p \rightarrow q, \neg q$
7	$\neg q \rightarrow \neg p$	none	$p \rightarrow q$

Quick check. In that derivation, which lines have p in their context?

Solution. Lines 3, 4 and 5, the lines of the box opened by p .

Now look at each line through its context. Line 3 is one of the formulas in its own context, so every row making the context true makes line 3 true. Line 4: every row making $p \rightarrow q$ and p true makes q true. Line 5: no row makes $p \rightarrow q, \neg q$, and p all true (the first and third force q), so the context is unsatisfiable and it entails even $\perp$ (Section 4.5). Line 6: every row making $p \rightarrow q$ and $\neg q$ true makes p false. Line 7: its context is the premise, and we know $p \rightarrow q \models \neg q \rightarrow \neg p$. So on this derivation, every line k satisfies

$$\Gamma_k \models \varphi_k.$$

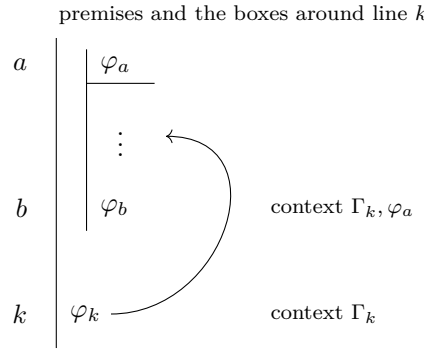
This is the statement that survives the boxes, and it is the one we shall prove for every line of every derivation.

Two facts about contexts, read off the scope rule, do the bookkeeping. The first says that citing an available line never involves an assumption the citing line does not have. The second says exactly what the context of a box's last line is, when the box is cited by a box rule.

Lemma 4.7.1. Contexts in a Derivation

Let lines i, k and $a \leq b$ belong to a derivation, and let Γ_j denote the context of line j .

- (a) If line i is available at line k , then $\Gamma_i \subseteq \Gamma_k$.
- (b) If line k is justified by a box rule citing the box whose first line is a and whose last line is b , then the formulas of Γ_b are exactly those of Γ_k together with φ_a , the assumption of that box. We write this $\Gamma_b = \Gamma_k, \varphi_a$.



Part (b): the last line of a cited box depends on everything line k depends on, and on the box's own assumption. The arrow is the citation.

Idea. Both parts compare two sets of boxes, since contexts differ only in the assumptions of the boxes around a line. For (a), the scope rule says that the boxes around line i are among the boxes around line k . For (b), the boxes around line b should be the cited box itself and the boxes around line k ; the danger is a further box inside the cited box that also ends on line b , and the box rules forbid exactly that.

The proof reads both parts off the scope rule; it is in the fold below.

🕒 **Second reading** · The proof: reading the scope rule

Proof. (a) The premises belong to every context. Any other formula of Γ_i is the assumption of a box containing line i . Since line i is available at line k , every box containing line i also contains line k , by the scope rule of Definition 1.4.2. So that assumption belongs to Γ_k .

(b) Let B be the box from line a to line b . Since line k cites B , the box B is available at line k . By the scope rule, it ends before line k , and the boxes containing B are exactly the boxes containing line k .

We find the boxes containing line b . Every box containing line k contains B , and so contains line b ; and B contains line b . Conversely, let C be a box containing line b . By Definition 1.4.2, two boxes are either disjoint or one inside the other; C and B share line b , so $C = B$, or C contains B , or C lies inside B . The last is impossible: then line b , the last line of B , would be a line of a box inside B , and every box rule requires the last line of its box to be a line of that box itself, not of a box inside it (Definition 1.4.1, Definition 2.2.1, Definition 2.5.1, Definition 1.6.1, Definition 2.4.1). So $C = B$, or C is one of the boxes containing B , which are the boxes containing line k .

Hence the boxes containing line b are B and the boxes containing line k . The assumption of B is φ_a , so Γ_b consists of the premises, the assumptions of the boxes containing line k , and φ_a , which is Γ_k together with φ_a .

■

In the derivation of Proposition 2.3.2, line 6 cites the box 3–5 by $\neg$ I, and $\Gamma_5 = p \rightarrow q, \neg q, p$ is $\Gamma_6 = p \rightarrow q, \neg q$ with the assumption p added, as (b) says. Line 5 cites lines 4 and 2, both available there, and Γ_4 and Γ_2 are contained in Γ_5 , as (a) says.

🕒 **Second reading** · *A box that ends inside another*

Warning. Part (b) needs the box to be **cited by a box rule**. Definition 1.4.2 allows a box to end on the same line as a box inside it, and then its last line has a larger context. In the derivation below, the box 2–4 is available at line 5, but $\Gamma_4 = p, q, r$ contains r , the assumption of the inner box, as well as q . No box rule may cite 2–4: its last line lies inside the box 3–4, not in 2–4 itself. (This is why the derivation of Example (Regrouping a disjunction) needs its line 11.)

1		p	
2			q
3			
4			
5		p	

R 1

R 1

Quick check. In the derivation of Proposition 2.3.2 above, which case of the soundness claim handles line 5? Which lines' contexts are involved, and why does $\Gamma_5 \models \perp$ hold?

Solution. Line 5 is $\perp$, by $\neg$ E citing lines 4 and 2: a one-step rule. Lines 4 and 2 are available at line 5, so by Lemma 4.7.1 (a), $\Gamma_4 = p \rightarrow q, \neg q, p$ and $\Gamma_2 = p \rightarrow q, \neg q$ are contained in $\Gamma_5 = p \rightarrow q, \neg q, p$. The claim for the earlier lines gives $\Gamma_4 \models q$ and $\Gamma_2 \models \neg q$. So a valuation making Γ_5 true would make both q and $\neg q$ true, which clause (V2) of Definition 4.4.2 forbids. No valuation makes Γ_5 true, and $\Gamma_5 \models \perp$ holds because there is no row to check.

4.7.4 The soundness theorem

We can now state the theorem that Chapters 1 and 2 kept pointing to.

Theorem 4.7.2. Soundness

Let Γ be a list of formulas and ψ a formula. If $\Gamma \vdash \psi$, then $\Gamma \models \psi$.

Try it first (ten minutes): suppose line k is $\varphi \rightarrow \psi$, by $\rightarrow$ I from a box that opens with φ and ends with ψ . Suppose you already know that every row making the context of line k and φ true makes ψ true. Why is $\varphi \rightarrow \psi$ true in every row making the context of line k true? What happens in the rows where φ is false?

Idea. An attempt at the $\rightarrow$ I case runs into two things. First, knowledge about the box's last line is knowledge about rows making its **own** context true, which contains the assumption φ ; that is why the claim must carry contexts, and why [Lemma 4.7.1](#) (b) is needed. Second, the rows in which φ is false are not covered by the box at all, and they need nothing: there $\varphi \rightarrow \psi$ is true by (V5). The same two moves handle every box rule. The Step roadmap:

- ① Prove the stronger claim $\Gamma_k \models \varphi_k$ for **every** line k of a derivation in primitive rules.
- ② Use strong induction on k (Theorem 4.2.2): a rule may cite **any** earlier line, not only the line just above, so the hypothesis must be available for all earlier lines at once.
- ③ Split on how line k is justified. A premise or an assumption belongs to its own context. A one-step rule: apply the hypothesis to the cited lines, move their contexts into Γ_k by [Lemma 4.7.1](#) (a), and use one clause of Definition 4.4.2. A box rule: apply the hypothesis to each box's last line, whose context is Γ_k plus the box's assumption by [Lemma 4.7.1](#) (b), and argue about the rows where that assumption is true or false.
- ④ The last line lies in no box, so its context is the premises, all in Γ ; adding the other formulas of Γ keeps the entailment.

Before reading on: in the case of RAA (Claim 3 below), where does the proof use that $v(\neg\varphi)$ is F exactly when $v(\varphi)$ is T, and which half of that fact does it use?

Proof. Step 1 (reduction to primitive rules). Suppose $\Gamma \vdash \psi$. As shown in the subsection on derivations of formulas, there is a derivation of ψ whose premises all belong to Γ and which uses primitive rules only. Fix one, with lines $1, \dots, n$; let φ_k be the formula on line k and Γ_k its context.

Step 2 (the induction). For $k \in \mathbb{N}$, let $S(k)$ be the statement: *if $1 \leq k \leq n$, then $\Gamma_k \models \varphi_k$* . We prove $S(k)$ for every k by Theorem 4.2.2. So let $k \in \mathbb{N}$, and assume $S(j)$ for every $j < k$. If $k = 0$ or $k > n$, there is nothing to prove, so let $1 \leq k \leq n$. Every line and every box that line k cites is available at line k , and so lies before line k (Definition 1.4.2). Hence the induction hypothesis $\Gamma_j \models \varphi_j$ holds for every line j that line k cites, and for the last line of every box it cites.

Let v be a valuation that makes every formula of Γ_k true. We show $v(\varphi_k) = \text{T}$; since v is arbitrary, this proves $\Gamma_k \models \varphi_k$. By Definition 1.4.2, line k is a premise, an assumption, or the conclusion of one primitive rule, and the three claims below treat these cases.

Step 3 (the cases).

Claim. If line k is a premise or an assumption, then $v(\varphi_k) = \text{T}$.

Proof of Claim. A premise belongs to every context. An assumption opens a box, that box contains line k , and so its assumption φ_k belongs to Γ_k . In both cases φ_k is one of the formulas that v makes true. ■

For the one-step rules we first record what the hypothesis gives for a cited line. If line k cites line i , then line i is available at line k , so $\Gamma_i \subseteq \Gamma_k$ by [Lemma 4.7.1](#) (a). Hence v makes every

formula of Γ_i true, and the induction hypothesis $\Gamma_i \models \varphi_i$ gives

$$v(\varphi_i) = \text{T} \quad \text{for every line } i \text{ cited by line } k. \quad (*)$$

Claim. If line k is justified by a one-step rule, then $v(\varphi_k) = \text{T}$.

Proof of Claim. We take the rules one at a time; in each, the cited lines are true under v by (*).

R. Line k repeats a line i , so $\varphi_k = \varphi_i$, and $v(\varphi_k) = \text{T}$.

$\wedge$ I. Line k is $\varphi \wedge \psi$, citing lines holding φ and ψ . Both are true under v , so $\varphi \wedge \psi$ is true by (V3).

$\wedge$ E. Line k is φ , citing a line holding $\varphi \wedge \psi$. That line is true under v , so by (V3) both conjuncts are true, and in particular φ . The second form, which writes ψ , is the same with the other conjunct, and we skip it.

$\rightarrow$ E. Line k is ψ , citing lines holding $\varphi \rightarrow \psi$ and φ . Both are true under v . By (V5), an implication with a true hypothesis and a false conclusion is false; since $\varphi \rightarrow \psi$ is true and φ is true, ψ is true.

$\vee$ I. Line k is $\varphi \vee \psi$ or $\psi \vee \varphi$, citing a line holding φ . That line is true under v , so both disjunctions are true by (V4).

$\neg$ E. Line k is $\perp$, citing lines holding φ and $\neg\varphi$. By (*) both would be true under v , and by (V2) $\neg\varphi$ is false whenever φ is true. So no valuation makes every formula of Γ_k true, and there is no v to consider: $\Gamma_k \models \perp$ holds because there is no row to check.

$\perp$ E. Line k is any formula φ , citing a line holding $\perp$. By (*), $v(\perp) = \text{T}$, which (V1) forbids. So again no valuation makes every formula of Γ_k true, and $\Gamma_k \models \varphi$ holds.

$\leftrightarrow$ E. Line k is ψ , citing lines holding $\varphi \leftrightarrow \psi$ and φ . Both are true under v ; by (V6), $v(\psi) = v(\varphi) = \text{T}$. The second form, which writes φ from $\varphi \leftrightarrow \psi$ and ψ , is the same with the sides exchanged, and we skip it.

■

For the box rules we record what the hypothesis gives for a cited box. If line k cites, by a box rule, the box from line a to line b , then $\Gamma_b = \Gamma_k, \varphi_a$ by [Lemma 4.7.1](#) (b), and the induction hypothesis for line b says

$$\Gamma_k, \varphi_a \models \varphi_b : \quad \text{every valuation making } \Gamma_k \text{ and } \varphi_a \text{ true makes } \varphi_b \text{ true.} \quad (**)$$

Claim. If line k is justified by a box rule, then $v(\varphi_k) = \text{T}$.

Proof of Claim. Recall that v makes every formula of Γ_k true.

$\rightarrow$ I. Line k is $\varphi \rightarrow \psi$, citing a box that opens with φ and ends with ψ . By (**), every valuation making Γ_k and φ true makes ψ true. If $v(\varphi) = \text{F}$, then $v(\varphi \rightarrow \psi) = \text{T}$ by (V5). If $v(\varphi) = \text{T}$, then v makes Γ_k and φ true, so $v(\psi) = \text{T}$, and again $v(\varphi \rightarrow \psi) = \text{T}$ by (V5).

$\neg$ I. Line k is $\neg\varphi$, citing a box that opens with φ and ends with $\perp$. By (**), every valuation making Γ_k and φ true makes $\perp$ true; since (V1) makes $\perp$ false under every valuation, no

valuation makes Γ_k and φ true. As v makes Γ_k true, it must make φ false, and then $v(\neg\varphi) = \text{T}$ by (V2).

RAA. Line k is φ , citing a box that opens with $\neg\varphi$ and ends with $\perp$. By $(**)$ and (V1), as in the previous case, no valuation makes Γ_k and $\neg\varphi$ true. Suppose $v(\varphi) = \text{F}$. Then $v(\neg\varphi) = \text{T}$ by (V2), so v makes Γ_k and $\neg\varphi$ true, which is impossible. Hence $v(\varphi) = \text{T}$.

$\vee\text{E}$. Line k is χ , citing a line i holding $\varphi \vee \psi$, a box that opens with φ and ends with χ , and a box that opens with ψ and ends with χ . By $(*)$, $v(\varphi \vee \psi) = \text{T}$, so by (V4), $v(\varphi) = \text{T}$ or $v(\psi) = \text{T}$. By $(**)$ for the first box, every valuation making Γ_k and φ true makes χ true; for the second, every valuation making Γ_k and ψ true makes χ true. If $v(\varphi) = \text{T}$, then v makes Γ_k and φ true, so $v(\chi) = \text{T}$. If $v(\psi) = \text{T}$, the second box gives $v(\chi) = \text{T}$ in the same way.

$\leftrightarrow\text{I}$. Line k is $\varphi \leftrightarrow \psi$, citing a box that opens with φ and ends with ψ , and one that opens with ψ and ends with φ . By $(**)$, every valuation making Γ_k and φ true makes ψ true, and every valuation making Γ_k and ψ true makes φ true. If $v(\varphi) = \text{T}$, the first box gives $v(\psi) = \text{T}$. If $v(\varphi) = \text{F}$, then $v(\psi) = \text{F}$ too, since $v(\psi) = \text{T}$ would give $v(\varphi) = \text{T}$ by the second box. In both cases $v(\varphi) = v(\psi)$, so $v(\varphi \leftrightarrow \psi) = \text{T}$ by (V6).

■

The three claims cover every way line k can be justified, so $S(k)$ holds. By Theorem 4.2.2, $S(k)$ holds for every $k \in \mathbb{N}$: every line k of the derivation satisfies $\Gamma_k \models \varphi_k$.

Step 4 (the last line). The last line n of a derivation lies in no box (Definition 1.4.2), so its context Γ_n consists of the premises alone, and $\varphi_n = \psi$. By the induction of Steps 2 and 3, $\Gamma_n \models \psi$. Every premise belongs to Γ (Definition 1.2.3), so a valuation making every formula of Γ true makes every formula of Γ_n true, and hence makes ψ true. This proves $\Gamma \models \psi$.

■

The proof is long, but each case is one step of the truth-table reasoning of Section 4.4, placed in the right context. Every primitive rule has passed. So has every derived rule, since each abbreviates primitive ones. *The rules never lead from truth to falsehood*, not merely line by line but for every derivation that can ever be written.

🕒 **Second reading** · Why induction on the position of a line

Why strong induction and the position k of a line? A rule may cite any available line, however far back: line 7 of the derivation of Proposition 2.3.2 cites the box that begins at line 2. So the hypothesis is needed for all earlier lines at once. A different route would induct on the length of the whole derivation, removing its last line. But the lines above the last line need not form a derivation, since a box may still be open there (in the derivation of Proposition 2.3.2, lines 1–6 end inside the box 2–6), and that route has to cut boxes open. Carrying the context of every line avoids the surgery.

Example. Running the claim on a case split

Here is a derivation of $p \vee q \vdash q \vee p$. List the context of each line, and check $\Gamma_k \models \varphi_k$ for each k , naming the case of the proof that handles it.

1	$p \vee q$	
2	p	asm, for $\vee E$
3	$q \vee p$	$\vee I$ 2
4	q	asm, for $\vee E$
5	$q \vee p$	$\vee I$ 4
6	$q \vee p$	$\vee E$ 1, 2–3, 4–5

Solution. The boxes are 2–3 and 4–5.

Line	Context Γ_k	Case	Why $\Gamma_k \models \varphi_k$
1	$p \vee q$	premise	φ_1 is in its context
2	$p \vee q, p$	assumption	φ_2 is in its context
3	$p \vee q, p$	$\vee I$ citing 2	$\Gamma_2 \subseteq \Gamma_3$; a row making p true makes $q \vee p$ true
4	$p \vee q, q$	assumption	φ_4 is in its context
5	$p \vee q, q$	$\vee I$ citing 4	as for line 3, with q
6	$p \vee q$	$\vee E$ citing 1 and both boxes	see below

For line 6: by [Lemma 4.7.1](#) (b), $\Gamma_3 = \Gamma_6, p$ and $\Gamma_5 = \Gamma_6, q$, and the claim for lines 3 and 5 says that rows making $p \vee q$ and p true, and rows making $p \vee q$ and q true, make $q \vee p$ true. A row making Γ_6 true makes $p \vee q$ true, so it makes p or q true, and one of the two boxes applies. Line 6 is the last line, and its context is the premise: so $p \vee q \models q \vee p$, as soundness promises.

4.7.5 Every rule, checked against the tables

Rule	Introduced	Case	Clause	Wrong version; its derivation; refuting row
R	1.4 (clause (D3))	Claim 2	none: the line is repeated	
$\wedge I, \wedge E$	1.2	Claim 2	(V3)	
$\rightarrow E$	1.3	Claim 2	(V5)	
$\rightarrow I$	1.4	Claim 3	(V5)	“a box from φ to ψ gives $\psi \rightarrow \varphi$ ”; the box $p \wedge q, p$ would give $\vdash p \rightarrow (p \wedge q)$; row $p = T, q = F$

Rule	Introduced	Case	Clause	Wrong version; its derivation; refuting row
$\vee I$	1.5	Claim 2	(V4)	
$\vee E$	1.6	Claim 3	(V4)	“the two boxes may end in different lines; write the first box’s”; from $p \vee q$, the boxes p and q would give $p \vee q \vdash p$; row $p = F, q = T$
$\neg E, \perp E$	2.1	Claim 2	(V2); (V1)	
$\neg I$	2.2	Claim 3	(V1), (V2)	“a box from φ to any line gives $\neg\varphi$ ”; the one-line box p would give $\vdash \neg p$; row $p = T$
$\leftrightarrow I, \leftrightarrow E$	2.4	Claim 3; Claim 2	(V6)	
RAA	2.5	Claim 3	(V1), (V2)	“a box from φ to $\perp$ gives φ ”; from $\neg p$, the box $p, \perp$ (by $\neg E$) would give $\neg p \vdash p$; row $p = F$

The thirteen primitive rules, with the case and clause each uses. Each wrong version changes one condition, and its case fails there: the row makes the premises true and the line false.

🕒 **Second reading** • *The four wrong derivations, written out*

Here are the four derivations the wrong versions would license, written as patterns with Chapter 2’s letters; read P and Q as the atoms p and q for the rows of the table. Each is correct except at its last line, which only the wrong version allows. Wrong $\rightarrow I$ reverses the implication its box proves; wrong $\vee E$ closes on a line that only the first box reaches; wrong $\neg I$ closes a box that never reached $\perp$; wrong RAA opens its box with φ instead of $\neg\varphi$.

1			$P \wedge Q$	asm, for $\rightarrow$ I	1		$P \vee Q$	
2			P	$\wedge$ E 1	2			P
3		$P \rightarrow (P \wedge Q)$	$\rightarrow$ I 1-2	3				Q
						4	P	$\vee$ E 1, 2-2, 3-3

1		$\neg P$		1		$\neg P$	
2		$\neg P$	asm, for $\neg I$	2		$\neg P$	asm, for RAA
		$\neg P$	$\neg I$ 1-1	3		$\perp$	$\neg E$ 2, 1
		$\neg P$		4		P	RAA 2-3

4.7.6 Refuting with one row

Soundness is most often used backwards. If the conclusion can be false while the premises are true, no derivation can exist.

Corollary 4.7.3. Refuting Row

Let Γ be a list of formulas and ψ a formula. If some valuation makes every formula of Γ true and ψ false, then $\Gamma \not\models \psi$.

Proof. Let v make every formula of Γ true and ψ false. Then $\Gamma \models \psi$ fails (Definition 4.5.3). If $\Gamma \vdash \psi$ held, [Theorem 4.7.2](#) would give $\Gamma \models \psi$. So $\Gamma \vdash \psi$ does not hold. ■

Here $\Gamma \not\models \psi$ means that there is **no** derivation of ψ from premises in Γ : not with the primitive rules, not with any derived rule, not in a thousand lines. One row has settled a question about infinitely many derivations.

Quick check. Does the row $p = q = T$ refute $p \rightarrow q \vdash q$? Does the row $p = q = F$?

Solution. The first does not, since its conclusion q is true. The second does: $p \rightarrow q$ is T and q is F. So $p \rightarrow q \not\models q$, by [Corollary 4.7.3](#).

Chapter 3's question. Section 3.1 searched for a derivation of $P \rightarrow Q \vdash Q \rightarrow P$ and stalled. The row $p = F, q = T$ makes $p \rightarrow q$ true (false hypothesis) and $q \rightarrow p$ false (true hypothesis, false conclusion). By [Corollary 4.7.3](#),

$$p \rightarrow q \not\models q \rightarrow p.$$

The letters of Chapter 3 stood for arbitrary statements, and the search asked for a derivation of the **pattern**: one derivation, written with P and Q , that works whatever they are. Such a derivation, with P and Q read as the atoms p and q , would derive $p \rightarrow q \vdash q \rightarrow p$, which we have just ruled out. So the pattern has no derivation, and the search was right to fail. (Particular instances can be derivable: with P and Q the same statement, $P \rightarrow P \vdash P \rightarrow P$ takes one line. A refuting row for atoms speaks about the pattern, as the warning after Lemma 4.5.3 said.)

What soundness certifies. Soundness turns every break this book has made into a proof that no derivation exists. Section 1.1 called a form bad when some choice of statements makes its counterexample set true (Definition 1.1.3). For a propositional form, a row making the

counterexample set true is such a choice, and by [Corollary 4.7.3](#) no derivation of the form exists, with any rule of Chapters 1 and 2. Chapter 1 promised exactly this, in Section 1.3, for affirming the consequent: the row $p = \text{F}, q = \text{T}$ of Example (Affirming the consequent is not valid) makes $p \rightarrow q$ and q true and p false, so

$$p \rightarrow q, q \not\vdash p.$$

So the repair of Argument B that Section 1.4 considered is not merely unjustified: it is impossible. The same holds for every break in the derive-or-break items from Section 1.5 on, for each claim of Section 3.2 that the search of Section 4.6 refuted, the four breaks of Exercise (B3: Derive or break, used later) among them, and for every refuting row of Sections 4.5 and 4.6. A break by statements about a number certifies the same: the truth values of those statements form a row, and since the connectives are truth-functional (Section 4.4), that row makes the counterexample set true.

Warning. A refuting row must make **every** premise true. For $p \rightarrow q$, $p \vdash q$, the row $p = \text{F}, q = \text{F}$ makes the conclusion false, but it also makes the premise p false, so it refutes nothing; and the sequent is derivable, by one line of $\rightarrow\text{E}$. A row in which some premise fails is simply not one of the rows Definition 4.5.3 looks at.

Warning. Soundness goes one way. It says that derivable sequents are valid; it does not say that a formula true in the row you happened to check is derivable. The atom p is true under the all-T valuation, but $\not\vdash p$: the all-F valuation makes it false, so [Corollary 4.7.3](#) (with Γ empty) applies. Whether **every** valid sequent is derivable is a separate question, answered in Section 4.8.

Quick check. Give a row showing $p \rightarrow r \not\vdash (p \vee q) \rightarrow r$.

Solution. Take $p = \text{F}, q = \text{T}, r = \text{F}$. The premise $p \rightarrow r$ is true, since p is false. The conclusion has hypothesis $p \vee q$, which is true, and conclusion r , which is false, so it is false. By [Corollary 4.7.3](#), no derivation exists. (The rows with p true fail: they make the premise force r true, or make the premise false.)

Satisfiable lists are consistent. Section 4.5 kept two words apart, and soundness joins them in one direction. A satisfiable list Δ is consistent (Definition 3.2.1): if $\Delta \vdash \perp$, soundness would give $\Delta \models \perp$, but a row making Δ true makes $\perp$ false. With Proposition 3.2.2 (b) this is [Corollary 4.7.3](#) once more: a row making the counterexample set $\Gamma, \neg\psi$ true shows it consistent, and so $\Gamma \not\vdash \psi$, as Exercise B4 of Section 3.2 anticipated. Section 4.8 proves the converse.

4.7.7

Test the subgoal

Section 3.1’s search works on subgoals: a goal, and the lines available for it. Soundness gives it a test. Before chasing a subgoal, look for a row making the available lines true and the subgoal false. If there is one, no line standing at this point can hold the subgoal: its context, the premises and the open assumptions, is among the available lines, so the row makes that context true and the subgoal false, while the proof of [Theorem 4.7.2](#) showed that every line k satisfies $\Gamma_k \vdash \varphi_k$. This is the move **test the subgoal**. Section 3.1 met it without the name (“Choosing a side of

$\vee$ can leave a goal that does not follow”): from $p \vee q$, the goal $q \vee p$ tempts $\vee I$ from q , and the row $p = T, q = F$ makes $p \vee q$ true and q false, so the search must split $p \vee q$ first.

Chapter 2 met a harder case. Section 2.5 did not aim at $\vee I$ for the De Morgan direction $\neg(P \wedge Q) \vdash \neg P \vee \neg Q$, because the step from the premise to either disjunct can be broken. Now such a break is a proof: the row $p = T, q = F$ makes $\neg(p \wedge q)$ true and $\neg p$ false, so $\neg(p \wedge q) \not\vdash \neg p$, and the row $p = F, q = T$ gives $\neg(p \wedge q) \not\vdash \neg q$. Both subgoals fail the test, so a derivation of Proposition 2.5.5 (d) cannot end with $\vee I$, which is why Section 2.5 reached for RAA. The test only refutes. A subgoal that passes it may still be hard to reach; whether it can be reached at all is Section 4.8’s question.

Quick check. A search for $p \rightarrow q, \neg q \vdash \neg p$ opens a box p for $\neg I$, with the goal $\perp$. Does the subgoal q pass the test?

Solution. Yes, trivially. The available lines $p \rightarrow q, \neg q, p$ are made true by no row at all (p and $p \rightarrow q$ force q), so no row refutes any subgoal there; and indeed q follows by $\rightarrow E$, then $\perp$ by $\neg E$. An unsatisfiable context lets every subgoal pass.

Remark (Looking ahead). Chapter 8 says what reductio asserts that the other rules do not, and who objected to it.

4.7.8 Exercises

4.7.8.1 A. Check your understanding

Exercise. A1

These parts are about the proof of [Theorem 4.7.2](#). Part (c) uses the derivation of Proposition 2.1.2, written with atoms:

1		$\neg p$	
2			asm, for $\rightarrow I$
3			
4			
5			

		p	

		$\perp$	$\neg E$ 2, 1
		q	$\perp E$ 3
		$p \rightarrow q$	$\rightarrow I$ 2–4

- Why does the proof use strong induction on k , rather than ordinary induction from line k to line $k + 1$?
- In the $\vee E$ case, the induction hypothesis is applied to the last lines b and d of the two boxes. What does it say about each, and which part of [Lemma 4.7.1](#) identifies the contexts?
- List the context of each of the five lines of the derivation above, and check $\Gamma_k \models \varphi_k$ for each.
- Suppose a rule “from $\varphi \rightarrow \psi$ and ψ , infer φ ” were added to the primitive rules. Which

part of the proof would fail? Give a valuation showing it.

(e) Where does Step 4 use that the last line of a derivation lies in no box?

Solution. (a) A rule may cite any available line, however far back, and the last line of a box that closed long ago. The hypothesis must therefore be available for every earlier line, not just for line $k - 1$. For example, line 7 of the derivation of Proposition 2.3.2 uses the claim for line 6, and line 6 uses it for line 5; but line 5 uses it for line 2, three lines back.

(b) It says $\Gamma_b \models \chi$ and $\Gamma_d \models \chi$. By Lemma 4.7.1 (b), $\Gamma_b = \Gamma_k, \varphi$ and $\Gamma_d = \Gamma_k, \psi$: every row making Γ_k and φ true makes χ true, and every row making Γ_k and ψ true makes χ true. It is **not** $\Gamma_k \models \chi$, which is what we are trying to prove.

(c) The box is 2–4.

Line	Context	Check
1	$\neg p$	premise, in its context
2	$\neg p, p$	assumption, in its context
3	$\neg p, p$	no row makes $\neg p$ and p true, so it entails $\perp$
4	$\neg p, p$	the same unsatisfiable context entails q
5	$\neg p$	a row making $\neg p$ true makes p false, so $p \rightarrow q$ true

Line 3 is the $\neg E$ case, line 4 the $\perp E$ case, and line 5 the $\rightarrow I$ case, with $\Gamma_4 = \Gamma_5, p$ by Lemma 4.7.1 (b).

(d) Claim 2, the one-step rules: it would need a new case, and the case is false. From $v(\varphi \rightarrow \psi) = T$ and $v(\psi) = T$ nothing follows about $v(\varphi)$. The valuation with $v(p) = F$, $v(q) = T$ makes $p \rightarrow q$ and q true and p false; with the new rule, the three-line derivation of p from $p \rightarrow q$ and q would have a last line whose context does not entail it.

(e) It concludes that the context of the last line is just the premises. Without that, the last line could depend on an undischarged assumption, and $\Gamma_n \models \psi$ would say nothing about the rows making the premises true.

Exercise. A2: Drill

The parts run from reading off a refuting row, through contexts, to testing a subgoal and deciding a claim.

- (a) Give a refuting row for $p \vee q \vdash p$.
- (b) Give one for $p \rightarrow q, \neg p \vdash \neg q$.
- (c) Does $p = T, q = F$ refute $p \wedge q \vdash q$?
- (d) Is there a refuting row for $p, p \rightarrow q \vdash q$?
- (e) Refute $\vdash p \vee q$.
- (f) Refute $p \rightarrow (q \vee r) \vdash p \rightarrow q$.

- (g) Line 5 of a derivation lies in the boxes opened at lines 2 and 4, and line 1 is the only premise. What is Γ_5 ?
- (h) Line k is justified by $\neg$ I citing the box 3–6. What is Γ_6 , in terms of Γ_k ?
- (i) A search for $p \vee q \vdash q \vee p$ considers ending with $\vee$ I from q . Test the subgoal q .
- (j) The list $p \rightarrow q, q \rightarrow \neg p$ is satisfiable: give a row, and say what follows about $\vdash$.
- (k) Which clause of Definition 4.4.2 does the $\vee$ E case of the proof of [Theorem 4.7.2](#) use?
- (l) Derive or refute: $(p \wedge q) \rightarrow r \vdash (p \rightarrow r) \wedge (q \rightarrow r)$.

Solution.

- (a) $p = \text{F}, q = \text{T}$.
- (b) $p = \text{F}, q = \text{T}$.
- (c) No: it makes the premise false.
- (d) No. The only row making both premises true has q true. (The sequent is derivable, by one line of $\rightarrow$ E.)
- (e) $p = q = \text{F}$. With no premises, every row qualifies.
- (f) $p = \text{T}, q = \text{F}, r = \text{T}$.
- (g) $\varphi_1, \varphi_2, \varphi_4$.
- (h) Γ_k, φ_3 , by [Lemma 4.7.1](#) (b).
- (i) The row $p = \text{T}, q = \text{F}$ makes $p \vee q$ true and q false: $p \vee q \not\vdash q$ ([Corollary 4.7.3](#)), so the subgoal is dead.
- (j) $p = \text{F}, q = \text{T}$ (or both F). So the list is consistent: $p \rightarrow q, q \rightarrow \neg p \not\vdash \perp$. (It is the list of Exercise (B4: Derive $\perp$, or make every member true, used later) (a) without its third member p ; with p , Section 3.2 derived $\perp$.)
- (k) (V4).
- (l) Refute: $p = \text{T}, q = \text{F}, r = \text{F}$ makes the premise true (false hypothesis) and $p \rightarrow r$ false, so the conclusion false; by [Corollary 4.7.3](#) no derivation exists.

4.7.8.2 B. Practice

Exercise. B1: Derivable or not

Determine which of the following are derivable. Justify your answer.

- (a) $(p \wedge q) \rightarrow r \vdash p \rightarrow r$.
- (b) $\neg(p \wedge q) \vdash \neg p$.
- (c) $p \leftrightarrow q, q \vee r \vdash p$.
- (d) $p \rightarrow (q \vee r) \vdash (p \rightarrow q) \vee (p \rightarrow r)$.

Solution. (a) Not derivable. The row $p = \text{T}, q = \text{F}, r = \text{F}$ makes the premise true (its hypothesis $p \wedge q$ is false) and the conclusion $p \rightarrow r$ false. By [Corollary 4.7.3](#) there is no derivation.

(b) Not derivable. The row $p = \text{T}, q = \text{F}$ makes $p \wedge q$ false, so $\neg(p \wedge q)$ true, and $\neg p$ false. By [Corollary 4.7.3](#) there is no derivation.

(c) Not derivable. The row $p = F, q = F, r = T$ makes $p \leftrightarrow q$ true (both sides false) and $q \vee r$ true, and p false.

(d) Derivable: Example (Choosing a side too early) in Section 3.1 derives it, with LEM on q . So no row can refute it: by [Theorem 4.7.2](#) it is valid. A search for a refuting row confirms this. The conclusion is false only if both $p \rightarrow q$ and $p \rightarrow r$ are false, that is, $p = T$ and $q = r = F$; but then $q \vee r$ is false and the premise is false.

The trap is (d). Its conclusion **looks** stronger than the premise, since it seems to decide in advance which disjunct follows from p ; but a derivation exists, and trying to refute it by a row is bound to fail.

Exercise. B2: Ordering the $\vee E$ case

Here are nine lines. Seven of them, in the right order, form the $\vee E$ case of the proof of [Theorem 4.7.2](#). Line k is χ , by $\vee E$ citing line i (holding $\varphi \vee \psi$), the box from line a to line b (opening with φ , ending with χ), and the box from line c to line d (opening with ψ , ending with χ). Give an admissible order, and say why each of the other two lines cannot be used.

- (i) By the induction hypothesis for lines b and d , which come before line k , $\Gamma_b \models \chi$ and $\Gamma_d \models \chi$.
- (ii) Hence $v(\chi) = T$ in either case; since v was arbitrary, $\Gamma_k \models \chi$.
- (iii) Let v be a valuation making every formula of Γ_k true.
- (iv) If $v(\psi) = T$, then v makes every formula of Γ_d true, so $v(\chi) = T$.
- (v) By [Lemma 4.7.1](#) (a), $\Gamma_i \subseteq \Gamma_k$, so v makes every formula of Γ_i true; by the induction hypothesis for line i , $v(\varphi \vee \psi) = T$, so $v(\varphi) = T$ or $v(\psi) = T$.
- (vi) By the induction hypothesis for line k , $\Gamma_k \models \chi$.
- (vii) By [Lemma 4.7.1](#) (b), Γ_b is Γ_k with φ added, and Γ_d is Γ_k with ψ added.
- (viii) If $v(\varphi) = T$, then v makes every formula of Γ_b true, so $v(\chi) = T$.
- (ix) Line b is inside the box, and the box was cited at line k , so the context of line b is Γ_k ; the induction hypothesis for line b then gives $\Gamma_k \models \chi$ directly.

Solution. An admissible order: (iii), (v), (vii), (i), (viii), (iv), (ii).

What each line needs: (v) and (viii), (iv) use the valuation fixed in (iii); (viii) and (iv) use (vii) and (i); (ii) uses the case split of (v) and both of (viii) and (iv). So any order is admissible in which (iii) comes before (v), (viii), and (iv); (vii) and (i) come before (viii) and (iv); (v) comes before (ii); and (viii) and (iv) come before (ii). For instance (vii), (i), (iii), (v), (iv), (viii), (ii) is also admissible: (vii) and (i) do not mention v , and the two cases may be taken in either order.

Line (vi) applies the induction hypothesis to line k itself. The hypothesis of Theorem 4.2.2 covers only the lines $j < k$; (vi) assumes what is to be proved.

Line (ix) uses the context of the citing line for the box's last line. It is false: line b lies inside the box, whose assumption φ is part of its context, so $\Gamma_b = \Gamma_k, \varphi$, not Γ_k . In the derivation of [Example \(Running the claim on a case split\)](#), line 3 has context $p \vee q, p$, while line 6 has context $p \vee q$, and $p \vee q \models q \vee p$ could not be read off from line 3 directly.

Exercise. B3: Derive or refute

For each claim, give a derivation, or a refuting row and the result that makes it a proof of non-derivability.

- (a) $p \leftrightarrow q \vdash \neg p \leftrightarrow \neg q$.
- (b) $p \leftrightarrow \neg q \vdash \neg(p \wedge q)$.
- (c) $p \rightarrow q, q \rightarrow r \vdash r \rightarrow p$.
- (d) $\neg p \vee q \vdash p \rightarrow q$.
- (e) $p \leftrightarrow q, \neg q \vdash p$.
- (f) $\neg(p \rightarrow q) \vdash q \rightarrow p$.

Solution.

(a) **Derive.**

1	$p \leftrightarrow q$	
2	$\neg p$	asm, for $\leftrightarrow$ I
3	q	asm, for $\neg$ I
4	p	$\leftrightarrow$ E 1, 3
5	$\perp$	$\neg$ E 4, 2
6	$\neg q$	$\neg$ I 3–5
7	$\neg q$	asm, for $\leftrightarrow$ I
8	p	asm, for $\neg$ I
9	q	$\leftrightarrow$ E 1, 8
10	$\perp$	$\neg$ E 9, 7
11	$\neg p$	$\neg$ I 8–10
12	$\neg p \leftrightarrow \neg q$	$\leftrightarrow$ I 2–6, 7–11

(b) **Derive.**

1	$p \leftrightarrow \neg q$	
2	$p \wedge q$	asm, for $\neg$ I
3	p	$\wedge$ E 2
4	q	$\wedge$ E 2
5	$\neg q$	$\leftrightarrow$ E 1, 3
6	$\perp$	$\neg$ E 4, 5
7	$\neg(p \wedge q)$	$\neg$ I 2–6

(c) **Refute.** $p = \text{F}, q = \text{F}, r = \text{T}$: both premises have false hypotheses, and $r \rightarrow p$ is $\text{T} \rightarrow \text{F}$. By [Corollary 4.7.3](#), no derivation exists.

(d) **Derive.**

1	$\neg p \vee q$		
2	p		asm, for $\rightarrow$ I
3	$\neg p$		asm, for $\vee$ E
4	$\perp$		$\neg$ E 2, 3
5	q		$\perp$ E 4
6	q		asm, for $\vee$ E
7	q		$\vee$ E 1, 3–5, 6–6
8	$p \rightarrow q$		$\rightarrow$ I 2–7

(e) **Refute.** $p = q = \text{F}$: $p \leftrightarrow q$ is T, $\neg q$ is T, and p is F. By [Corollary 4.7.3](#), no derivation exists.

(f) **Derive**, without RAA.

1	$\neg(p \rightarrow q)$		
2	q		asm, for $\rightarrow$ I
3	p		asm, for $\rightarrow$ I
4	q		R 2
5	$p \rightarrow q$		$\rightarrow$ I 3-4
6	$\perp$		$\neg$ E 5, 1
7	p		$\perp$ E 6
8	$q \rightarrow p$		$\rightarrow$ I 2-7

(c) is the converse of the chain, a way of affirming a conclusion; (e) is a look-alike of $p \leftrightarrow q$, $\neg q \vdash \neg p$, which (a) gives by $\leftrightarrow$ E; (f) looks paradoxical and is not: its premise is false in every row but one, and there $q \rightarrow p$ is true.

Exercise. B4: Test the subgoal

A search for $p \rightarrow (q \vee r)$, $\neg q \vdash p \rightarrow r$ has reached this stage:

1		$p \rightarrow (q \vee r)$	
2		$\neg q$	
3			p asm, for $\rightarrow$ I
			$\vdots$
9			r
10		$p \rightarrow r$ $\rightarrow$ I 3–9	

A student plans to reach r by deriving q first.

(a) Test that subgoal.

(b) Name the move that works, then complete it.

Solution.

- (a) The available lines are $p \rightarrow (q \vee r)$, $\neg q$, p , and the subgoal is q . The row $p = \text{T}$, $q = \text{F}$, $r = \text{T}$ makes the three lines true and q false. So no line of the box itself, whose context is $p \rightarrow (q \vee r)$, $\neg q$, p , can hold q : every line satisfies $\Gamma_k \models \varphi_k$, by the proof of [Theorem 4.7.2](#). The plan is dead. (q may still stand as the assumption of a case box inside it, whose lines have a larger context; that is how (b) uses it.)
- (b) Use the premise by its main connective: $\rightarrow\text{E}$ gives $q \vee r$; split it, closing the case q by the clash with $\neg q$.

1		$p \rightarrow (q \vee r)$		
2		$\neg q$		
3			p asm, for $\rightarrow$ I	
4			$q \vee r$ $\rightarrow$ E 1, 3	
5				q asm, for $\vee$ E
6				$\perp$ $\neg$ E 5, 2
7				r $\perp$ E 6
8				r asm, for $\vee$ E
9			r	$\vee$ E 4, 5–7, 8–8
10		$p \rightarrow r$	$\rightarrow$ I 3–9	

(Proposition 2.3.4 gives line 9 from lines 4 and 2 in one step, as a derived rule; either is correct.)

4.7.8.3 C. Going deeper

Exercise. C1: A safe rule and an unsafe one

- (a) Add to the primitive rules a one-step rule X: from a line $\varphi \wedge \psi$, infer $\varphi \vee \psi$. Show that [Theorem 4.7.2](#) still holds for derivations that may use X.
- (b) Add instead a one-step rule Y: from a line $\varphi \vee \psi$, infer $\varphi \wedge \psi$. Show that soundness fails: name the case of the proof that breaks, give a derivation using Y of a sequent with a refuting row, and give the row.

Solution (Hint 1). The proof of [Theorem 4.7.2](#) treats one rule per case. Adding a rule adds one case to Claim 2 and changes nothing else.

Solution (Hint 2). For (b), start a derivation from the premise p , and look for a way to write $p \vee q$ first.

Solution. (a) The proof of [Theorem 4.7.2](#) goes through with one more case in Claim 2. Every other part is unchanged: Step 1 and Step 4 do not look at the rules, [Lemma 4.7.1](#) (a) holds for any cited line, and (b) concerns only box rules, and X is not one. *The case X.* Line k is $\varphi \vee \psi$, citing a line i holding $\varphi \wedge \psi$. By (*), $v(\varphi \wedge \psi) = T$, so $v(\varphi) = T$ by (V3), and $v(\varphi \vee \psi) = T$ by (V4). So every line of a derivation with X still satisfies $\Gamma_k \models \varphi_k$, and soundness holds.

(b) The new case of Claim 2 fails: from $v(\varphi \vee \psi) = T$, (V4) gives only that one disjunct is true, and (V3) needs both. With Y, this is a derivation of $p \vdash q$:

1	p	
2	$p \vee q$	$\vee I$ 1
3	$p \wedge q$	Y 2
4	q	$\wedge E$ 3

The row $p = T, q = F$ makes the premise true and the conclusion false, so $p \not\models q$, although $p \vdash q$ with Y. The failure is at line 3: its context is p , and the row makes p and line 2 true but line 3 false.

Marking scheme. (10 points)

- 2: (a) says that only one case is added, and why the rest of the proof (Steps 1 and 4, [Lemma 4.7.1](#)) is unaffected.
- 3: (a) the case for X, using the hypothesis for the cited line through [Lemma 4.7.1](#) (a) and the clauses (V3), (V4).
- 2: (b) names the failing case, with the reason (one true disjunct does not give a true conjunction).
- 3: (b) a correct derivation with Y and a refuting row for its sequent.

Fatal errors (0 for the step): a “refuting row” that makes a premise false; in (a), checking X only on atoms without the context argument.

Exercise. C2: Exit check: soundness

Close the section first, and write from memory.

Write out, in full, the $\rightarrow I$, $\neg I$ and $\vee E$ cases of the proof of [Theorem 4.7.2](#): state the claim being proved for every line, set up the induction, and prove the three cases, naming the reason for each step.

Solution. The claim. Fix a derivation in primitive rules with lines $1, \dots, n$, and write φ_k for the formula on line k and Γ_k for its context (the premises and the assumptions of the boxes containing line k). We prove $\Gamma_k \models \varphi_k$ for every k with $1 \leq k \leq n$, by [Theorem 4.2.2](#) applied to “if $1 \leq k \leq n$, then $\Gamma_k \models \varphi_k$ ” (for $k = 0$ it holds vacuously): fix k with $1 \leq k \leq n$, assume $\Gamma_j \models \varphi_j$ for every line $j < k$, and let v be a valuation making every formula of Γ_k true. Every line and box that line k cites lies before line k , so the hypothesis applies to the cited lines and to the last lines of the cited boxes. We show $v(\varphi_k) = T$.

$\rightarrow I$. Line k is $\varphi \rightarrow \psi$, citing the box from line a (the assumption φ) to line b (the formula ψ).

By [Lemma 4.7.1](#) (b), $\Gamma_b = \Gamma_k, \varphi$. Since $b < k$, the hypothesis gives $\Gamma_k, \varphi \models \psi$. If $v(\varphi) = F$, then $v(\varphi \rightarrow \psi) = T$ by (V5). If $v(\varphi) = T$, then v makes every formula of Γ_k, φ true, so $v(\psi) = T$, and $v(\varphi \rightarrow \psi) = T$ by (V5).

$\neg I$. Line k is $\neg\varphi$, citing the box from line a (the assumption φ) to line b (the formula $\perp$). By [Lemma 4.7.1](#) (b) and the hypothesis for line b , $\Gamma_k, \varphi \models \perp$. Since $v(\perp) = F$ under every valuation (V1), no valuation makes all of Γ_k, φ true. As v makes Γ_k true, $v(\varphi) = F$, so $v(\neg\varphi) = T$ by (V2).

$\vee E$. Line k is χ , citing line i (holding $\varphi \vee \psi$) and the boxes from a to b (φ to χ) and from c to d (ψ to χ). Line i is available at line k , so $\Gamma_i \subseteq \Gamma_k$ by [Lemma 4.7.1](#) (a); v makes Γ_i true, and the hypothesis for line i gives $v(\varphi \vee \psi) = T$, so $v(\varphi) = T$ or $v(\psi) = T$ by (V4). By [Lemma 4.7.1](#) (b), $\Gamma_b = \Gamma_k, \varphi$ and $\Gamma_d = \Gamma_k, \psi$, and the hypothesis for lines b and d gives $\Gamma_k, \varphi \models \chi$ and $\Gamma_k, \psi \models \chi$. If $v(\varphi) = T$, then v makes Γ_k, φ true, so $v(\chi) = T$; if $v(\psi) = T$, likewise by the second box.

Conclusion. With the other cases, the induction gives $\Gamma_k \models \varphi_k$ for every line. The last line lies in no box, so its context is the premises, which all belong to Γ ; a valuation making Γ true makes the premises true, and so makes the last line ψ true: $\Gamma \models \psi$.

Marking scheme. (10 points)

- 2: states the claim $\Gamma_k \models \varphi_k$ for **every** line, with contexts.
- 2: uses the course-of-values hypothesis on earlier lines only, and says why the cited lines and boxes are earlier.
- 2: identifies the context of each box's last line as Γ_k plus the box's assumption, citing [Lemma 4.7.1](#) (b) (and (a) for the cited line in $\vee E$).
- 1 + 1 + 1: the semantic step in each of the three cases (the rows where the assumption is false for $\rightarrow I$; no row satisfies the box's context for $\neg I$; the split on which disjunct is true for $\vee E$).
- 1: the conclusion for the last line.

A proof that inducts on the length of the derivation instead, and handles open boxes correctly, earns the same points step for step. Fatal errors (0 for the step): applying the hypothesis to line k itself; using Γ_k as the context of a box's last line.

Exercise. C3: A changed table

Replace (V4) by: $v(\varphi \vee \psi) = T$ exactly when **exactly one** of $v(\varphi)$, $v(\psi)$ is T. Keep every other clause and every rule.

- Which cases of the proof of [Theorem 4.7.2](#) still go through?
- For each rule whose case fails, give a derivation and a row that is a counterexample for the new table.
- Is $p \vdash p \vee q$ still derivable? Is it valid for the new table? What does this show?

Solution.

- Every case except $\vee I$. The rules that do not mention $\vee$ use other clauses, which are unchanged. The $\vee E$ case needs only that a true disjunction has a true side, and that still holds: exactly one side is true.

- (b) $\vee I$: the derivation 1 p ; 2 $p \vee q$, $\vee I$ 1. The row $p = q = T$ makes p true and, for the new table, $p \vee q$ false.
- (c) Derivable: the rules did not change. Not valid for the new table: the row of (b). So soundness is a fact about the rules **and** the tables together; change a table, and the same rules can lead from truth to falsehood.

Marking scheme. (10 points)

- 4: (a), with the argument for the $\vee E$ case.
- 4: (b), the derivation and the row, checked.
- 2: (c), both answers and the conclusion.

Fatal errors (0 for the step): in (a), declaring the $\vee E$ case broken without a row that breaks it.

Exercise. C4: Challenge: $\perp$ read as a letter

Read $\perp$ as an ordinary letter. A **letter valuation** w gives T or F to every atom **and to** $\perp$, computes $\wedge, \vee, \rightarrow, \leftrightarrow$ by (V3)–(V6), and sets $w(\neg\varphi) = T$ exactly when $w(\varphi) = F$ or $w(\perp) = T$. Write $\Gamma \models_{\ell} \psi$ when every letter valuation making every formula of Γ true makes ψ true.

- (a) Show that every line of a derivation that uses no $\perp E$ and no RAA, and cites no derived rule whose own derivation uses them, satisfies $\Gamma_k \models_{\ell} \varphi_k$. Give the cases $\neg E$ and $\neg I$, and say why the others are those of [Theorem 4.7.2](#).
- (b) Deduce that $\perp \vdash p$ has no derivation without $\perp E$ and RAA (and without derived rules derived with them).
- (c) Derive $\perp \vdash p$ with RAA and without $\perp E$, and say why (b) must exclude RAA as well.

Solution.

- (a) Step 1 of the proof expands each line citing a derived rule into a pasted derivation of that rule, which by hypothesis uses neither $\perp E$ nor RAA; so the expanded derivation is in the remaining primitive rules (and a derived rule derived from derived rules of this kind expands the same way, one layer at a time). Steps 1, 2 and 4 do not look at the clauses, and the cases of premises, assumptions, R and the rules for $\wedge, \vee, \rightarrow, \leftrightarrow$ use only (V3)–(V6), which are unchanged; they go through word for word with w for v .
- $\neg E$: w makes Γ_k true, so by (*) it makes φ and $\neg\varphi$ true. Now $w(\neg\varphi) = T$ with $w(\varphi) = T$ forces $w(\perp) = T$, and line k is $\perp$.
- $\neg I$: by (**), every letter valuation making Γ_k and φ true makes $\perp$ true. If $w(\varphi) = F$, then $w(\neg\varphi) = T$. If $w(\varphi) = T$, then w makes Γ_k and φ true, so $w(\perp) = T$, and again $w(\neg\varphi) = T$.
- (b) The letter valuation with $w(\perp) = T$ and $w(p) = F$ makes the premise true and the conclusion false. So $\perp \models_{\ell} p$ fails, and by (a) no derivation of $\perp \vdash p$ without $\perp E$ and RAA, and without derived rules derived with them, exists. (Many derived rules of Chapter 2 use one of the two, Proposition 2.5.1 (b) (with RAA) and Proposition 2.1.1 (with $\perp E$) among them; a derivation that avoids the two only on its face is not covered.)
- (c)

1	$\perp$	
2	$\neg p$	asm, for RAA
3	$\perp$	R 1
4	p	RAA 2–3

So with RAA the rule $\perp E$ is derivable, and (a) must fail for RAA. It does: with $w(\perp) = T$, every negation is true, so the box of lines 2–3 satisfies (**), and yet $w(p) = F$.

Marking scheme. (10 points)

- 2: the carried-over cases named, with the reason (Steps 1, 2, 4 and the clauses (V3)–(V6) unchanged), and Step 1's restriction to derived rules derived without $\perp E$ and RAA.
- 2: the $\neg E$ case.
- 3: the $\neg I$ case, both values of $w(\varphi)$.
- 1: (b), the letter valuation and the conclusion.
- 2: (c), the derivation and the failing case.

Fatal errors (0 for the step): in (b), using the ordinary valuations, under which $\perp \models p$ holds.

Exercise. C5: Challenge: two three-valued tables

A **Kleene valuation** u_K assigns one of the values 0, $\frac{1}{2}$, 1 to each atom and is extended by $u_K(\perp) = 0$, $u_K(\neg\varphi) = 1 - u_K(\varphi)$, $u_K(\varphi \wedge \psi) = \min\{u_K(\varphi), u_K(\psi)\}$, $u_K(\varphi \vee \psi) = \max\{u_K(\varphi), u_K(\psi)\}$, $u_K(\varphi \rightarrow \psi) = \max\{1 - u_K(\varphi), u_K(\psi)\}$. (This extension differs from the three-valued valuation of the optional excursion below, whose $\neg$ gives $\neg\varphi$ the value 0 at $\frac{1}{2}$.) For a set D of **designated** values, say that Γ D -entails ψ if every Kleene valuation u_K giving every formula of Γ a value in D gives ψ a value in D .

- For $D = \{1\}$, show that the case $\rightarrow I$ fails: give a box that is correct for D -entailment and a u_K at which the line that $\rightarrow I$ writes is undesignated.
- For $D = \{\frac{1}{2}, 1\}$, show that the case $\rightarrow E$ fails.
- Is $p \vee \neg p$ designated under every u_K in (a)? In (b)?

Solution.

- The one-line box p : p D -entails p . $\rightarrow I$ writes $p \rightarrow p$, and at $u_K(p) = \frac{1}{2}$, $u_K(p \rightarrow p) = \max\{\frac{1}{2}, \frac{1}{2}\} = \frac{1}{2} \notin D$.
- The premises $p \rightarrow q$, p and the conclusion q , at $u_K(p) = \frac{1}{2}$, $u_K(q) = 0$: $u_K(p \rightarrow q) = \max\{\frac{1}{2}, 0\} = \frac{1}{2} \in D$ and $u_K(p) = \frac{1}{2} \in D$, but $u_K(q) = 0 \notin D$.
- In (a), no: $u_K(p \vee \neg p) = \max\{\frac{1}{2}, \frac{1}{2}\} = \frac{1}{2}$ at $u_K(p) = \frac{1}{2}$. In (b), yes: $\max\{x, 1 - x\} \geq \frac{1}{2}$ for each of the three values x .

So the first choice of D loses excluded middle and the box rule for $\rightarrow$; the second keeps excluded middle and loses modus ponens.

Marking scheme. (10 points)

- 4: (a), the box, the valuation and the computation.

- 3: (b), the valuation and the three values.
- 3: (c), both parts, with the computation.

Fatal errors (0 for the step): a “failure” in (a) or (b) at a valuation that does not designate every premise.

4.7.9 Optional excursion: reductio says something new

Optional: nothing later depends on this excursion. It answers a question Chapter 2 raised.

A **three-valued valuation** u assigns to each atom one of the values 0, $\frac{1}{2}$, 1, ordered $0 < \frac{1}{2} < 1$. Think of 1 as “true”, 0 as “false”, and $\frac{1}{2}$ as “not yet settled”.

$u(\varphi \rightarrow \psi)$	$u(\psi) = 0$	$u(\psi) = \frac{1}{2}$	$u(\psi) = 1$	$u(\varphi)$	$u(\neg\varphi)$
$u(\varphi) = 0$	1	1	1	0	1
$u(\varphi) = \frac{1}{2}$	0	1	1	$\frac{1}{2}$	0
$u(\varphi) = 1$	0	$\frac{1}{2}$	1	1	0

An implication gets 1 when its conclusion is at least as true as its hypothesis, and otherwise the value of its conclusion; a negation is 1 only on 0.

Chapter 2 introduced RAA in Section 2.5 because the rules before it stalled on $\neg\neg P \vdash P$, and said that Chapter 4 would prove that no derivation of P from $\neg\neg P$ avoids RAA. Truth tables cannot do it. The row method only proves non-derivability from **all** the rules, and with RAA the sequent is derivable, so no two-valued row refutes it. The trick is to run the proof of [Theorem 4.7.2](#) with a different notion of truth, one with three values, under which every rule except RAA still passes.

A three-valued valuation is extended to all formulas by recursion (Theorem 4.3.4):

$$\begin{aligned}
 u(\perp) &= 0, & u(\neg\varphi) &= u(\varphi \rightarrow \perp), \\
 u(\varphi \wedge \psi) &= \min\{u(\varphi), u(\psi)\}, & u(\varphi \vee \psi) &= \max\{u(\varphi), u(\psi)\}, \\
 u(\varphi \rightarrow \psi) &= \begin{cases} 1 & \text{if } u(\varphi) \leq u(\psi), \\ u(\psi) & \text{otherwise,} \end{cases} \\
 u(\varphi \leftrightarrow \psi) &= \min\{u(\varphi \rightarrow \psi), u(\psi \rightarrow \varphi)\}.
 \end{aligned}$$

So $u(\neg\varphi) = 1$ if $u(\varphi) = 0$, and $u(\neg\varphi) = 0$ otherwise. On the values 0 and 1 alone these are exactly the tables of Definition 4.4.2, with 1 for T and 0 for F. For a list Γ , let $u(\Gamma)$ be the least of the values $u(\gamma)$ for γ in Γ , and $u(\Gamma) = 1$ if Γ is empty. Write $\Gamma \models_3 \psi$ if $u(\Gamma) \leq u(\psi)$ for every three-valued valuation u : the conclusion is at least as true as the weakest premise.

Theorem 4.7.4. Reductio Is Not Derivable from the Other Rules

Optional: nothing later depends on this.

Neither $\neg\neg p \vdash p$ nor $\vdash p \vee \neg p$ has a derivation that uses only the primitive rules other than RAA (and derived rules derived from those).

Idea. Repeat the proof of [Theorem 4.7.2](#) with $\models_3$ in place of $\models$: every line of a derivation without RAA satisfies $\Gamma_k \models_3 \varphi_k$. Then show that both sequents fail for $\models_3$, at $u(p) = \frac{1}{2}$. The one case that breaks is RAA, and that is the point.

Proof. As in Step 1 of the proof of [Theorem 4.7.2](#), a derived rule derived without RAA abbreviates a pasted derivation in the primitive rules other than RAA, so it is enough to consider those. Fix a derivation that uses only the primitive rules other than RAA, with lines $1, \dots, n$, formulas φ_k and contexts Γ_k . We show $\Gamma_k \models_3 \varphi_k$ for every k with $1 \leq k \leq n$, by Theorem 4.2.2 exactly as in Step 2 of the proof of [Theorem 4.7.2](#), the cases $k = 0$ and $k > n$ being vacuous. Fix k , assume the claim for all earlier lines, and fix a three-valued valuation u ; write $m = u(\Gamma_k)$. We must show $m \leq u(\varphi_k)$. Two facts do the bookkeeping. If line k cites line i , then $\Gamma_i \subseteq \Gamma_k$ ([Lemma 4.7.1](#) (a)), and the least value over a longer list is at most the least over a shorter one, so $m \leq u(\Gamma_i) \leq u(\varphi_i)$. If line k cites a box from line a to line b by a box rule, then $\Gamma_b = \Gamma_k, \varphi_a$ ([Lemma 4.7.1](#) (b)), so the hypothesis for line b says $\min\{m, u(\varphi_a)\} \leq u(\varphi_b)$. Below, $x = u(\varphi)$, $y = u(\psi)$, $z = u(\chi)$.

Premises and assumptions: φ_k is in Γ_k , so $m \leq u(\varphi_k)$. R: $m \leq u(\varphi_i) = u(\varphi_k)$. $\wedge$ I: $m \leq x$ and $m \leq y$, so $m \leq \min\{x, y\}$. $\wedge$ E: $m \leq \min\{x, y\} \leq x$, and likewise $\leq y$. $\vee$ I: $m \leq x \leq \max\{x, y\}$.

$\rightarrow$ E: $m \leq u(\varphi \rightarrow \psi)$ and $m \leq x$. If $x \leq y$, then $m \leq x \leq y$. Otherwise $u(\varphi \rightarrow \psi) = y$, so $m \leq y$. $\leftrightarrow$ E: $m \leq u(\varphi \leftrightarrow \psi) \leq u(\varphi \rightarrow \psi)$ and $m \leq x$, so $m \leq y$ as for $\rightarrow$ E; the second form is symmetric.

$\neg$ E: $m \leq x$ and $m \leq u(\neg\varphi)$. If $x = 0$, then $m \leq 0$; if $x > 0$, then $u(\neg\varphi) = 0$ and $m \leq 0$. So $m = 0 = u(\perp)$. $\perp$ E: $m \leq u(\perp) = 0$, so $m = 0 \leq u(\neg\varphi)$.

$\rightarrow$ I: the hypothesis gives $\min\{m, x\} \leq y$. If $x \leq y$, then $u(\varphi \rightarrow \psi) = 1 \geq m$. Otherwise $y < x$ and $u(\varphi \rightarrow \psi) = y$; since $\min\{m, x\} \leq y < x$, the minimum is not x , so it is m , and $m \leq y$. $\leftrightarrow$ I: the same argument for each box gives $m \leq u(\varphi \rightarrow \psi)$ and $m \leq u(\psi \rightarrow \varphi)$, so m is at most their minimum.

$\neg$ I: the hypothesis gives $\min\{m, x\} \leq u(\perp) = 0$. If $x = 0$, then $u(\neg\varphi) = 1 \geq m$. If $x > 0$, the minimum is 0 and is not x , so $m = 0 \leq u(\neg\varphi)$.

$\vee$ E: $m \leq \max\{x, y\}$, and the boxes give $\min\{m, x\} \leq z$ and $\min\{m, y\} \leq z$. If $x \geq y$, then $m \leq x$, so $\min\{m, x\} = m$ and $m \leq z$. If $y > x$, then $m \leq y$ and $m \leq z$ from the second box.

This covers every rule used, so $\Gamma_k \models_3 \varphi_k$ for every line, and in particular for the last line, whose context is the premises. As in Step 4 of the proof of [Theorem 4.7.2](#), a list containing all the premises has least value at most theirs, so every sequent $\Gamma \vdash \psi$ derived without RAA satisfies $\Gamma \models_3 \psi$.

Now let $u(p) = \frac{1}{2}$. Then $u(\neg p) = 0$, since $\frac{1}{2} > 0$, and $u(\neg\neg p) = 1$, since $u(\neg p) = 0$. So $u(\neg\neg p) = 1 > \frac{1}{2} = u(p)$, and $\neg\neg p \models_3 p$ fails. Also $u(p \vee \neg p) = \max\{\frac{1}{2}, 0\} = \frac{1}{2} < 1$, the value of the empty list, so $\models_3 p \vee \neg p$ fails. Hence neither sequent is derivable without RAA. ■

The case that was not written is RAA itself. Its box would give $\min\{m, u(\neg\varphi)\} \leq 0$, and we

would need $m \leq x$. At $u(\varphi) = \frac{1}{2}$ the box gives nothing, since $u(\neg\varphi) = 0$ already, while m may be 1. This is the three-valued face of the question asked in the Idea of [Theorem 4.7.2](#): the two-valued RAA case used that $v(\varphi) = \text{F}$ forces $v(\neg\varphi) = \text{T}$, and with a third value, φ and $\neg\varphi$ can both fail to be true.

Remark. The same values settle Peirce’s law, $((p \rightarrow q) \rightarrow p) \rightarrow p$, which Section 2.5 derived with RAA (Proposition 2.5.4): at $u(p) = \frac{1}{2}$, $u(q) = 0$, we get $u(p \rightarrow q) = 0$, $u((p \rightarrow q) \rightarrow p) = 1$, and the whole formula has value $\frac{1}{2}$, so it too needs RAA. The values do not refute everything Chapter 2 derived with RAA: the De Morgan direction $\neg(p \wedge q) \vdash \neg p \vee \neg q$ passes this test, and so does $\vdash (p \rightarrow q) \vee (q \rightarrow p)$. Neither of these is derivable without RAA either, but showing that needs a different kind of model, and this book states it without proof, as Section 2.5’s Looking-ahead box announced. RAA says something the other rules cannot.

A technique, named. The proof is an instance of a general technique: to show that a rule cannot be derived from the others, find a new notion of “true” under which the other rules are sound and that rule is not. Here the notion has three values; Exercise C4 uses two values with $\perp$ read as an ordinary letter, and Exercise C5 two other three-valued tables. Each notion is a different semantics. For the two-valued tables the rules are not only sound but complete, which is Section 4.8’s theorem.