Mathematical Notation

Mathematical Notation

A Student's Field Guide

Gaurav Tiwari

gauravtiwari.org

First edition, 2026.

gauravtiwari.org

To everyone who has stared at a symbol
and been too embarrassed to ask.

Contents

Preface

Nobody ever taught you to read this.

Somewhere in your first year, the writing changed. School mathematics was mostly words and numbers with the occasional symbol. University mathematics is symbols with the occasional word, and at no point did anyone stop and say what the symbols mean, how you say them out loud, or which ones are traps.

You were expected to absorb it. Most people do, roughly, over about two years, by pattern-matching and by being wrong in front of other people. That works, and it's a slow and unpleasant way to learn something that could have been a book.

This is that book.

What it is

A field guide. Every symbol you're likely to meet from first year through early graduate work, with four things for each one: the glyph, how you say it aloud, what it means, and what it does *not* mean.

That last field is the reason the book exists. Most notation trouble isn't ignorance, it's a near-miss: reading $\Rightarrow$ as "causes", reading $\subset$ as strict when the author meant it isn't, reading a bar as an average when it means a closure, assuming log is base 10 in a paper where it's base 2. Knowing roughly what a symbol means is what gets you into trouble, because it's enough to keep reading and not enough to be right.

How to use it

Three ways, and all three are intended.

Read Part I straight through. It's short, and it's about grammar rather than vocabulary: how an expression is built, what binds what, why quantifier order changes the meaning, and how to say things aloud. If you only read one part, read this one. It's the part that makes the rest usable.

Browse Part II. Nine chapters of symbol entries grouped by subject. You can read them in order, and most people won't. Find your symbol, read the entry, go back to whatever you were doing.

Keep Part III and the Toolkit for when something goes wrong. Part III is about collisions and misreadings, and the Toolkit has the look-up-by-shape index for when you have a symbol and no name for it, which is the single most common way a student gets stuck.

A note on fields

Notation is not universal, and books that pretend otherwise do real damage. The same symbol means different things in different subjects, the same idea gets different symbols, and a convention that's absolutely standard in one department is unheard of two floors up.

So where a symbol collides, it's flagged. Where a field does something its own way, it's labelled. Appendix F collects the differences by subject, and Appendix D is the collision table: one symbol, every meaning, sorted.

What it isn't

Not a mathematics textbook. It tells you what $\int_a^b f$ means as notation, not how to evaluate it. It tells you that ε is conventionally a small positive quantity, not how to run an epsilon-delta proof.

Not a typesetting manual either, though Appendix E gives the LaTeX command for every symbol in the book, because you'll eventually need to write these as well as read them.

Gaurav Tiwari
gauravtiwari.org

Part I

How Notation Works

1

Nobody Teaches You to Read This

Here's a line from a first-year analysis course.

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n > N : |a_n - L| < \varepsilon$$

If you can say that out loud, in words, at normal speaking speed, you can skip this chapter.

Most people can't, and the reason isn't that they don't understand convergence. It's that nobody has ever asked them to say a formula aloud, so the symbols have stayed visual shapes rather than becoming language.

1.1 Why this particular gap exists

School mathematics is mostly prose with numbers in it. University mathematics is mostly notation with prose in it. The switch happens over about six weeks and nobody announces it.

Worse, the people teaching you genuinely cannot see the problem. Your lecturer has read $\forall \varepsilon > 0$ perhaps fifty thousand times. It isn't notation to them any more, it's a word. Asking them what it means gets you a correct answer to a different question, because they've forgotten that the difficulty was ever mechanical.

> *Fluency in notation is a separate skill from understanding the mathematics. You can have one without the other in either direction, and almost nobody tells you they're separate.*

1.2 The four ways notation actually goes wrong

Not one problem, four. They need different fixes, so it's worth knowing which one you have.

Modes 3 and 4 are what this book is mostly about, because they're the ones a symbol list won't touch.

1.3 Mode 2, and why saying it aloud matters so much

This sounds like a soft skill and it's the most practical thing in the book.

Working memory holds sounds much better than shapes. When you read $|a_n - L| < \varepsilon$ as a picture, you're storing an image and you'll lose it within a line or two. When you read it as "the distance from a_n to L is less than epsilon", you're storing a sentence, and sentences survive.

Try it now, on that convergence definition:

1 You don't know the symbol
The easy case. Look it up. Part II, or the look-up-by-shape index.

↓

2 You can't say it aloud
You recognise it and have no words. So you can't hold it in your head, can't discuss it, can't check it.

↓

3 You know it approximately
The dangerous one. Close enough to keep reading, wrong enough to reach a false conclusion.

↓

4 You can't parse the structure
You know every symbol and still can't see what applies to what. A grammar problem, not a vocabulary one.

Figure 1.1: The four failure modes. Only the first is about not knowing the symbol, which is why simply looking symbols up doesn't fix as much as you'd expect.

> For every epsilon greater than zero, there exists a natural number N, such that for every n greater than N, the distance from a_n to L is less than epsilon.

Thirty-five words, and once it's a sentence you can hear that the last clause is inside the second, which is inside the first. That nesting is the whole content of the definition, and it's invisible when the line is a row of shapes.

HOW IT IS ACTUALLY USED

Watch two researchers at a blackboard. They almost never read a formula silently. They say it, badly, with placeholder words: "so this thing here, sum over i of the whatever, is bounded by…" Fluent readers of notation are constantly converting to speech. It isn't a beginner's crutch, it's how the professionals do it.

1.4 Mode 3: the near-miss, with three examples

Each of these is a real, common, mark-costing misreading. Each is a small error that produces a confidently wrong answer.

The arrow that isn't causation. $P \Rightarrow Q$ says: whenever P holds, Q holds. It says nothing about why, nothing about time order, and nothing at all about the case where P is false. Students read it as "P causes Q" or "P then Q" and go on to conclude that $\neg P \Rightarrow \neg Q$, which does not follow. Chapter 6.

The subset that might be equal. In most modern writing $A \subset B$ allows $A = B$. In some older and some European texts it means strictly smaller. If you assume strict and the author didn't, you will construct a counterexample to a true theorem and spend an evening confused. Chapter 7.

The log with no base. In pure mathematics log usually means natural log. In engineering it often means base 10. In computer science it nearly always means base 2. In statistics it means natural log. All four communities consider their reading obvious. Chapter 16.

None of those is exotic. All three appear in first-year courses, and all three are invisible until somebody names them.

1.5 Mode 4: structure, briefly

Consider these two, which use exactly the same symbols:

$$\forall x\, \exists y : x < y \qquad \text{and} \qquad \exists y\, \forall x : x < y$$

The first says every number has something bigger than it, which is true of the integers. The second says there's a single number bigger than everything, which is false. Same six symbols, opposite truth values, and the only difference is the order of two quantifiers.

That's a grammar problem, and Chapters 2 and 3 are about it.

1.6 What this book gives you

Part I	The grammar. How expressions are built, what binds what, and how to say things aloud
Part II	The symbols, by subject. Glyph, read-aloud, meaning, and the misreading
Part III	Collisions, the common misreadings, decorations, and how to write notation others can read
Toolkit	Greek alphabet said aloud, look-up by shape, a read-aloud card, the collision table, LaTeX commands, conventions by field

1.7 One habit to start now

Keep a notation page for each module. One sheet, at the front of the notebook. Every time a symbol appears that you can't say aloud, write it down with the date and the lecture.

Two things happen. You'll look most of them up within a week, because having written it down makes the gap feel small and closable. And the ones you never look up turn out to be the ones you never needed, which is useful information too.

PRACTISE THIS

Take the convergence definition at the top of this chapter and say it aloud, right now, without looking at the sentence version. Then find any displayed formula in your current lecture notes and do the same. If you stall on a symbol, write it on your notation page. That page is the reading list this book is for.

2

The Grammar of an Expression

An expression is built, not written. It has a structure, and the structure is what you have to see before the meaning arrives.

This chapter is about seeing it.

2.1 Everything is one of four things

Every symbol in a formula does exactly one job.

Job	Examples	What to ask
Operand	x, 3, π, A	What is this a name for?
Operator	$+$, $\times$, $\int$, ∇	What does it take in, and what comes out?
Relation	$=$, $<$, $\in$, $\subseteq$	What two things is it comparing? Is the whole line a claim?
Grouping	(), [], {}, | |, the bar of a fraction	Where does it start and stop?

The single most useful distinction on that table is **operator versus relation**, because it tells you whether you're looking at a *thing* or a *claim*.

$x + 3$ is a thing. It has a value. You can't argue with it.

$x + 3 = 7$ is a claim. It's true or false, or it's a condition on x. You can prove it, disprove it, or solve it.

Students lose marks by treating a claim as a thing (writing a chain of equations that never asserts anything) and by treating a thing as a claim ("prove $x + 3$"). Both errors dissolve the moment you can spot which symbol is the relation.

> *Find the relation symbol first. It's the main verb of the sentence. Everything to its left is one noun phrase; everything to its right is another. If there's no relation symbol, the line isn't a statement.*

2.2 Precedence, which nobody states and everybody assumes

Written mathematics leaves out brackets that a computer would demand, because there's an agreed order of operations. Here it is, tightest first.

So $\sin x^2$ means $\sin(x^2)$, because superscripts bind tighter than function application. And $\forall x\, P(x) \Rightarrow Q(x)$ is ambiguous enough that careful authors bracket it, because quantifiers are loosest and it's genuinely unclear whether the $\forall x$ covers the whole implication.

1. **Superscripts and subscripts** x^2, a_n
2. **Function application** $\sin x$, $f(x)$, $\log n$
3. **Multiplication and division** ab, a/b, $a \cdot b$
4. **Addition and subtraction** $a+b$, $a-b$
5. **Relations** $=, <, \in, \subseteq$
6. **Logical connectives, then quantifiers** $\neg$, then $\wedge$, then $\vee$, then $\Rightarrow$, then $\forall$ and $\exists$. Loosest of all.

Figure 2.1: Precedence, tightest binding at the top. Anything higher on this list grabs its arguments before anything lower gets a chance.

2.3 Two notations that hide their brackets

Fractions. The bar is a bracket. $\dfrac{a+b}{c}$ is $(a+b)/c$, and writing it inline as $a+b/c$ changes the value. When you convert a displayed fraction to one line, put the brackets back.

Radicals. The bar over the root does the same job. $\sqrt{a+b}$ is not $\sqrt{a}+b$.

Both are the commonest source of arithmetic errors when copying a formula from a screen into a calculator.

2.4 $\sin^2 x$, and the exception that catches everyone

For trigonometric functions, $\sin^2 x$ means $(\sin x)^2$. The exponent is on the output.

But $\sin^{-1} x$ means the *inverse function*, $\arcsin x$, not $1/\sin x$. Same-looking notation, two different jobs, and -1 is the only exponent that behaves this way.

MISREAD

$\sin^{-1}(0.5) = 1/\sin(0.5) = 2.09$

READ CORRECTLY

$\sin^{-1}(0.5) = \arcsin(0.5) = \pi/6 \approx 0.524$

This is genuinely bad notation and it's universal, so you have to know it. Some authors avoid it by writing arcsin, and you should too when you have the choice.

2.5 Reading an expression from the outside in

The instinct is to read left to right. The better instinct is to find the outermost structure first, then descend.

Take:

$$\sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}$$

1. **Outermost:** it's a sum. Everything else is inside it.

2. **What varies:** k, from 1 to n. So the answer depends on n, not on k.
3. **The summand:** a fraction. Bottom is k. Top is $(-1)^{k+1}$, which is $+1$ when k is odd and -1 when k is even.
4. **Say it:** "the sum from k equals one to n of minus one to the k plus one, over k", or better, "the alternating harmonic sum up to n".

Four steps, and the last one is the payoff: once you've descended, you can usually name the whole thing in a phrase, and the phrase is what you carry forward.

2.6 The scope question

Scope is how far a symbol's influence reaches. It's the question that resolves most Mode 4 confusion.

$\sum$, $\prod$, $\int$	Extends over the whole following expression unless bracketed. $\int f(x) + g(x)\,dx$ is ambiguous; the dx marks the end
$\forall$, $\exists$	Extends to the end of the statement, or to the closing bracket
$\sqrt{\ }$, fraction bar	Exactly what's under or over it
A subscript	Only the symbol it's attached to. x_{n+1} is one term; $x_n + 1$ is a term plus one
d or ∂	Up to the next d, or to the end

HOW IT IS ACTUALLY USED

When a formula is genuinely ambiguous, working mathematicians don't parse it. They look at what makes sense in context and read that. This is completely normal and it's why you should be generous with brackets in your own writing: your reader doesn't have your context.

2.7 The parse drill

Three questions, in order, on any expression you're stuck on.

1. **Is this a thing or a claim?** Find the relation symbol, or establish there isn't one.
2. **What's the outermost operator?** Sum, integral, fraction, function, quantifier.
3. **What varies, and over what range?** Everything else is fixed for the purposes of this line.

Answer those three and the expression is parsed, whether or not you yet know what it means.

PRACTISE THIS

Open your lecture notes at the densest displayed formula you can find. Run the three parse questions on it and write the answers in the margin. Then say the whole thing aloud in one sentence. Do this on one formula a day for a fortnight and Mode 4 stops happening to you.

3

Quantifiers, and Why Order Decides Everything

Two symbols, $\forall$ and $\exists$, and more marks are lost to them than to any other notation in a mathematics degree.

Not because they're hard. Because their order matters and nothing in the way they're printed tells you that.

3.1 The two symbols

The quantifiers

$\forall$	**for all, for every, for each** The statement that follows holds for every element of the stated set. No exceptions. *Not: "for any" in the loose English sense. English "any" can mean every ("any fool knows that") or some ("do you have any?"). $\forall$ only ever means every.*
$\exists$	**there exists, there is** At least one element makes the statement true. It says nothing about how many, and nothing about which. *Not: exactly one, and not "we can construct one". Existence proofs often produce no example at all.*
$\exists!$	**there exists a unique** Exactly one. Both existence and uniqueness are being claimed, and a full proof has to establish both.
$\nexists$	**there does not exist** No element makes the statement true. Equivalent to $\forall$ applied to the negation.

3.2 Order changes the meaning, always

Here it is with the smallest possible example, over the integers.

$\forall x\, \exists y : y > x$	For every x there's some y bigger. **True.** y is allowed to depend on x: take $y = x + 1$
$\exists y\, \forall x : y > x$	There's one y bigger than every x. **False.** y is fixed first, and then x can be $y + 1$

The rule underneath is worth memorising, because it generalises to every quantifier pair you will ever meet:

> *A later variable may depend on an earlier one. An earlier variable may not depend on a later one. Reading a formula left to right is reading it in order of who gets to know about whom.*

So in $\forall x\, \exists y$, the y is chosen after x is known, and may be built from it. In $\exists y\, \forall x$, the y is pinned down before x is even mentioned, so it must work for all of them at once. That's a much stronger claim, and it's why the two are so often the difference between a theorem and a falsehood.

3.3 Where you meet it for real

Continuity. This is the definition that separates students who read quantifiers from students who don't.

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$$

versus *uniform* continuity:

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall a\, \forall x : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$$

The only difference is where $\forall a$ sits. In the first, δ is chosen after a is known, so it's allowed to depend on the point. In the second, δ comes before a, so one δ has to work everywhere at once.

That is the entire content of the distinction between continuity and uniform continuity, and it's a matter of quantifier order. Read the order and the concept is obvious. Miss it and you'll spend a term thinking the two definitions look the same.

HOW IT IS ACTUALLY USED

This is the single highest-value thing in the book. If you take away one habit, make it this: whenever you meet a definition with more than one quantifier, write out which variables are allowed to depend on which. Three lines in the margin. It converts most analysis definitions from opaque to obvious.

3.4 Negation flips them

To negate a quantified statement, swap each quantifier and negate the inside.

Statement	Its negation
$\forall x : P(x)$	$\exists x : \neg P(x)$
$\exists x : P(x)$	$\forall x : \neg P(x)$
$\forall x\, \exists y : P(x, y)$	$\exists x\, \forall y : \neg P(x, y)$
$P \Rightarrow Q$	$P \wedge \neg Q$

That last row surprises people. The negation of "if P then Q" is not "if P then not Q". It's "P holds and Q doesn't", which is exactly what a counterexample is.

Applied to continuity, negating the definition gives you the recipe for proving a function *isn't* continuous: there's an ε such that for every δ, there's an x within δ of a whose image is further than ε away. Every step of that proof is dictated by the negation rule.

3.5 Implicit quantifiers, which are everywhere

Most quantifiers in real mathematical writing are never written down.

As written	As meant
"$(x+1)^2 = x^2+2x+1$"	For all real x
"$\sin^2\theta + \cos^2\theta = 1$"	For all θ
"Let f be continuous. Then f is bounded on $[a,b]$."	For all such f, and all $a < b$
"$x^2 = 4$"	There exists such an x, or: solve for x. Context decides

The last row is the ambiguous one. An identity and an equation to solve look identical on the page, and only the surrounding sentence tells you which is meant. When you write, say which: "for all x" or "find all x such that".

3.6 The colon, the comma, and the dot

Between the quantifier and the statement, authors write various things. They all mean "such that" or nothing at all.

$\forall x : P(x)$	Colon. Common in logic and computer science
$\forall x, P(x)$	Comma. Common in analysis
$\forall x \,.\, P(x)$	Dot. Common in formal logic and type theory
$\forall x\ P(x)$	Nothing at all. Very common; rely on spacing
$\forall x \in A$	The set membership is part of the quantifier: "for all x in A"
$\exists x$ s.t. $P(x)$	Written out. Older texts

None of these carries meaning. Don't spend time on the difference.

3.7 Bounded quantifiers

$\forall\varepsilon > 0$ is shorthand for $\forall\varepsilon\,(\varepsilon > 0 \Rightarrow \ldots)$, and $\exists\delta > 0$ is shorthand for $\exists\delta\,(\delta > 0 \wedge \ldots)$.

Note that the two expand differently: the universal one becomes an implication, the existential one becomes a conjunction. That asymmetry falls straight out of the negation rules above, and it explains why an empty range makes a $\forall$ statement vacuously true and an $\exists$ statement automatically false.

"Every element of the empty set is purple" is true. Annoying, and true.

PRACTISE THIS

Find the definition of convergence, continuity, or a limit in your current notes. Write out, in the margin, which variable is chosen before which, and which is allowed to depend on which. Then negate the whole definition using the flip rule. Those two exercises are most of what an analysis course asks you to do.

4

What Binds What

In $\sum_{k=1}^{n} a_k$, the answer depends on n and not on k.

That's obvious once you see it and invisible until you do, and the idea behind it, *binding*, runs through every part of mathematics.

4.1 Free and bound

A variable is **bound** if some symbol has captured it and given it a range to run over. It's **free** otherwise.

Expression	Bound	Free
$\sum_{k=1}^{n} a_k$	k	n, and the sequence a
$\int_0^1 f(x,t)\,dx$	x	t, and f
$\forall x : x > y$	x	y
$\{n \in \mathbb{N} : n > m\}$	n	m
$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$	h	a, f

The test is simple and mechanical:

> *A bound variable can be renamed throughout without changing anything. A free variable cannot. If swapping every k for a j leaves the meaning identical, k was bound.*

So $\sum_{k=1}^{n} a_k$ and $\sum_{j=1}^{n} a_j$ are the same object. But $\sum_{k=1}^{n} a_k$ and $\sum_{k=1}^{m} a_k$ are not, because n was free.

4.2 Why this is worth knowing

Three practical payoffs, all of them things students get wrong.

You can't use a bound variable outside its scope. Writing "$\sum_{k=1}^{n} a_k$, and therefore $k < n$" is meaningless. There is no k outside the sum. This is a common and heavily penalised error in first proofs.

Reusing a letter shadows it. If you're inside a proof where n means one thing and you open a sum over n, the inner n hides the outer one and your reader has to guess which is which. Pick a fresh letter. This costs nothing and prevents a whole class of confusion.

Dependence tells you what to differentiate. In $\int_0^1 f(x,t)\,dx$, the result is a function of t alone, so $\frac{d}{dt}$ makes sense and $\frac{d}{dx}$ does not. Knowing which variables survive the operation is knowing what kind of object you have.

4.3 The binders

Every symbol below captures a variable. When you see one, ask immediately which letter it has taken.

Symbols that bind a variable

$\sum_k$	**sum over** k Binds k. The result depends on the range and on whatever else appears in the summand.
$\prod_k$	**product over** k Binds k, same as the sum.
$\int \dots dx$	**integral with respect to** x Binds x. The dx is not decoration; it names the bound variable and marks where the integrand ends.
$\forall x$	**for all** x Binds x over the statement that follows.
$\exists x$	**there exists** x Binds x over the statement that follows.
$\lim_{x\to a}$	**limit as** x **tends to** a Binds x. Leaves a free.
$\{x : P(x)\}$	**the set of** x **such that** P **of** x Binds x inside the braces.
$\lambda x.\, e$	**lambda** x, e Binds x; the whole thing is a function. Common in logic, computer science, and type theory.
$\max_{x\in S}$	**max over** x **in** S Binds x. So do min, sup, inf, arg max, and arg min.

4.4 Dummy variables, and why the letter is a choice

A bound variable is often called a *dummy*. The name is fair: it's a placeholder, and the choice of letter is a courtesy to the reader rather than a mathematical fact.

Conventions worth following, because breaking them makes readers work:

i, j, k, ℓ, m, n	Integer indices and counters
x, y, z	Real variables
s, t	Time, or a parameter
u, v, w	Vectors, or substitution variables
f, g, h	Functions
A, B, C	Sets or matrices, depending on the field
ε, δ	Small positive quantities
α, β, γ	Angles, or coefficients
λ	An eigenvalue, or a Lagrange multiplier, or a parameter
θ	An angle, or a parameter to be estimated
p, q	Primes, probabilities, or propositions

Using n for a real number or x for an integer index isn't wrong. It just makes every reader pause, and a page full of small pauses is a page people put down.

HOW IT IS ACTUALLY USED

Referees and markers notice variable naming more than authors expect. Not because it's a rule, but because consistent naming is a reliable signal that the writer has the structure clear in their own head. A proof that uses k for three different things in two pages reads as a proof that was assembled rather than understood.

4.5 Substitution, and the trap in it

When you substitute into an expression, you may only replace free variables, and you must not accidentally capture one.

The classic failure:

MISREAD

Start with $\exists y : y > x$, which is true over the integers for every x.
Substitute $x = y$.
Get $\exists y : y > y$, which is false.

READ CORRECTLY

Start with $\exists y : y > x$.
Rename the bound variable first: $\exists z : z > x$.
Now substitute $x = y$: $\exists z : z > y$, which is true.

The substituted y was captured by the existing binder. Renaming the bound variable first is the standard fix, and it's why compilers, proof assistants, and careful humans all rename before substituting.

You'll meet this again in any course involving formal logic, lambda calculus, or automated proof, where it's given the name *capture-avoiding substitution*.

4.6 Reading a formula for binding

Two passes, quick.

1. **Circle every binder.** $\sum$, $\int$, $\forall$, $\exists$, lim, set braces, max.

2. **For each, name the variable it took** and how far its influence runs.

What's left uncircled is free, and the free variables are what the whole expression is a function of. That last sentence is worth rereading: it's how you work out what kind of object you're looking at.

PRACTISE THIS

Take three displayed formulas from your notes. For each, list the bound variables and the free ones, then say in words what the expression is a function of. If you can't name what it's a function of, you haven't finished parsing it.

5

Saying It Aloud

The skill that makes everything else work.

You can't discuss what you can't pronounce, you can't hold it in working memory, and you can't tell when you've misread it. Reading aloud is the error-check.

5.1 Why silent reading fails on notation

Ordinary prose is read silently because you already have the sounds. You learned the words aurally years before you read them, so the silent reading is running on a soundtrack you don't notice.

Notation has no soundtrack. Nobody taught you the sounds, so when you read $\bigcup_{i\in I} A_i$ silently you're storing a picture, and pictures degrade within a line.

The fix is to build the soundtrack deliberately. It takes a few weeks and then it's permanent.

> *If you can't say a formula aloud at normal speaking speed, you haven't read it. Recognising it isn't reading it.*

5.2 The default readings

Not the only ones, but the ones you'll hear most often. Say each of these out loud once now.

Written	**Said**
x^2	"x squared"
x^3	"x cubed"
x^n	"x to the n"
x^{n+1}	"x to the n plus one"
x_n	"x sub n", or just "x n"
a_{ij}	"a i j"
$f(x)$	"f of x"
$f'(x)$	"f prime of x"
$\dot{x}$	"x dot"
$\bar{x}$	"x bar"
$\hat{x}$	"x hat"
$\tilde{x}$	"x tilde"
$\lvert x\rvert$	"mod x", or "the absolute value of x"
$\lVert v\rVert$	"norm v", or "the norm of v"
$\frac{a}{b}$	"a over b"
$\sqrt{x}$	"root x"
$\sqrt[3]{x}$	"cube root of x"

continued on next page

Written	Said
$\int_a^b f$	"the integral from a to b of f"
$\frac{dy}{dx}$	"dy by dx", or "dy over dx"
$\frac{\partial f}{\partial x}$	"partial f by partial x", or "*deef deex*"
∇f	"grad f"
$\nabla \cdot v$	"div v"
$\nabla \times v$	"curl v"
$n!$	"n factorial"
$\binom{n}{k}$	"n choose k"
$A^\top$	"A transpose"
A^{-1}	"A inverse"
$\langle u, v\rangle$	"the inner product of u and v"
$x \in A$	"x in A", or "x is an element of A"
$A \subseteq B$	"A is contained in B", or "A is a subset of B"
$A \cup B$	"A union B"
$A \cap B$	"A intersect B"
$f : A \to B$	"f from A to B", or "f maps A to B"
$x \longmapsto x^2$	"x maps to x squared"
$a \equiv b \pmod{n}$	"a is congruent to b mod n"
$\sum_{k=1}^{n}$	"the sum from k equals one to n"
$\lim_{x\to 0}$	"the limit as x tends to zero"
$O(n^2)$	"big oh of n squared"
$\mathbb{R}$	"the reals", or "R"
$\emptyset$	"the empty set"
∞	"infinity"

Appendix C has this list plus a hundred more, formatted as a card to keep in a notebook.

5.3 Three habits of fluent readers

They chunk. Nobody says "the sum from k equals one to n of one over k" twice. The second time it's "the harmonic sum". Building names for recurring chunks is most of what fluency is.

They use placeholders. "This thing here", "that whole factor", "the bracket". You do not need a name for every sub-expression to talk about the structure, and reaching for the placeholder is faster than stalling.

They read structure before content. "It's an integral of a product, and the second factor is exponential." That's a complete useful reading, delivered before any of the details.

HOW IT IS ACTUALLY USED

Listen to a lecturer read a long formula. They'll typically say the shape first ("so this is a sum of squares"), then descend into the terms. They almost never read left to right, symbol by symbol, which is exactly what students do when asked to read one out. Copying the professional habit is free and immediate.

5.4 The regional variations you'll hear

Nobody agrees, and none of it matters. Say what your department says.

$\lvert x \rvert$	"mod x" (UK, India) vs "absolute value of x" (US)
$\emptyset$	"empty set" vs "null set" vs "fee"
$\mathbb{Z}$	"zed" (UK, India) vs "zee" (US)
$\frac{dy}{dx}$	"dy by dx" (UK) vs "dy dx" (US)
ln	"lon" vs "L N" vs "natural log"
sin	"sine", and sinh is "shine" or "sinch"
tanh	"tanch", "than", or "tan aitch". All three are said
π	"pie" in English, "pee" in Greek
χ	"kye" (UK, US) vs "khee" (Greek)
$\aleph$	"aleph", with the stress usually on the first syllable

Do not spend energy on these. The only cost of picking the wrong one is a half-second of clarification.

5.5 The drill

Five minutes, three times a week, for a month.

1. Open your notes at a page with displayed formulas.
2. Read three of them aloud, at speaking speed, without pausing to work anything out.
3. Where you stall, mark the symbol and look it up. Two minutes, maximum.
4. Read the same three again the next session, from cold.

Step 4 is the one that makes it work, and it's the one people skip. Rereading from cold is what moves the sounds from recognition into recall.

5.6 Reading to another person

The strongest version of the drill: explain a formula out loud to somebody who doesn't know it.

You'll discover within thirty seconds which parts you can say and which parts you've been sliding past. There's no faster diagnostic, and it works even when the other person is a housemate doing an unrelated degree, because the failures are in your own mouth rather than in their understanding.

PRACTISE THIS

Pick the longest displayed formula in your current lecture notes. Read it aloud to somebody who isn't taking the course. Anywhere you say "and then this bit" and move on, mark it. Those marks are your reading list for Part II.

Part II

The Symbols

6

Logic and Proof

The symbols that hold an argument together.

These appear in every subject, and they're the ones most often read approximately, because each has a close English word that means something slightly different.

6.1 Connectives

Connectives

$\neg$	**not** Negation. Also written $\sim$ in older texts and in some logic courses.
$\wedge$	**and** Both sides hold. Also written & informally, and $\cdot$ in Boolean algebra.
$\vee$	**or** At least one side holds. This is *inclusive* or: both is allowed. *Not: the exclusive "or" of English ("tea or coffee"). Exclusive or is* $\oplus$ *or* $\veebar$.
$\Rightarrow$	**implies, if…then** If the left side is true, the right side is true. Written $\rightarrow$ in many logic courses and $\supset$ in older ones. *Not: "causes". It carries no mechanism and no time order.* $\Rightarrow$ *says nothing whatever about the case where the left side is false.*
$\Leftrightarrow$	**if and only if, iff** Both directions. Equivalently, the two sides have the same truth value. Often written $\equiv$ in logic, and "iff" in prose.
$\Leftarrow$	**is implied by, if** The reverse implication. Rarely written alone; usually one half of a two-part proof.

6.2 Turnstiles, and the distinction they draw

These two look similar, appear in the same courses, and mean genuinely different things. Getting them the wrong way round is a standard exam error in a first logic course.

Provability and truth

$\vdash$	**proves, entails, turnstile** Syntactic. There is a derivation of the right side from the left, using the rules of the system. A statement about what you can *write down*.

$\models$ **models, satisfies, double turnstile**

Semantic. The right side is true in every model where the left side is true. A statement about what is *the case*.

Not: the same as $\vdash$. Soundness and completeness theorems are precisely the claims that the two coincide, and they are theorems, not definitions.

6.3 Definition versus claim

The most consequential distinction in the chapter, and the one most often lost.

Introducing versus asserting

$:=$ **is defined to be**

Definition. The left side is being introduced as a name for the right side. Also written $\equiv$, $\triangleq$, $\doteq$, or $\stackrel{\text{def}}{=}$.

Not: an equation you could be asked to prove. There is nothing to prove; a name is being assigned.

$=$ **equals**

A claim, or a condition. Depending on context: an identity that holds always, or an equation to be solved.

$\equiv$ **is identically equal to, is congruent to, is defined as**

Three distinct jobs depending on field. In analysis: holds for all values, not just some. In number theory: congruence, always with a modulus. In logic and physics: definition.

Not: one single meaning. Check the context every time. See Chapter 16.

6.4 Proof punctuation

Marking the argument

$\therefore$ **therefore**

Conclusion follows. Common in handwriting and school work, rare in printed mathematics.

$\because$ **because**

Reason follows. Rarer still, and worth avoiding in written work.

$\blacksquare$ **end of proof, QED**

Marks the end. Also $\square$, $\square$, “QED”, or “//”.

$\square$ **end of proof, or a modal operator**

End of proof in most texts. In modal logic it means “necessarily”, and $\Diamond$ means “possibly”.

6.5 Quantifiers, in brief

Covered fully in Chapter 3. Repeated here so the chapter stands alone as a reference.

Quantifiers

$\forall$	**for all** Every element, no exceptions. *Not: English "any", which is ambiguous.*
$\exists$	**there exists** At least one. *Not: exactly one, which is* $\exists!$.
$\exists!$	**there exists a unique** Exactly one. Two things to prove.

6.6 Sets of truth values and Boolean notation

Truth values

$\top$	**top, true, verum** The always-true proposition. In order theory, the greatest element.
$\bot$	**bottom, false, falsum, contradiction** The always-false proposition. In order theory, the least element. In a proof by contradiction, what you're aiming to derive.
$\oplus$	**exclusive or, XOR** Exactly one side holds. In other contexts entirely, direct sum. See Chapter 12.

6.7 The three implications students conflate

Given $P \Rightarrow Q$, three related statements exist and only one of them follows.

Name	**Statement**	**Does it follow?**
Converse	$Q \Rightarrow P$	No
Inverse	$\neg P \Rightarrow \neg Q$	No
Contrapositive	$\neg Q \Rightarrow \neg P$	**Yes**, always

MISREAD

All squares are rectangles. This shape is not a square, so it is not a rectangle.

READ CORRECTLY

All squares are rectangles. This shape is not a rectangle, so it is not a square.

The first uses the inverse and is wrong. The second uses the contrapositive and is valid. Proof by contrapositive is a standard technique precisely because the contrapositive is always equivalent, and it is often much easier to prove.

HOW IT IS ACTUALLY USED

In practice, working mathematicians write $\Rightarrow$ far less than students expect. Printed proofs are mostly English sentences: "Since f is continuous, …", "It follows that …", "Hence …". The arrows live on blackboards and in your own notes. A proof written entirely in symbols is usually harder to check, not easier, and most journals will ask you to convert it to prose.

6.8 Vacuous truth, which is real and annoying

If P is false, then $P \Rightarrow Q$ is true, whatever Q says.

"Every element of the empty set is purple" is a true statement. So is "every element of the empty set is not purple".

This isn't a quirk of notation, it's what makes the definition of implication consistent, and it shows up in real proofs: a claim about all elements of a set you later prove is empty is automatically true, and that's sometimes exactly the argument.

PRACTISE THIS

Take one theorem from your current notes and write its converse, inverse, and contrapositive. Decide which are true. Then find one $\Rightarrow$ in your own written work and check you haven't quietly used its converse.

7

Sets and Relations

Set notation is the vocabulary the rest of mathematics is written in, and it contains the single most common genuine ambiguity in the subject.

That ambiguity is $\subset$, and it gets a section of its own.

7.1 Membership and containment

Membership

$\in$	**is an element of, is in, belongs to** $x \in A$ says x is one of the things A contains. *Not:* $\subseteq$. *Membership and containment are different relations, and mixing them is the commonest set-theory error.* $\{1\} \subseteq \{1, 2\}$ *is true;* $\{1\} \in \{1, 2\}$ *is false.*
$\notin$	**is not an element of**
$\ni$	**contains as an element** $A \ni x$ is $x \in A$ written backwards. Uncommon, and used to keep a sentence flowing.

Containment

$\subseteq$	**is a subset of, is contained in** Every element of the left set is in the right one. Equality is allowed.
$\subsetneq$	**is a proper subset of** Contained, and not equal. Unambiguous, and worth using in your own writing for exactly that reason.
$\subset$	**is a subset of, or is a proper subset of** **This symbol is ambiguous.** Most modern authors use it for $\subseteq$. Some, especially in older and continental European texts, use it for $\subsetneq$. *Not: something you can rely on without checking. See the next section.*
$\supseteq$	**contains, is a superset of** The mirror image of $\subseteq$.

7.2 The $\subset$ problem, and how to handle it

Two schools, both established, both confident.

School	What $A \subset B$ means to them
The common modern one	$A \subseteq B$. Equality allowed. Analogous to $\leq$
The strict one	$A \subsetneq B$. Equality not allowed. Analogous to $<$

You cannot tell which one an author means from the symbol. You have to look for one of these:

1. An explicit statement in the notation section or first chapter. Most careful books say.
2. A place where the author writes $A \subset A$. If they do, they mean the inclusive reading.
3. The presence of $\subsetneq$ elsewhere. If they have a separate symbol for proper, then $\subset$ is inclusive.

> *In your own writing, use $\subseteq$ and $\subsetneq$, and never $\subset$. It costs one keystroke and removes the ambiguity permanently.*

7.3 Building sets

Constructors

$\{ a, b, c \}$ **the set containing a, b and c**

Roster notation. Order doesn't matter and repeats don't count: $\{1, 2, 2\} = \{1, 2\}$.

Not: a list or a tuple, where order and repetition do matter. Those are written with round brackets.

$\{ x \in A : P(x) \}$ **the set of x in A such that P of x**

Set-builder notation. The colon is sometimes a vertical bar: $\{ x \mid P(x) \}$. Both mean "such that".

$\emptyset$ **the empty set**

The set with no elements. Also written $\varnothing$, or $\{\}$.

Not: the number zero, and not the Greek letter ϕ, which it is frequently and wrongly called.

$\mathcal{P}(A)$ **the power set of A**

The set of all subsets of A, including $\emptyset$ and A itself. Also written 2^A, which is a good notation: if A has n elements, $\mathcal{P}(A)$ has 2^n.

$|A|$ **the cardinality of A, the size of A**

How many elements. Also written $\#A$ or $n(A)$.

Not: absolute value, though it is the same glyph. Context decides. See Chapter 16.

7.4 Operations

Combining sets

$\cup$ **union**

Everything in either set. $\bigcup_{i \in I} A_i$ for an indexed family.

$\cap$ **intersection**

Everything in both. $\bigcap_{i\in I} A_i$ for an indexed family.

$\setminus$ **set minus, difference**

$A \setminus B$ is everything in A that isn't in B. Also written $A - B$.

Not: requiring $B \subseteq A$. The elements of B outside A are simply ignored.

A^c **the complement of A**

Everything not in A, relative to an understood universal set. Also written $\bar{A}$, A', or $\complement A$.

Not: well defined without a stated universe. "Everything not in A" needs to know what everything is.

$\times$ **Cartesian product**

$A \times B$ is the set of ordered pairs (a, b). A^n is n copies. This is where $\mathbb{R}^3$ comes from.

Not: multiplication, though it is the same glyph.

$\sqcup$ **disjoint union**

Union, with the claim that the sets do not overlap, or with the elements tagged to keep them distinct.

7.5 Comparing things

These five are read as "roughly the same" by students and mean five different strengths of sameness.

Degrees of sameness

$=$ **equals**

The same object.

$\cong$ **is isomorphic to, is congruent to**

Structurally the same. In algebra, an isomorphism exists. In geometry, same shape and size.

Not: the same object. Two isomorphic groups can be built from entirely different things.

$\simeq$ **is isomorphic to, is homotopy equivalent to**

Weaker than $\cong$ in topology. In many other areas used interchangeably with $\cong$, which is unhelpful but common.

$\sim$ **is similar to, is related to, is distributed as**

The general-purpose relation symbol. Means an equivalence relation, or similarity of triangles, or asymptotic equivalence, or a probability distribution, depending entirely on field.

Not: one thing. This is the most overloaded symbol in mathematics.

$\approx$ **is approximately equal to**

Numerically close. Not a precise statement unless the author says how close.

Not: $\sim$, which in analysis means the ratio tends to 1, a genuinely different claim.

$\propto$ **is proportional to**

Equal up to an unspecified constant factor.

7.6 Relations and orders

Order relations

$\leq$ **is less than or equal to**

$\preceq$ **precedes or equals**

A general partial order. Used when the order isn't the usual numeric one.

$\sqsubseteq$ **is a sub-thing of**

Another general order symbol, common in computer science for substructures and in domain theory.

$\perp$ **is perpendicular to, is orthogonal to, is independent of**

Geometry: perpendicular. Linear algebra: orthogonal. Statistics: $\perp\!\!\!\perp$ for independence. Order theory: the bottom element.

HOW IT IS ACTUALLY USED

Set notation is where the gap between how it's taught and how it's used is widest. Textbooks introduce $\cup$, $\cap$ and complements with Venn diagrams and then almost never use them again. What working mathematicians actually use constantly is set-builder notation and indexed unions: $\{\, x \in X : f(x) > 0 \,\}$ and $\bigcup_n A_n$. Get fluent with those two and you have most of what you'll need.

7.7 The number sets, previewed

Covered fully in Chapter 8. Listed here because they appear in every set-builder expression you'll meet.

$\mathbb{N}$	The natural numbers. Whether 0 is included depends on the author
$\mathbb{Z}$	The integers
$\mathbb{Q}$	The rationals
$\mathbb{R}$	The reals
$\mathbb{C}$	The complex numbers

PRACTISE THIS

Open any textbook you're using and find out, in the first ten pages, whether the author's $\subset$ allows equality and whether their $\mathbb{N}$ contains zero. Write both answers on the inside cover. You will need them all term, and nobody will tell you again.

8

Numbers and Number Systems

Blackboard bold letters, the brackets that modify a number, and the one question you must answer about $\mathbb{N}$ before you use anybody's notes.

8.1 The number sets

Blackboard bold

$\mathbb{N}$	**the naturals, N** The counting numbers. **Whether** 0 **is included is not settled.** French and much of computer science and set theory include it; much of English-language analysis and number theory does not. *Not: a settled convention. Look for* $\mathbb{N}_0$, $\mathbb{Z}^+$, *or* $\mathbb{Z}_{\geq 0}$, *all of which are unambiguous.*
$\mathbb{Z}$	**the integers, Z** Whole numbers, positive, negative and zero. From the German *Zahlen*.
$\mathbb{Q}$	**the rationals, Q** Ratios of integers. From *quotient*.
$\mathbb{R}$	**the reals, R**
$\mathbb{C}$	**the complex numbers, C** Numbers $a + bi$ with $a, b \in \mathbb{R}$.
$\mathbb{F}_q$	**the finite field with q elements** Also written GF(q), for Galois field. q must be a prime power.
$\mathbb{H}$	**the quaternions** Named for Hamilton. Multiplication is not commutative.

Restricting a number set

$\mathbb{R}^+$	**the positive reals** Usually strictly positive. Occasionally non-negative. Ambiguous enough that careful writers use $\mathbb{R}_{>0}$ or $\mathbb{R}_{\geq 0}$.
$\mathbb{Z}^\times$	**the units, the invertible elements** In a ring, the elements with multiplicative inverses. For $\mathbb{Z}$ that's just ± 1. *Not: the non-zero elements, though for a field the two coincide.*

$\mathbb{R}^n$ **R to the n**

Ordered n-tuples of reals. This is the Cartesian product of Chapter 7, so $\mathbb{R}^3$ is triples.

8.2 The chain, and what each step adds

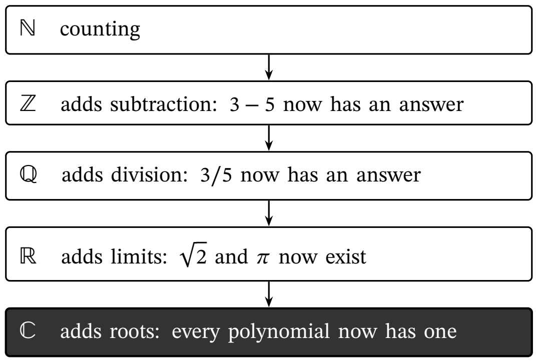

Figure 8.1: The standard chain of number systems, with what each extension was invented to solve. Each set contains all the previous ones.

That chain is worth carrying because it tells you which set a problem needs. A question about solving $x^2 + 1 = 0$ is a question about $\mathbb{C}$; a question about whether a sequence converges is a question about $\mathbb{R}$.

8.3 Brackets that modify a number

Size and rounding

$|x|$ **mod x, the absolute value of x**

Distance from zero. For a complex number, distance from the origin: $|a + bi| = \sqrt{a^2 + b^2}$.

$\lfloor x \rfloor$ **the floor of x**

The greatest integer not exceeding x. $\lfloor 2.7 \rfloor = 2$, and $\lfloor -2.7 \rfloor = -3$.

Not: rounding, and not truncation. For negative numbers all three differ.

$\lceil x \rceil$ **the ceiling of x**

The least integer not below x. $\lceil 2.1 \rceil = 3$.

$\|v\|$ **the norm of v**

Length of a vector. Double bars distinguish it from absolute value. Chapter 12.

8.4 Signs and comparison

Signs

$\pm$ **plus or minus**

Both cases at once, as in the quadratic formula. In experimental work, a measurement with its uncertainty: 9.81 ± 0.02.

$\mp$ **minus or plus**

The opposite choice to a $\pm$ elsewhere in the same expression. The pairing is the whole point.

Comparison

$\ll$ **is much less than**

Informal in analysis, and made precise in asymptotics. In number theory it is Vinogradov's notation and means exactly $O(\cdot)$.

Not: a defined quantity of "much". Unless it is Vinogradov's, it is a hint to the reader.

$\gg$ **is much greater than**

$\neq$ **is not equal to**

$\asymp$ **is asymptotically equivalent to, is of the same order**

Bounded above and below by constant multiples.

8.5 Number theory notation

Divisibility and congruence

$a \mid b$ a **divides** b

There is an integer k with $b = ak$. **This is a statement**, true or false.

Not: a/b, which is a number. The vertical bar here is a relation, not division, and it points the opposite way from the intuition: $2 \mid 6$, *not* $6 \mid 2$.

$a \nmid b$ a **does not divide** b

$a \equiv b \pmod{n}$ a **is congruent to** b **mod** n

n divides $a - b$. The modulus is part of the statement and must always be written.

$\gcd(a, b)$ **the gcd of** a **and** b

Greatest common divisor. Sometimes written (a, b), which collides badly with ordered pairs and open intervals.

$n!$ n **factorial**

$1 \times 2 \times \cdots \times n$, with $0! = 1$ by convention.

$\binom{n}{k}$ n **choose** k

The number of k-element subsets of an n-element set. Also written ${}^nC_k$ or $C(n, k)$ in school texts.

Not: a fraction. The absence of a bar is the whole difference, which is why the school notations survive.

8.6 Intervals, and the bracket that means two things

Written	Means	Endpoints
$[a, b]$	closed	both included
(a, b)	open	neither included
$[a, b)$	half-open	a yes, b no
$]a, b[$	open	continental European style for (a, b)

(a, b) is genuinely ambiguous: it is the open interval, and it is the ordered pair, and in some texts it is the gcd. Nothing in the symbol tells you which. Context always does, but you have to look.

HOW IT IS ACTUALLY USED

The $\mathbb{N}$ question comes up more than you'd think, and it is not pedantry. An induction proof with base case $n = 0$ and one with base case $n = 1$ are different proofs, and a definition of a sequence as $(a_n)_{n\in\mathbb{N}}$ has a different first term depending on the author. Find out which convention your lecturer uses in week one.

8.7 Infinity, which is not a number

∞ is not an element of $\mathbb{R}$. Writing ∞ in a limit, in an interval endpoint, or in an integral bound is shorthand for a statement about unbounded behaviour, not a value being plugged in.

So $[0, \infty)$ is half-open with good reason: there is nothing at the right end to include. And $\lim_{x\to\infty} f(x) = L$ is a claim about what happens as x grows without bound, in which no arithmetic involving ∞ takes place.

The extended reals $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ do treat it as an element, in measure theory and elsewhere, with carefully restricted arithmetic. $\infty - \infty$ remains undefined even there.

PRACTISE THIS

Check three sources you use, your lecture notes and two textbooks, for whether $0 \in \mathbb{N}$. Note the answers. Then find one place in your own written work where you used $\subset$, (a, b), or $\mathbb{N}$ without saying which reading you meant, and fix it.

9

Functions and Mappings

Two arrows that look alike and do different jobs, plus the vocabulary for saying what a function does to a set.

9.1 The two arrows

This is the distinction most first-year students have never had explained, and it takes one sentence.

Arrows

$\to$ **to, from…to**

Used between *sets*. $f : A \to B$ says f takes things from A and produces things in B.

$\longmapsto$ **maps to**

Used between *elements*. $x \longmapsto x^2$ says this particular input goes to that particular output.

Not: $\to$. The two are not interchangeable: $f : \mathbb{R} \to \mathbb{R}$, $x \longmapsto x^2$ is the full, correct way to define a function, and it uses both.

So the complete definition of the squaring function reads:

$$f : \mathbb{R} \to \mathbb{R}, \qquad x \longmapsto x^2$$

Said aloud: "f from the reals to the reals, x maps to x squared." The first part says what kind of function it is; the second says what it does.

> *$\to$ goes between sets. $\longmapsto$ goes between elements. If you can remember only one thing about function notation, remember that.*

9.2 The three sets of a function

Domain	Where the inputs come from. The A in $f : A \to B$
Codomain	Where the outputs are declared to live. The B. Chosen by whoever defined the function
Image or range	Where the outputs *actually* land. A subset of the codomain, possibly a proper one

For $f : \mathbb{R} \to \mathbb{R}$, $x \longmapsto x^2$, the codomain is all of $\mathbb{R}$ and the image is only $[0, \infty)$. That gap is exactly what "not surjective" means, and the distinction is why the two words exist.

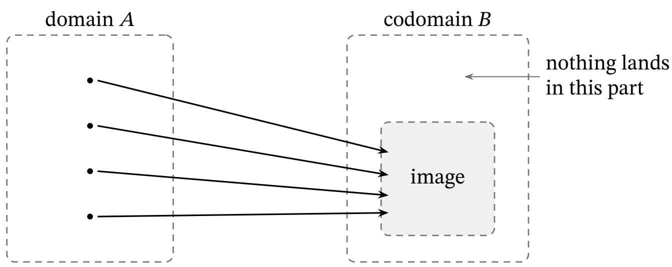

Figure 9.1: Domain, codomain and image. Surjective means the image fills the codomain; injective means no two arrows land on the same point.

9.3 Properties of functions

The three properties

$A \hookrightarrow B$ **injective, one-to-one**

Different inputs give different outputs. Equivalently, $f(x) = f(y)$ forces $x = y$. The hooked arrow is the standard notation.

Not: "one-to-one correspondence", which is the older name for bijective. *This clash is why the Latin words are preferred.*

$A \twoheadrightarrow B$ **surjective, onto**

Every element of the codomain is hit by something. Image equals codomain. The two-headed arrow is the standard notation.

$A \xrightarrow{\sim} B$ **bijective, a bijection**

Both. Exactly then does an inverse function exist.

9.4 Operating on functions

Combining and inverting

$g \circ f$ g **composed with** f**,** g **after** f**,** g **of** f

Do f first, then g: $(g \circ f)(x) = g(f(x))$.

Not: f first reading from the left. Composition is written right to left, which is the opposite of the reading order and catches everyone at least once.

f^{-1} f **inverse**

The inverse function, when f is bijective. $f^{-1}(f(x)) = x$.

Not: $1/f$. Also, for a general f, $f^{-1}(S)$ means the preimage of a set, which is defined even when no inverse function exists.

$f|_S$ f **restricted to** S

The same function, with its domain cut down to S. Useful for making a function injective, or continuous, on a smaller set.

$f(S)$ **the image of** S **under** f

The set of all $f(x)$ for $x \in S$. A set, not a value.

$f^{-1}(S)$ **the preimage of** S

Everything that maps into S. Always defined, for any f, whether or not f is invertible.

Not: evidence that f^{-1} exists as a function. This is the single most common misreading in a first topology course.

id_A **the identity on** A

$x \longmapsto x$. Sometimes written 1_A or $\mathbb{1}$.

9.5 Structure-specific vocabulary

Different subjects name the image and the "sent to zero" set differently. The concepts are the same.

Field	Image	Sent to zero/identity
Linear algebra	$\mathrm{im}\, T$, column space	$\ker T$, null space
Group theory	$\mathrm{im}\, \varphi$	$\ker \varphi$
Analysis	range, $f(A)$	the zero set, $f^{-1}(0)$

Named subsets

$\ker f$ **the kernel of** f

Everything sent to zero, or to the identity. In linear algebra it's a subspace; in group theory a normal subgroup.

$\mathrm{im}\, f$ **the image of** f

Everything actually produced.

$\mathrm{supp}\, f$ **the support of** f

Where f is not zero, or the closure of that set. Common in analysis and signal processing.

9.6 Arrows with decorations

Specialised arrows

$\hookrightarrow$ **injects into, embeds in**

An injective map, often an inclusion of a subobject.

$\twoheadrightarrow$ **surjects onto**

A surjective map, often a quotient.

$\xrightarrow{\sim}$ **is isomorphic to, via**

An isomorphism. The label above the arrow usually names the map.

$\rightsquigarrow$ **leads to, converges to, reduces to**

No fixed meaning. Used for rewriting, for weak convergence, and for informal implication. Always check the surrounding text.

9.7 Multi-argument functions and currying

$f : A \times B \to C$ takes a pair. You'll also see $f : A \to (B \to C)$, which takes one argument and returns a function.

The two carry the same information, and the second is standard in computer science, type theory, and functional programming, where it's called *currying*. In mathematics you'll meet it in the guise of $f(x)(y)$ or $f_x(y)$.

HOW IT IS ACTUALLY USED

Composition order trips up almost everyone, and the reason it's written backwards is worth knowing. It's so that $(g \circ f)(x) = g(f(x))$ matches the way function application already nests. The cost is that the diagram $A \xrightarrow{f} B \xrightarrow{g} C$ reads left to right while the formula reads right to left. Category theorists sometimes write $f; g$ or fg for the left-to-right version precisely because of this.

PRACTISE THIS

Take three functions from your notes and write each in the full two-part form: $f : A \to B$, $x \longmapsto \ldots$. Then for each, state the image and say whether it equals the codomain. If you have never done this, it takes ten minutes and it makes surjectivity stop being an abstraction.

10

Sums, Products, and Integrals

The big operators. Each one binds a variable, each one has a range, and each one has a scope problem that catches people.

10.1 Sums and products

The big operators

$\sum_{k=m}^{n} a_k$	**the sum from k equals m to n of a sub k** Add up the terms as k runs from m to n inclusive. Binds k; the result depends on m, n, and the sequence.
$\sum_{k\in S} a_k$	**the sum over k in S** Same, with the index running over a set rather than a range. Needs the set to be finite, or the sum to converge absolutely, for the answer not to depend on order.
$\prod_{k=1}^{n} a_k$	**the product from k equals one to n** Multiply instead of add. Empty product is 1, just as the empty sum is 0.
$\coprod$	**coproduct, disjoint union** In category theory the dual of a product. In set theory, disjoint union. Not a multiplication.
$\bigcup_{i\in I} A_i$	**the union over i in I** Everything in any of the A_i. $\bigcap$ likewise for intersection.

10.2 Empty ranges, and why the conventions are what they are

If the upper limit is below the lower one, the sum is empty.

Empty sum	0
Empty product	1
Empty union	∅
Empty intersection	Everything, so it is only defined relative to a stated universe

These are not arbitrary. Each is chosen so that splitting a range in two and recombining gives the right answer, which is the property every induction proof relies on.

10.3 The scope problem

Where does the summand end?

$$\sum_{k=1}^{n} k + 1$$

This is genuinely ambiguous, and both readings occur in print. It might be $\left(\sum_{k=1}^{n} k\right)+1$, or $\sum_{k=1}^{n}(k+1)$. They differ by $n-1$.

> *Bracket the summand whenever it contains a + or − at the top level. This costs two characters and removes an ambiguity your reader cannot resolve.*

The same applies to $\int$, where the dx conventionally closes the integrand, which is one of the reasons not to drop it.

10.4 Integrals

Integral signs

$\int f(x)\,dx$	**the integral of f of x dee x** Indefinite integral, or antiderivative. The answer is a family of functions, which is why the constant of integration appears.
$\int_a^b f(x)\,dx$	**the integral from a to b** Definite integral. The answer is a number. x is bound, so $\int_a^b f(x)\,dx$ and $\int_a^b f(t)\,dt$ are identical.
$\iint$, $\iiint$	**double, triple integral** Over a region in two or three dimensions.
$\oint$	**contour integral, closed line integral** Around a closed curve. Central in complex analysis and in electromagnetism.
$\int_\Omega$	**the integral over Ω** Over a named region rather than an interval. The region's dimension is implied by Ω.

10.5 The d, and what it is doing

dx is not a multiplication and it is not decoration. It does three jobs.

1. **Names the variable of integration**, which matters as soon as there is more than one letter present. $\int f(x,t)\,dx$ and $\int f(x,t)\,dt$ are different objects.
2. **Marks where the integrand ends**, resolving the scope problem above.
3. **Carries the change of variables**, which is why substitution works: $u = g(x)$ gives $du = g'(x)\,dx$, and the algebra of that substitution is exactly what the notation was designed to make automatic.

Physicists and engineers frequently write $\int dx\, f(x)$, with the dx in front. Same meaning, and it's arguably clearer for long integrands because it announces the variable before the work starts.

10.6 Derivative notations, and the four that coexist

Notation	Named for	Where you see it
$\frac{dy}{dx}$	Leibniz	Calculus, physics. Best for the chain rule and substitution
$f'(x)$	Lagrange	Analysis, when the variable is obvious
$\dot{x}$	Newton	Mechanics, and always with respect to time
$D_x f$	Euler	Differential equations, operator theory

All four mean the same thing. You will meet all four, sometimes on the same page, and the choice is about which one makes the manipulation at hand least painful.

Derivatives

$\frac{dy}{dx}$ **dee y by dee x**

Ordinary derivative. Treat as a single symbol, though the notation is deliberately suggestive of a ratio.

Not: a fraction you may split, in general. The suggestion is useful and it is not a licence.

$\frac{\partial f}{\partial x}$ **partial f by partial x, dee f dee x**

Partial derivative: differentiate with respect to x, holding the other variables fixed. The curly ∂ is the signal that other variables exist.

$f''(x)$ **f double prime**

Second derivative. Beyond three primes, switch to $f^{(4)}$.

$f^{(n)}(x)$ **the nth derivative of f**

The brackets around n are essential: f^n would mean a power or an iterate.

$\dot{x}$, $\ddot{x}$ **x dot, x double dot**

First and second derivative with respect to time. Almost exclusively mechanics.

∇f **grad f, del f**

The vector of partial derivatives. Chapter 14.

δ **delta, variation**

In the calculus of variations, a variation rather than a derivative. Also a small change, a Dirac delta, or a Kronecker delta. Heavily overloaded.

10.7 Δ, δ, d, and ∂

Four related symbols that students use interchangeably and that mean four different sizes of change.

Δx	A finite change. Measurable, subtractable: $x_2 - x_1$
δx	A small change, usually infinitesimal, often a variation in the calculus-of-variations sense
dx	An infinitesimal, or a differential form, depending how rigorous the text is being
∂x	Only ever appears in $\partial f/\partial x$. It is not a quantity and cannot be used alone

That last row is worth stating plainly, because "cancel the ∂s" is a thing students try. You cannot. ∂ has no standalone meaning.

HOW IT IS ACTUALLY USED

Physicists manipulate dy/dx as a fraction constantly, splitting it, multiplying through, and cancelling. Mathematicians wince and then do the same thing in private. The practice is justified rigorously by differential forms, and separately by the chain rule, and it almost always gives the right answer. Almost. It breaks for partial derivatives, where $\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1$ rather than 1, which is a genuinely surprising result and a good reason to know when you are being informal.

10.8 Index gymnastics

Two habits that make sums much easier to read, and to write.

Reindexing. $\sum_{k=1}^{n} a_{k+1} = \sum_{j=2}^{n+1} a_j$. The substitution $j = k + 1$ shifts both the limits and the term. Getting the limits right is the whole skill, and the check is to test the first and last terms of each.

Swapping order. $\sum_i \sum_j = \sum_j \sum_i$ for finite sums, always. For infinite sums it requires absolute convergence, and the counterexamples when it fails are famous enough that any analysis course will show you one.

PRACTISE THIS

Take a summation from your notes and reindex it so the lower limit is zero. Check by writing out the first two and last two terms of both versions. Then find one integral in the same notes and say, out loud, what the dx is doing there.

11

Limits, Convergence, and Growth

Arrows that mean approach, and the family of symbols for saying how fast something grows.

11.1 Limits

Limit notation

$\lim_{x\to a} f(x)$	**the limit as x tends to a of f of x** The value $f(x)$ approaches as x gets close to a. Says nothing about $f(a)$ itself, which may differ or may not exist. *Not: $f(a)$. The whole point of a limit is that it is defined without evaluating at the point.*
$\lim_{x\to a^+}$	**the limit from the right, from above** Only $x > a$ is considered. Also written $x \downarrow a$ or $x \searrow a$.
$\lim_{x\to a^-}$	**the limit from the left, from below** Also written $x \uparrow a$ or $x \nearrow a$.
$\limsup_n$	**lim sup, the limit superior** The largest value the sequence approaches infinitely often. Always exists in the extended reals, which is why it is useful when lim may not exist.
$\liminf_n$	**lim inf, the limit inferior** The smallest such value. lim exists exactly when $\limsup = \liminf$.
$\sup S$	**the supremum, the sup, the least upper bound** The smallest number that is $\geq$ everything in S. Need not be an element of S: $\sup(0,1) = 1$. *Not:* max. *A maximum must be attained; a supremum need not be. This distinction is the whole reason both words exist.*
$\inf S$	**the infimum, the greatest lower bound** Likewise, against min.

11.2 Convergence arrows

Approach

$\to$	**tends to, approaches, converges to** $a_n \to L$ says the sequence converges to L. The same glyph as the function arrow of Chapter 9, doing an unrelated job.
$\longrightarrow$	**tends to** Identical meaning; the length is typographic, used when a label sits above the arrow.
$\xrightarrow{n\to\infty}$	**tends to, as n tends to infinity** The condition above the arrow says in what sense.
$\nearrow$	**increases to** Monotonically increasing convergence. $\searrow$ for decreasing. Common in measure theory.
$\rightrightarrows$	**converges uniformly to** Uniform convergence, in some texts. Others write $\xrightarrow{\text{unif}}$ or say it in words.
$\rightsquigarrow$	**converges weakly to, converges in distribution** Weak convergence. Also written $\Rightarrow$ in probability, which collides badly with implication. Check the context.

11.3 The convergence modes, and their notations

In analysis and probability, a sequence can converge in several inequivalent senses. The notation is how the author tells you which.

$\to$	Pointwise, or plain convergence
$\xrightarrow{\text{unif}}$	Uniformly: one rate works everywhere
$\xrightarrow{L^p}$	In p-th mean
$\xrightarrow{p}$	In probability
$\xrightarrow{d}$	In distribution
$\xrightarrow{\text{a.s.}}$	Almost surely, with probability one

The label above the arrow is not decoration. It is the entire content of the claim, and a proof that works for one mode routinely fails for another.

> *When you see an arrow with something written above it, read the label first. It tells you which theorem you are allowed to use next.*

11.4 Asymptotic and growth notation

The most misused family in the book, largely because computer science and mathematics use the same symbols with different strictness.

Growth rates

$O(g(n))$	**big oh of g of n** Grows no faster than g, up to a constant factor, for large n. An **upper** bound. *Not: a tight bound. n is $O(n^2)$, and that statement is true and unhelpful.*

$o(g(n))$ **little oh of g of n**

Grows *strictly* slower: the ratio tends to zero. Strictly stronger than O.

$\Omega(g(n))$ **big omega**

A **lower** bound: grows at least as fast as g.

Not: the same Ω as in probability, where it is the sample space, or in physics, where it is a solid angle or a region.

$\Theta(g(n))$ **big theta**

Both bounds: grows exactly like g, up to constants. This is usually what people mean when they say "$O(n \log n)$" out loud.

$\sim$ **is asymptotically equal to**

The ratio tends to 1. Stronger than Θ, which allows any constant.

Not: $\approx$. This is a precise statement with a proof obligation.

$\asymp$ **is of the same order as**

Bounded above and below by constant multiples. Effectively Θ, written by mathematicians.

$\ll$ **is much less than, or is O of**

In analytic number theory, Vinogradov's notation, exactly synonymous with O. Elsewhere, an informal hint.

11.5 The equals sign in $f(n) = O(g(n))$

This is an abuse of notation, universally used, and worth understanding rather than resenting.

$O(g(n))$ is properly a *set* of functions. The honest statement is $f \in O(g)$. Everyone writes $=$ anyway.

The consequence is that the relation is not symmetric. You may write $n = O(n^2)$. You may not write $O(n^2) = n$. Chains like $f(n) = O(n) = O(n^2)$ read left to right only, and reversing any step is invalid.

MISREAD

Since $f(n) = O(n)$ and $g(n) = O(n)$, we have $f(n) = g(n)$.

READ CORRECTLY

Since $f(n) = O(n)$ and $g(n) = O(n)$, we have $f(n) + g(n) = O(n)$.

The first treats $=$ as transitive and equality-like. It isn't.

11.6 Where the analysts and the computer scientists differ

	In analysis	**In computer science**
Limit taken	Usually $x \to 0$	Usually $n \to \infty$
O used for	Error terms in expansions	Running time
Ω	Occasionally Hardy's stronger version	Lower bound on running time
Tightness	Care taken	"O" often said when Θ is meant

That last row is the one to watch. A CS lecturer saying "quicksort is $O(n \log n)$ on average" almost always means Θ. It's not wrong, it's just weaker than what they intend, and in an exam you may be expected to state the tight bound.

HOW IT IS ACTUALLY USED

The most common way this notation is misused in student work is claiming a lower bound by proving an upper one. "The algorithm is $O(n^2)$, so it cannot be faster than n^2" is exactly backwards: O is a ceiling, and being under a ceiling says nothing about the floor. If you want to say an algorithm is genuinely slow, you need Ω.

11.7 Taylor and error terms

You will meet O and o in expansions long before you meet them in algorithms.

$$e^x = 1 + x + \frac{x^2}{2} + O(x^3) \qquad \text{as } x \to 0$$

Note the direction: here the limit is $x \to 0$, so $O(x^3)$ means *small*, not large. The same symbol, the opposite intuition, and the only thing distinguishing them is the stated limit. Which is why careful authors always write "as $x \to 0$" or "as $n \to \infty$".

PRACTISE THIS

Find an O in your notes and write down, explicitly, what the limit is and whether the bound is upper or lower. Then check whether the author meant Θ. In algorithms courses the answer is yes surprisingly often.

12

Linear Algebra and Matrices

Superscripts that aren't powers, brackets that aren't grouping, and a transpose symbol that three fields write three ways.

12.1 Decorating a matrix

Matrix superscripts

$A^{\intercal}$ — *A* **transpose**

Rows become columns. Also written A^t, A', or $A^\top$. The prime version collides with derivatives and is best avoided.

Not: a power. None of the superscripts in this group is an exponent.

A^* — *A* **star, the conjugate transpose,** *A* **adjoint**

Transpose and complex-conjugate. For real matrices it's just the transpose. In functional analysis, the adjoint operator.

$A^\dagger$ — *A* **dagger**

Conjugate transpose in physics, where A^* usually means conjugate only. In numerical linear algebra, the Moore-Penrose pseudoinverse. Two genuinely different meanings.

A^{-1} — *A* **inverse**

The multiplicative inverse, when it exists.

Not: $1/A$, which is not defined for matrices, and not the reciprocal of the entries.

$\bar{A}$ — *A* **bar,** *A* **conjugate**

Entry-wise complex conjugate, no transpose.

12.2 Named quantities

Scalars extracted from a matrix

$\det A$ — **the determinant of** *A*

Also written $|A|$, which collides with absolute value and with cardinality. Prefer det.

$\operatorname{tr} A$ — **the trace of** *A*

Sum of the diagonal entries. Equal to the sum of the eigenvalues, with multiplicity.

rank A — **the rank of A**

Dimension of the column space. Equals the dimension of the row space, which is a theorem, not a definition.

λ — **lambda, an eigenvalue**

By near-universal convention. $Av = \lambda v$ for an eigenvector $v \neq 0$.

12.3 Products, and the four of them

Students meet four things called "product" in one term. They take different inputs and give different outputs, and that's the fastest way to tell them apart.

Product	Takes	Gives
$\langle u, v\rangle$ inner	two vectors	a scalar
$u \times v$ cross	two vectors in $\mathbb{R}^3$	a vector
$u \otimes v$ outer/tensor	two vectors	a matrix, or a tensor
AB matrix	two matrices	a matrix

Products

$\langle u, v\rangle$ — **the inner product of u and v**

Also written $u \cdot v$ (the dot product), (u, v), or $\langle u \mid v\rangle$ in physics. Gives a scalar.

$u \times v$ — **u cross v**

Cross product. Only in three dimensions, and the result is perpendicular to both inputs. Also the Cartesian product of Chapter 7, and ordinary multiplication.

$\otimes$ — **tensor product, outer product, Kronecker product**

Combines two spaces into a bigger one. In quantum mechanics, combining systems.

$\oplus$ — **direct sum**

Combines two spaces side by side: $\dim(U \oplus V) = \dim U + \dim V$. Also exclusive-or in logic, and modulo-2 addition.

$\odot$ — **Hadamard product, element-wise product**

Multiply entry by entry. Common in machine learning, rare in pure mathematics.

12.4 Spaces and spans

Subspaces

span(S) — **the span of S**

All linear combinations of the vectors in S. Always a subspace, even when S isn't.

$\ker T$ — **the kernel of T, the null space**

Everything mapped to zero.

$\operatorname{im} T$	**the image of T, the column space, the range** Everything the map produces.
$V^{\perp}$	**V perp, the orthogonal complement** Every vector orthogonal to everything in V.
V^*	**V star, the dual space** The space of linear maps from V to the scalars. Same glyph as conjugate transpose, entirely different object.
$\dim V$	**the dimension of V** Size of any basis.

12.5 Norms and the subscripts on them

Measuring size

$\|\|v\|\|$	**the norm of v** Length. Double bars, to distinguish from absolute value.
$\|\|v\|\|_2$	**the two-norm, the Euclidean norm** $\sqrt{\sum v_i^2}$. The default when no subscript appears.
$\|\|v\|\|_1$	**the one-norm, the taxicab norm** $\sum \|v_i\|$. Central in compressed sensing and in Lasso regression.
$\|\|v\|\|_\infty$	**the infinity norm, the max norm** $\max_i \|v_i\|$.
$\|\|A\|\|_F$	**the Frobenius norm** Treat the matrix as a long vector and take the two-norm.

12.6 Writing a matrix

Brackets are not standardised, and none of the variants carries meaning.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

The first two are the same matrix; round brackets are more common in mathematics, square in engineering and numerical work. The third, with plain vertical bars, is *not* a matrix. It's the determinant, a number. That's a real distinction hiding in a choice of bracket.

12.7 Index notation

Entries and indices

a_{ij} **$a\ i\ j$**

The entry in row i, column j. **Row first, always.** Also written A_{ij}, $(A)_{ij}$, or $A[i][j]$ in code.

Not: column first. The convention is universal in mathematics and reversed in some graphics libraries, which is a real source of bugs.

δ_{ij} **the Kronecker delta**

1 if $i = j$, else 0. The identity matrix, written entry-wise.

Not: the Dirac delta $\delta(x)$, which is a completely different object.

ε_{ijk} **the Levi-Civita symbol**

$+1$ for even permutations of 123, -1 for odd, 0 if any index repeats. How cross products and determinants get written in index notation.

$A_i{}^j$ **mixed indices**

In tensor work, upper and lower indices mean different things: contravariant and covariant. Position matters and is not decoration.

HOW IT IS ACTUALLY USED

Einstein summation is the convention that a repeated index implies a sum: $a_i b_i$ means $\sum_i a_i b_i$, with the $\sum$ left out. It is universal in relativity and continuum mechanics and absent from most linear algebra courses, so the first time you meet it the formulas look like they are missing something. They are. The sum sign was deleted on purpose, and the rule is that you restore one for every index that appears twice.

12.8 Vectors: four ways to write one

$\boldsymbol{v}$	Bold. Standard in printed mathematics and engineering
$\vec{v}$	Arrow. Standard in handwriting and in physics teaching, since bold cannot be handwritten
$\underline{v}$	Underlined. The handwritten stand-in for bold in some UK departments
v	Plain, when the context makes it obvious. Common in pure mathematics, where everything is a vector anyway

Unit vectors take a hat: $\hat{v}$, and the standard basis is $\hat{\imath}, \hat{\jmath}, \hat{k}$ in physics, e_1, e_2, e_3 in mathematics.

PRACTISE THIS

Take one matrix identity from your notes, say $(AB)^\top = B^\top A^\top$, and write it out in index notation with explicit sums. Then write it again using Einstein summation. Doing this once makes index-heavy papers readable for good.

13

Probability and Statistics

A vertical bar that isn't division, a tilde that isn't approximation, and the hat-versus-no-hat distinction that the whole subject rests on.

13.1 The basic operators

Probability

$\Pr(A)$ — **the probability of A**

Also written $P(A)$, $\mathbb{P}(A)$, or $\mathrm{P}(A)$. All identical.

$\Pr(A \mid B)$ — **the probability of A given B**

Conditional probability. **The bar is not division.** It reads "given".

Not: a ratio you can rearrange freely. $\Pr(A \mid B)$ *and* $\Pr(B \mid A)$ *are different numbers, and confusing them is the base-rate fallacy.*

$\mathbb{E}[X]$ — **the expectation of X, the expected value, the mean**

Also $E[X]$, EX, or $\langle X \rangle$ in physics. Square brackets are conventional because the argument is a random variable, not a number.

$\mathrm{Var}(X)$ — **the variance of X**

$\mathbb{E}[(X - \mathbb{E}X)^2]$. Also written σ^2 or $\mathbb{V}(X)$.

$\mathrm{Cov}(X, Y)$ — **the covariance**

Ω — **the sample space**

The set of all outcomes. Same glyph as big-omega growth notation and as a physical region.

13.2 The tilde, which is not approximation

Distribution

$X \sim N(\mu, \sigma^2)$ — **X is distributed as, X follows**

Declares which distribution a random variable has. Reads "is distributed as".

Not: "is approximately". In statistics $\sim$ almost always means "is distributed as", and it is one of the most common misreadings by students arriving from a pure-mathematics course.

$X \perp\!\!\!\perp Y$ **X is independent of Y**

Independence. Also written $X \perp Y$, which collides with orthogonality, and sometimes just stated in words.

$X_i \overset{\text{iid}}{\sim} F$ **are independent and identically distributed as F**

Three claims at once: independent, identically distributed, and following F. Nearly every statistical method assumes it.

13.3 Common distributions, and their parameters

The parameter list is part of the name, and the conventions differ enough to matter.

$N(\mu, \sigma^2)$	Normal. **Usually the second parameter is the variance**, not the standard deviation. Some software takes the standard deviation instead, and this causes real errors
$\text{Bin}(n, p)$	Binomial: n trials, success probability p
$\text{Poisson}(\lambda)$	Poisson: rate λ
$U(a, b)$	Uniform on the interval
$\text{Exp}(\lambda)$	Exponential. λ is the rate, so the mean is $1/\lambda$. Some texts parameterise by the mean instead
$\text{Beta}(\alpha, \beta)$, $\text{Gamma}(k, \theta)$	Parameter conventions vary widely. Always check

> *Whenever you meet a distribution written with parameters, check what the parameters mean in that source. Rate versus mean, and variance versus standard deviation, are the two that silently break results.*

13.4 Hats, bars, and the population-sample distinction

This is the notational heart of statistics, and once you see it the subject reorganises itself.

Quantity	Population	Sample estimate
Mean	μ	$\bar{x}$ or $\hat{\mu}$
Variance	σ^2	s^2 or $\hat{\sigma}^2$
Proportion	p	$\hat{p}$
Regression coefficient	β	$\hat{\beta}$
Any parameter	θ	$\hat{\theta}$

Greek letters are truths about the world that you do not know. Roman letters and hatted symbols are numbers you computed from data.

> *Greek is the unknown truth. Hats and Roman are your estimate of it. Every inference procedure in statistics is a claim about how far the second is from the first.*

Decorations

$\bar{x}$ — x **bar, the sample mean**

The average of your data.

Not: μ, which is the thing you are trying to estimate.

$\hat{\theta}$ — **theta hat, the estimate of theta**

A number computed from the sample. Itself a random variable, which is why it has a distribution and a standard error.

$\tilde{x}$ — x **tilde**

Usually the median, sometimes a transformed or predicted value. Less standardised than the others.

$\hat{y}$ — y **hat, the fitted value, the prediction**

In regression, the model's output for a given input.

13.5 Test statistics and inference

Inference vocabulary

H_0 — **H nought, the null hypothesis**

The default claim being tested, usually "no effect".

H_1 — **H one, the alternative hypothesis**

Also H_a.

p — **the p-value**

The probability of data at least this extreme *if H_0 is true.*

Not: the probability that H_0 is true. That is a different quantity and requires a prior. This is the single most common misinterpretation in all of applied statistics.

α — **alpha, the significance level**

The false-positive rate you are willing to accept. Conventionally 0.05, for no principled reason.

β — **beta, the type II error rate**

Power is $1 - \beta$. Note that β also means a regression coefficient, in the same textbooks, often on the same page.

$\mathcal{L}(\theta)$ — **the likelihood of theta**

The probability of the observed data, viewed as a function of the parameter.

Not: the probability of θ. Likelihood is not a probability distribution over θ.

13.6 Bayesian notation

Bayes

$\pi(\theta)$ — **the prior**

Your belief about θ before seeing data. Sometimes $p(\theta)$.

$p(\theta \mid x)$ **the posterior**

Belief after seeing the data. The output of a Bayesian analysis.

$\propto$ **is proportional to**

Ubiquitous in Bayesian work: posterior $\propto$ likelihood $\times$ prior, dropping the normalising constant because it does not depend on θ.

13.7 The bar, which does three jobs

$\Pr(A \mid B)$	"given". Conditioning
$\{x \mid P(x)\}$	"such that". Set-builder
$a \mid b$	"divides". Number theory
$\lvert x\rvert$	Absolute value, when there are two of them

Four jobs for one glyph, and only position and context distinguish them. Chapter 16 has the full list.

HOW IT IS ACTUALLY USED

Statisticians are notably relaxed about notation compared with pure mathematicians, and it catches people arriving from a maths degree. $\mathbb{E}X$, $\mathbb{E}[X]$ and $\mathbb{E}(X)$ all appear; P, Pr and $\mathbb{P}$ are used interchangeably; and the same paper will write $N(\mu, \sigma^2)$ and $N(\mu, \sigma)$ if two authors wrote different sections. Read for meaning and check the parameterisation, rather than expecting consistency.

PRACTISE THIS

Take a result from a statistics course and mark every symbol as either a population quantity or a sample estimate. Then find the p-value in it and write out, in a full sentence, what event has that probability. If the sentence does not start "the probability of data this extreme, assuming the null", rewrite it until it does.

14

Physics and Engineering Conventions

Where the physicists do it differently, and why.

Much of this chapter is about the same mathematics you've already met, written by a community with different priorities: dimensional consistency, coordinate independence, and being able to write it on a blackboard at speed.

14.1 Vector calculus

The del operator

∇ — **del, nabla**

The vector of partial derivatives. Not a quantity on its own; it is an operator waiting for something to act on.

∇f — **grad f, the gradient**

Takes a scalar field, gives a vector field. Points in the direction of steepest increase.

$\nabla \cdot \boldsymbol{v}$ — **div v, the divergence**

Takes a vector field, gives a scalar field. How much the field is spreading out at a point.

$\nabla \times \boldsymbol{v}$ — **curl v**

Takes a vector field, gives a vector field. How much it circulates. Only defined in three dimensions.

$\nabla^2 f$ — **the Laplacian of f**

$\nabla \cdot \nabla f$. Also written Δf, which collides badly with "a finite change in f".

The input and output types are the fastest way to keep these straight:

Operator	Takes	Gives
grad	scalar	vector
div	vector	scalar
curl	vector	vector
Laplacian	scalar	scalar

If you write $\nabla \cdot f$ with f a scalar, or ∇f with f a vector, the expression is meaningless. Type-checking these is a fast way to catch algebra errors.

14.2 Constants you're expected to recognise

Named constants

$\hbar$ — **h bar, the reduced Planck constant**
$h/2\pi$. Quantum mechanics is written in terms of $\hbar$ almost throughout.

c — **the speed of light**
Also a generic constant, a specific heat, or a concentration. Context.

ε_0, μ_0 — **epsilon nought, mu nought**
Permittivity and permeability of free space.

k_B — **k B, the Boltzmann constant**
The subscript matters: k alone is usually a wavenumber or a spring constant.

Ω — **ohm, or a solid angle, or a region**
Three unrelated meanings, all common in physics.

14.3 Dirac notation

Standard in quantum mechanics, and impenetrable until somebody tells you it's just inner products.

Bra-ket

$|\psi\rangle$ — **ket psi**
A vector in a complex vector space, representing a state.

$\langle\psi|$ — **bra psi**
The dual vector: the conjugate transpose of the ket.

$\langle\phi \mid \psi\rangle$ — **bra ket, the inner product of phi and psi**
A complex number. The bracket, which is where "bra" and "ket" come from.

$|\psi\rangle\langle\phi|$ — **ket bra**
An *operator*, not a number. Outer product. The order of bra and ket is the whole difference.

$\langle\hat{A}\rangle$ — **the expectation value of A**
$\langle\psi|\hat{A}|\psi\rangle$. The same angle brackets as the classical average, deliberately.

$\hat{A}$ — **A hat, the operator A**
In quantum mechanics a hat means an operator, not a unit vector and not an estimate. Three fields, three meanings for one decoration.

14.4 Units, and writing them properly

Notation rules that matter in lab reports and are marked.

$9.81\ \mathrm{m\,s^{-2}}$	Upright roman for units, always. Italic is for variables
Space before the unit	5 kg, not 5kg. The exceptions are °, ′ and ″
$\mathrm{m\,s^{-1}}$	Preferred over m/s in formal work; unambiguous for compound units
Capitalisation	Units named after people take a capital symbol and a lowercase name: 5 newtons, 5 N
25 ▪C	Degree symbol attached to the C, with a space before the number's unit

> *Variables are italic. Units, and named functions like* sin, log, det, *are upright. That single distinction is what makes m (a mass) different from* m *(a metre), and it is the most frequently ignored typographic rule in student reports.*

14.5 Notation that means something different here

The collisions that matter most when moving between a maths course and a physics one.

Symbol	In mathematics	In physics
$\hat{x}$	An estimate, or a unit vector	An operator, or a unit vector
A^*	Conjugate transpose	Complex conjugate only
$A^\dagger$	Pseudoinverse	Conjugate transpose
Δ	The Laplacian, or a change	Almost always a change
$\langle\cdot\rangle$	Inner product, or span	A time or ensemble average
δ	Kronecker or Dirac delta	Also a variation, also a small change
j	An index	$\sqrt{-1}$, in electrical engineering, since i is current
$\equiv$	Identity or congruence	Nearly always a definition

That j row is not a joke. Electrical engineering uses j for the imaginary unit throughout, and a signals textbook will write $e^{j\omega t}$ where a mathematics text writes $e^{i\omega t}$.

14.6 Approximation notation, used heavily

Physicists approximate constantly and have notation for the different kinds.

$\approx$	Approximately equal. The general-purpose one
$\simeq$	Approximately, often with the sense "to the order we care about"
$\sim$	Of the same order of magnitude. "$\sim 10^{-9}$ m" means roughly a nanometre
$\propto$	Proportional to, dropping constants
$\approx$ with $\ll$	"since $v \ll c$, $\gamma \approx 1$". The inequality justifies the approximation

HOW IT IS ACTUALLY USED

Physicists write $\sim$ for order-of-magnitude far more often than mathematicians, who reserve it for the precise asymptotic statement that the ratio tends to one. Reading a physics paper with the mathematician's definition in mind makes the claims look absurdly strong. Reading a mathematics paper with the physicist's makes them look vague. Both communities are being precise about different things.

14.7 Coordinate and index conventions

Tensor indices

x^μ

x upper mu, a contravariant component

Upper index. In relativity, μ runs 0 to 3 with 0 the time coordinate.

Not: a power. Upper indices in tensor work are labels, not exponents, which is genuinely confusing and completely standard.

x_μ

x lower mu, a covariant component

Lower index. Related to the upper version by the metric.

$g_{\mu\nu}$

the metric tensor

Defines distances. Raises and lowers indices.

∂_μ

partial mu

Shorthand for $\partial/\partial x^\mu$. Note the index position flips when the derivative is written this way.

A repeated index, one up and one down, implies a sum. That's the Einstein convention of Chapter 12, and in relativity it's not optional: the sums are never written.

PRACTISE THIS

Take a formula from a physics course and type-check it: mark each quantity as scalar, vector, or operator, and confirm every grad, div and curl has the right kind of input. Then check the units balance on both sides. Those two checks catch most algebra errors before you ever look at the numbers.

15

Four Fields That Change the Vocabulary

By the time you reach category theory, combinatorics, order theory, or formal languages, most of the symbols are familiar shapes doing unfamiliar jobs. The arrow may describe a morphism, a derivation, or a transition. A star may mean repeated concatenation rather than conjugation. A wedge may mean meet rather than logical and.

This chapter gives you the compact vocabulary each field expects before the first lecture is over.

15.1 Category theory

Category theory is written with objects and arrows. The symbols often hide the kind of object on purpose because the same statements are meant to work for sets, groups, spaces, and many other categories.

Categories, objects, and arrows

$\mathcal{C}, \mathcal{D}$	**the categories C and D** Collections of objects together with morphisms between them. Script capitals usually name categories. *Not: ordinary sets unless the context says the category itself is being treated as a set.*
$A \in \mathrm{Ob}(\mathcal{C})$	**A is an object of C** A is one of the objects of the category $\mathcal{C}$. *Not: A set element in the ordinary membership sense; many categories are too large to be sets.*
$f : A \to B$	**f from A to B** A morphism with domain A and codomain B. What morphisms are depends on the category: functions, homomorphisms, continuous maps, and so on. *Not: Necessarily a function between underlying sets, though in familiar concrete categories it often is.*
$\mathrm{Hom}_{\mathcal{C}}(A, B)$	**Hom C from A to B** The collection of morphisms $A \to B$ in $\mathcal{C}$. Also written $\mathcal{C}(A, B)$. *Not: A homomorphism. The whole expression is a collection, not one map.*
$g \circ f$	**g after f** Composition: first apply $f : A \to B$, then $g : B \to C$. *Not: Multiplication. The rightmost morphism acts first.*

id_A

the identity on A

The identity morphism $A \to A$, neutral for composition.

Not: The number 1, even when the category is algebraic.

Functors and natural transformations

$F : \mathcal{C} \to \mathcal{D}$

the functor F from C to D

Maps objects of $\mathcal{C}$ to objects of $\mathcal{D}$ and morphisms to morphisms while preserving identities and composition.

Not: Merely a function between two collections. It must respect the categorical structure.

$\mathcal{C}^{\mathrm{op}}$

C op, the opposite category

Has the same objects as $\mathcal{C}$ and every arrow reversed.

Not: An inverse or a complement of $\mathcal{C}$.

$\eta : F \Rightarrow G$

eta from F to G

A natural transformation between functors. Its components are morphisms $\eta_A : F(A) \to G(A)$ that commute with the relevant arrows.

Not: Ordinary logical implication. The double arrow lives between functors here.

$F \dashv G$

F is left adjoint to G

An adjunction between functors, usually witnessed by a natural correspondence between two Hom collections.

Not: Perpendicularity.

HOW IT IS ACTUALLY USED

A commutative diagram says that every directed path with the same start and finish gives the same composite morphism. The picture is an equation system. “The square commutes” means $h \circ f = k \circ g$ for the four arrows around that square.

15.2 Combinatorics

Combinatorics compresses counting instructions. The danger is treating an index set, a coefficient extractor, or a partition symbol as ordinary algebra.

Finite sets and selections

$[n]$

bracket n

Usually the set $\{1, 2, \dots, n\}$ in combinatorics and theoretical computer science.

Not: A closed interval. There is no comma and the entries are integers.

$\binom{n}{k}$

n choose k

The number of k-element subsets of an n-element set.

Not: The fraction n/k, or an ordered selection.

$\binom{n}{k_1,\dots,k_r}$

the multinomial coefficient

The number of ways to split n labelled objects into ordered groups of sizes $k_1, \dots, k_r$, where the sizes sum to n.

Not: A vector of binomial coefficients.

S_n **S n, the symmetric group**

The group of all permutations of $[n]$.

Not: A sequence term unless the surrounding text defines one.

$\lambda \vdash n$ **lambda partitions n**

λ is an integer partition of n.

Not: Divisibility. The turnstile shape has a field-specific meaning here.

$(n)_k$ **n falling k**

Often the falling factorial $n(n-1)\cdots(n-k+1)$. Some authors use the same notation for a rising factorial, so check the definition.

Not: An ordered pair.

Generating functions and graphs

$[x^n]F(x)$ **the coefficient of x to the n in F**

Extracts the coefficient multiplying x^n in the series $F(x)$.

Not: Evaluation of F at x^n. The brackets are an operator.

$G = (V, E)$ **G with vertex set V and edge set E**

A graph specified by its vertices and edges.

Not: A coordinate pair in the plane.

$N(v)$ **the neighbourhood of v**

The set of vertices adjacent to v. Closed neighbourhood is often $N[v] = N(v) \cup \{v\}$.

Not: A numeric function unless the author defines it that way.

$d(v), \deg(v)$ **the degree of v**

The number of edges incident with v, or neighbours of v in a simple graph.

Not: A derivative.

$G[S]$ **G induced by S**

The subgraph of G induced by the vertex set S.

Not: An array lookup.

15.3 Order theory

An order is a relation, not necessarily a numerical comparison. The first habit is to ask whether the order is partial or total and whether the symbol is strict.

Ordered sets

$(P, \preceq)$ **P ordered by precedes**

A partially ordered set: $\preceq$ is reflexive, antisymmetric, and transitive on P.

Not: A metric comparison. Two elements may be incomparable.

$x \prec y$ **x strictly precedes y**

Usually $x \preceq y$ and $x \neq y$. Some authors use it for a cover relation instead, so check.

Not: Automatically meaning x is numerically smaller.

$x \parallel y$ **x is incomparable with y**

Neither $x \preceq y$ nor $y \preceq x$.

Not: Parallel vectors or lines in this context.

$x \lessdot y$ **y covers x**

$x \prec y$ with no element strictly between them.

Not: An approximation.

$x \wedge y$ **x meet y**

The greatest lower bound of x and y when it exists.

Not: Logical and, although meet behaves like and in Boolean lattices.

$x \vee y$ **x join y**

The least upper bound of x and y when it exists.

Not: Logical or, though the analogy is intentional in Boolean lattices.

$\downarrow x$ **down-set generated by x**

Usually $\{y : y \preceq x\}$. Dually, $\uparrow x = \{y : x \preceq y\}$.

Not: A limit or decreasing sequence.

> *Minimal is not the same as minimum. A minimum lies below every element. A minimal element merely has nothing strictly below it. A partial order can have several minimal elements and no minimum.*

15.4 Formal languages and automata

Formal-language notation treats strings as mathematical objects. Product means concatenation, a star means arbitrary finite repetition, and the empty string is not the empty set.

Alphabets, words, and languages

Σ **capital sigma, the alphabet**

A finite set of symbols from which strings are built.

Not: A summation. Here it is a set.

ε **epsilon, the empty string**

The unique string of length zero. Often written λ in older texts.

Not: The empty set, or a small positive real.

Σ^* **sigma star**

The set of all finite strings over Σ, including ε.

Not: A dual, conjugate, or repeated multiplication of a number.

Σ^+ **sigma plus**

All nonempty finite strings over Σ: $\Sigma^* \setminus \{\varepsilon\}$.

Not: A positive part in the numeric sense.

$|w|$ **the length of w**

The number of symbols in the word w.

Not: Absolute value, though the same bars deliberately measure size.

uv **u concatenated with v**

The word obtained by writing v immediately after u.

Not: Numeric multiplication.

$L \subseteq \Sigma^*$ **L is a language over Sigma**

A formal language is any set of finite strings over the alphabet.

Not: A grammar or an automaton. Those can describe or recognise a language but are not the same object.

L^* **Kleene star of L**

All strings formed by concatenating finitely many words from L, including zero words, which gives ε.

Not: The set of repeated characters from one word.

Grammars and automata

$u \Rightarrow_G v$ **u derives v in G**

One derivation step under grammar G. The reflexive-transitive closure is often $\Rightarrow_G^*$.

Not: Logical implication between propositions.

$L(G)$ **the language generated by G**

All terminal strings derivable from the grammar's start symbol.

Not: A function value in the ordinary sense, though the notation is function-shaped.

$\delta(q, a)$ **delta of q comma a**

The transition function sends state q and input symbol a to a next state, or to a set of states in a nondeterministic automaton.

Not: A small change.

$q_0,\ F$ **q nought and F**

Usually the start state q_0 and the set F of accepting states.

Not: Universal notation. These letters are conventional but must still be defined.

15.5 The transfer habit

When one of these fields is new, make a four-column page:

Glyph	Type	Field meaning	Old collision
$[n]$	set	$\{1, \dots, n\}$	interval
$\wedge$	operation	meet	logical and
Σ^*	set	all finite words	conjugate
$\eta : F \Rightarrow G$	family of maps	natural transformation	implication

The last column is the useful one. New notation rarely fails because the glyph is invisible. It fails because an old meaning arrives first and quietly controls the sentence.

PRACTISE THIS

Take the first page of notes from one of these fields. Mark every symbol whose old meaning differs from its local one, then state the type of the whole expression. If you can say what kind of thing each side is before computing, the notation has stopped being a wall.

Part III

Where It Goes Wrong

16

One Symbol, Many Meanings

There aren't enough symbols.

Mathematics has a few hundred usable glyphs and many thousands of concepts, so collisions are inevitable. What makes them dangerous is that each community is so used to its own reading that nobody flags it.

16.1 The worst offenders

| | and |, the vertical bar

Form	Means	Where
$\lvert x\rvert$	absolute value	everywhere
$\lvert z\rvert$	modulus	complex analysis
$\lvert A\rvert$	cardinality	set theory
$\lvert A\rvert$	determinant	linear algebra
$\lvert G\rvert$	order of a group	algebra
$a \mid b$	*a* divides *b*	number theory
$\Pr(A \mid B)$	given	probability
$\{x \mid P(x)\}$	such that	set-builder
$f\vert_S$	restricted to	analysis

Nine meanings. The saving grace is that position disambiguates: a bar *between* two things is a relation, a bar *around* something is a size, and a subscripted bar is a restriction.

∼, the tilde

is distributed as	statistics. **The most common meaning by volume**
is asymptotically equal to	analysis: the ratio tends to 1
is of the same order	physics: within a factor of ten or so
is similar to	geometry: same shape
is equivalent to	any equivalence relation
is approximately	informal writing
negation	older logic texts

This is the single most overloaded symbol in the book, and the readings are not close to each other. A statistician's $X \sim N(0,1)$ and an analyst's $f(n) \sim n^2$ have nothing in common.

$\equiv$

is defined as	physics, logic, and much of applied work
is identically equal to	analysis: holds for all values, not just some
is congruent to	number theory, always with a modulus
is logically equivalent to	logic

$\subset$

Covered in Chapter 7, and it belongs here too: it means $\subseteq$ to most modern authors and $\subsetneq$ to a substantial minority, with no way to tell from the glyph.

log

base e	pure mathematics, statistics, most of physics
base 10	engineering, chemistry, school textbooks
base 2	computer science, information theory

All three communities consider theirs obvious. In an algorithms course $\log n$ is base 2; in an analysis course it's base e; in a chemistry paper $\log[H^+]$ is base 10. Note that the base rarely matters for $O(\log n)$, since bases differ by a constant factor, and it matters enormously for a numerical answer.

(a, b)

ordered pair	everywhere
open interval	analysis
gcd	number theory
inner product	some linear algebra texts
the group generated by	algebra, though usually with angle brackets

16.2 The full collision list

Appendix D has every collision in the book as one sorted table, which is the format you want when you're staring at a symbol and need to work out which reading applies.

The commonest others, in brief:

$\times$	multiplication, cross product, Cartesian product
$\cdot$	multiplication, dot product, function composition, a placeholder for an argument
$\oplus$	direct sum, exclusive or, modulo-2 addition
Ω	sample space, growth lower bound, solid angle, ohm, a region
δ	Kronecker delta, Dirac delta, a small change, a variation
Δ	a change, the Laplacian, a triangle, a discriminant
λ	eigenvalue, Poisson rate, wavelength, Lagrange multiplier, lambda abstraction
σ	standard deviation, a permutation, a sigma-algebra, a stress component
ϕ, φ	the golden ratio, Euler's totient, an angle, a homomorphism, a potential
e	Euler's number, the identity of a group, a basis vector, the charge of an electron
i	$\sqrt{-1}$, an index. And j for $\sqrt{-1}$ in electrical engineering
$*$	convolution, adjoint, dual, a general binary operation, the Kleene star
$'$	derivative, a different variable, a complement, a transpose, minutes of arc
$\hat{}$	unit vector, an estimate, a quantum operator, a Fourier transform
$\square$	end of proof, necessity in modal logic, the d'Alembertian in relativity

16.3 How to resolve a collision, in order

Four moves, cheapest first. Most collisions resolve at step 1.

1 What subject is this?
A tilde in a statistics course is "distributed as". That single question settles most cases.

2 What are the operands?
| around a set is cardinality; around a matrix, a determinant. Type-check what is on each side.

3 Check the notation section
Most textbooks have one, usually unread, in the front matter or an appendix.

4 Try both and see which is true
If one reading makes the theorem false, that was not the reading. Legitimate and fast.

Figure 16.1: Resolving an ambiguous symbol. Do not skip step 1; the answer is usually there and it takes thirty seconds.

Step 4 deserves defending. It sounds like guessing and it isn't: you're using the fact that published theorems are usually true as evidence about notation. Working mathematicians do this constantly and quickly.

HOW IT IS ACTUALLY USED

The reason nobody fixes these collisions is that notation is not legislated. It spreads by imitation, each community optimises for its own most frequent case, and a symbol that has meant one thing in a field for eighty years is not going to change because it clashes with something two buildings away. Learning to read across the boundary is easier than waiting for the boundary to move.

16.4 The one you'll hit first

If you're in first year: it's $\subset$, and then log.

If you've just started statistics after a mathematics degree: it's $\sim$, and it will cost you an afternoon before you notice.

If you've moved from mathematics into physics or engineering: it's the $\hat{}$ and $\dagger$ pair, and the fact that Δ almost never means the Laplacian.

PRACTISE THIS

Start a collision page at the back of your notation sheet. Every time you meet a symbol that means something different from what you expected, write the symbol, the two meanings, and the course. By the end of a year it will be the most-used page in the notebook.

17

Fourteen Misreadings That Cost Marks

Every one of these is common, every one produces a confidently wrong answer, and every one is invisible to the person making it.

They're ordered roughly by how often markers see them.

17.1 1. Reading $\Rightarrow$ as causation

$P \Rightarrow Q$ is a claim about truth values, not about mechanism or time. From it you may conclude $\neg Q \Rightarrow \neg P$ and nothing else.

The cost: using the converse or the inverse in a proof. Both are invalid and both feel obvious.

17.2 2. Getting quantifier order backwards

$\forall x\, \exists y$ and $\exists y\, \forall x$ are different claims, and the second is much stronger.

The cost: proving the wrong statement, usually the easier one, and losing every mark for the question. Chapter 3.

17.3 3. Assuming $\subset$ is strict

Most modern authors allow equality.

The cost: an evening spent constructing a counterexample to a true theorem.

17.4 4. Treating $\log$ as base 10

Or as base e, or as base 2, in a course that means one of the others.

The cost: a numerical answer wrong by a factor of 2.3 or 3.32. Painful because the method was right.

17.5 5. Reading $\sim$ as "approximately"

In statistics it means "is distributed as". In analysis it means the ratio tends to exactly one.

The cost: in statistics, not realising a distribution has been specified. In analysis, thinking a precise statement is a vague one.

17.6 6. Reading $\sin^{-1}$ as $1/\sin$

$\sin^{-1}$ is the inverse function. $\sin^2$ *is* the square. Same position, different rule.

The cost: wrong by a lot, on a question you understood.

17.7 7. Confusing $\in$ and $\subseteq$

Elements versus subsets. $1 \in \{1,2\}$ and $\{1\} \subseteq \{1,2\}$ are both true; $\{1\} \in \{1,2\}$ is false.

The cost: a whole question in a first set theory or analysis course, where this distinction is most of what's being tested.

17.8 8. Using a bound variable outside its scope

Writing "$\sum_{k=1}^{n} a_k$, hence $k \leq n$". There is no k outside the sum.

The cost: marks for rigour, and a marker who now reads the rest of the proof suspiciously.

17.9 9. Claiming a lower bound from O

O is a ceiling. "The algorithm is $O(n^2)$, so it can't be faster" is exactly backwards.

The cost: the entire point of a complexity question.

17.10 10. Reading p as the probability the null is true

The p-value is the probability of data this extreme *given* the null.

The cost: conceptual, and it survives into professional work. This misreading appears in published papers.

17.11 11. Treating ∂ as a quantity

You can informally manipulate dy/dx as a fraction. You cannot do the same with partials: $\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1$, not 1.

The cost: a sign error that propagates through a thermodynamics or multivariable calculus problem.

17.12 12. Assuming $0 \in \mathbb{N}$, or assuming it isn't

Depends entirely on the author.

The cost: an induction proof with the wrong base case, or an off-by-one in a sequence definition.

17.13 13. Reading $f^{-1}(S)$ as requiring an inverse

The preimage is defined for every function. f^{-1} as a *function* requires bijectivity; f^{-1} as a *preimage* does not.

The cost: being unable to start most topology questions, since continuity is defined by preimages.

17.14 14. Dropping the modulus from a congruence

$a \equiv b$ says nothing without $\pmod{n}$.

The cost: marks, and genuine ambiguity, since $\equiv$ has three other meanings.

17.15 The one-page check

Before you hand anything in, scan for these six specifically. They're the ones that appear in written work rather than only in reading.

Every $\Rightarrow$	Have you used only it, and not its converse?
Every quantifier pair	Is the order the one you meant?
Every bound variable	Used only inside its binder?
Every $\subset$	Replaced by $\subseteq$ or $\subsetneq$?
Every $\equiv$	Modulus stated, if it's a congruence?
Every O	Upper bound where you meant upper bound?

HOW IT IS ACTUALLY USED

Markers see the same handful of these every year, and the pattern is consistent: they cluster in the work of students who understand the material well. That isn't a coincidence. A student who is unsure writes slowly and checks; a student who is confident writes at speed and lets a near-miss through. If you are good at your subject, these are the errors that are costing you, not the conceptual ones.

PRACTISE THIS

Take a marked problem sheet you got back and find every lost mark. Sort them into "I did not understand this" and "I wrote what I did not mean". If the second pile is bigger than you expected, that's this chapter's list, and it's much cheaper to fix than the first pile.

18

Decorations

Hats, bars, tildes, primes, stars, and daggers.

A decoration is a modifier attached to a symbol, and unlike most notation it carries almost no fixed meaning. What it means is set by the field and often by the individual author, which makes decorations the place where you most need to check rather than assume.

18.1 The hat

$\hat{x}$

unit vector	physics, geometry: $\hat{v} = v/\|v\|$
an estimate	statistics: $\hat{\theta}$ is computed from data
an operator	quantum mechanics: $\hat{H}$ is the Hamiltonian
Fourier transform	analysis: $\hat{f}(\xi)$
a fitted value	regression: $\hat{y}$

Five meanings, and the first three are the ones you'll meet in the same year if you do joint honours. A statistician's $\hat{\sigma}$ and a physicist's $\hat{\sigma}$ (a Pauli matrix) share nothing but the mark.

18.2 The bar

$\bar{x}$

the mean	statistics: $\bar{x}$ is the sample average
complex conjugate	complex analysis: $\bar{z} = a - bi$
closure	topology: $\bar{A}$ is A with its limit points
complement	set theory: $\bar{A}$ is everything not in A
negation	logic and circuit design
an average over time	physics

Closure and complement are near-opposites, both written with a bar, and both appear in a first topology course. The context is usually clear and occasionally isn't, which is why many topologists prefer $\mathrm{cl}(A)$ and A^c.

18.3 The prime

x'

derivative	calculus: $f'(x)$
a different object	everywhere: "let x' be another such point". The most common use by volume
transpose	older statistics and econometrics: A' for $A^\intercal$
complement	set theory: A'
the derived set	topology
minutes of arc	astronomy, surveying. $''$ for seconds

> *When you write a prime, make sure the context tells the reader which of these it is. If you're differentiating in a document that also uses primes for "a second point of the same kind", switch one of them to a different notation.*

18.4 Stars and daggers

A^* and $A^\dagger$

	A^*	$A^\dagger$
Linear algebra	conjugate transpose	pseudoinverse
Physics	complex conjugate	conjugate transpose
Functional analysis	adjoint, or dual space	adjoint
Optimisation	the optimal value: x^*	
Formal languages	Kleene star: zero or more	

This is a genuine trap when moving between a mathematics and a physics course, because both use both symbols and they swap meanings.

18.5 Tildes, dots, and the rest

$\tilde{x}$	the median; a transformed variable; a modified version of an object; an equivalence class representative
$\dot{x}$	time derivative, in mechanics. Almost never anything else, which makes it one of the few reliable decorations
$\ddot{x}$	second time derivative
$\vec{v}$	a vector, in handwriting and in physics teaching
$\underline{v}$	a vector, in some UK departments; a matrix in others
$\boldsymbol{v}$	a vector, in print
$\mathring{A}$	the interior, in topology. Rare
x°	degrees; or the interior; or a composition

18.6 Sub- and superscripts as labels

Not everything raised or lowered is arithmetic.

x_1, x_2	Labels for different objects. Not multiplication
a_{ij}	Two labels: row and column
$\mathbb{R}^n$	Not a power: a Cartesian product
$f^{(3)}$	The third derivative. Brackets distinguish it from a power
f^3	A cube, or a threefold composition, depending on the field
x^μ	A tensor index, not a power
A^c	A complement, not a power
L^2	A named function space, not a square
H_0	A hypothesis label
$\log_2$	A base

That f^3 row is a real ambiguity. In analysis it usually means $(f(x))^3$; in dynamical systems it usually means $f(f(f(x)))$. Careful authors write $f^{\circ 3}$ for the composition.

HOW IT IS ACTUALLY USED

The reliable signal is not the decoration, it's the field. Before reading a decorated symbol, ask what subject the document belongs to and read the decoration the way that subject reads it. A single paper almost never uses a decoration two ways; different papers routinely do.

18.7 Writing decorations yourself

Three rules that save your reader real effort.

1. **Declare it.** One sentence, once: "throughout, a hat denotes an estimator". Costs a line, prevents every misreading.
2. **Don't stack.** $\hat{\tilde{x}}'$ is unreadable and probably means you need a different letter.

3. **Don't reuse.** If a hat means "unit vector" in section 2, it may not mean "estimate" in section 5.

PRACTISE THIS

Go through one chapter of your current notes and list every decorated symbol with what it means there. If any decoration appears with two meanings in the same chapter, that's worth raising in a tutorial: it's usually an accident, and it's usually confusing more people than you.

19

Writing Notation Others Can Read

Up to here the book has been about reading. This chapter is about not inflicting the same problems on somebody else.

The audience is whoever marks your work, and eventually whoever referees your paper. Both are reading fast, and both will resolve an ambiguity in the way least favourable to you.

19.1 Declare your conventions, once

The single highest-value habit. One short paragraph at the start of any substantial piece of work.

> Throughout, $\mathbb{N}$ includes 0, $\subseteq$ denotes containment with equality permitted, log is base 2, and a hat denotes an estimator. Vectors are bold; matrices are uppercase roman.

Four lines. It removes every ambiguity in this book that a reader could otherwise hit, and it signals that you know the ambiguities exist.

> *State the convention for anything that has more than one common reading. The reader should never have to work out what you meant, because when they guess, they guess wrong.*

19.2 Choose the unambiguous variant

Where two notations exist and one is clearer, use the clear one. It costs nothing.

Instead of	**Write**	**Because**
$\subset$	$\subseteq$ or $\subsetneq$	the ambiguity disappears
$\lvert A\rvert$ for a determinant	$\det A$	the bars are overloaded
A' for transpose	$A^\top$	primes mean derivatives
$\sin^{-1}$	arcsin	it isn't a reciprocal
(a, b) for gcd	$\gcd(a, b)$	it isn't a pair
log where the base matters	ln, $\log_2$, $\log_{10}$	three communities disagree
f^3 for composition	$f^{\circ 3}$	it looks like a cube
Δ for the Laplacian	∇^2	Δ means "change"
$\equiv$ for a definition	$:=$	$\equiv$ has three other jobs

19.3 Bracket generously

Brackets you didn't need cost a reader nothing. Brackets you omitted cost them a re-read.

MISREAD

$\sum_{k=1}^{n} k+1$ $\qquad \int f(x)+g(x)\,dx$ $\qquad a/bc$

READ CORRECTLY

$\left(\sum_{k=1}^{n} k\right)+1$ $\qquad \int\big(f(x)+g(x)\big)\,dx$ $\qquad a/(bc)$

The third is worth dwelling on. a/bc is read as $(a/b)c$ by strict precedence and as $a/(bc)$ by most physicists. Both readings are defensible, which is exactly why you must not write it.

19.4 One letter, one job

Within a single document, a symbol should mean one thing.

Don't reuse n	for a sample size and a summation index
Don't reuse λ	for an eigenvalue and a Lagrange multiplier in the same derivation
Don't shadow	if k is fixed in the theorem, don't open a sum over k
Don't switch	if it's ε in section 2, it isn't ϵ in section 4

If you run out of letters, you have options before reuse: subscripts, a different alphabet, or, most often, a sign that the argument should be split into two lemmas.

19.5 Follow the conventions for letter choice

Covered in Chapter 4 and worth repeating: i, j, k for integer indices, x, y, z for reals, f, g, h for functions, ε, δ for small positive quantities.

Breaking these isn't wrong, and it makes every reader pause. Pauses accumulate.

19.6 Upright versus italic

Italic	Variables: x, f, n, θ
Upright	Named functions: sin, log, det, tr, max Units: kg, $\mathrm{m\,s^{-1}}$ Multi-letter names: Var, Cov, iid Constants some authors treat as fixed rather than variable: e, i, d in $\mathrm{d}x$

The rule that catches people: a multi-letter name in italics is read as a product. Writing *sin* says $s \times i \times n$. In LaTeX the fix is `\sin`, or `\operatorname{...}` for anything without a built-in command.

19.7 Display versus inline

Put an expression on its own line when it's the object of the sentence, when it's long, or when you'll refer back to it. Number it if you'll refer back to it.

Keep it inline when it's a small part of a sentence. A page where every formula is displayed is as hard to read as one where none is.

And write around your formulas. An equation dropped between two paragraphs with no introduction and no comment makes the reader do the connecting work.

MISREAD

$$\|x\|_1 \leq \sqrt{n}\,\|x\|_2$$

Therefore the result follows.

READ CORRECTLY

Comparing the two norms on $\mathbb{R}^n$ gives

$$\|x\|_1 \leq \sqrt{n}\,\|x\|_2,$$

where the factor $\sqrt{n}$ is attained when every coordinate is equal. The bound is therefore tight, which is what the argument needs.

Note the comma after the display. A displayed equation is part of the sentence and takes the punctuation the sentence needs.

HOW IT IS ACTUALLY USED

The mark of a well-written mathematical document is that you can read the prose alone, skipping every display, and still follow the argument. The formulas then supply the detail. Documents written the other way round, where the prose is glue between formulas, are the ones that get sent back.

19.8 Notation you should not invent

Sometimes you genuinely need a new symbol. Usually you don't.

Before inventing one, check: does the field already have a notation for this? Could a word do the job? Could an existing symbol with a subscript work?

If you do invent, three rules: define it prominently, use it consistently, and don't pick a glyph that already means something in your field. A new meaning for $\otimes$ in a paper about vector spaces will be misread by everyone.

PRACTISE THIS

Add a four-line conventions paragraph to the front of your next problem sheet or report. Then read the document once, looking only at whether any symbol does two jobs. Both take ten minutes and both are visible to whoever marks it.

20

Writing Symbols by Hand

Printed symbols arrive with different fonts and generous spacing. Handwritten symbols arrive through one pen, at lecture speed, on a page that may be read six months later. The collisions are different.

The goal is not beautiful handwriting. It is recoverable handwriting: you should be able to tell which symbol you meant without replaying the lecture.

20.1 Design differences, not decorative ones

Two symbols that differ only by size or neatness will eventually collide. Give each pair a structural distinction.

Pair	Risk	Reliable distinction
v, ν	Latin v / Greek nu	pointed v; rounded ν with a lead-in
u, μ	Latin u / Greek mu	give μ a long descending first stroke
$\phi, \emptyset$	phi / empty set	vertical-stroke ϕ; slashed oval for $\emptyset$
p, ρ	Latin p / rho	closed bowl on p; open curved top on ρ
$x, \times$	variable / multiplication	cursive x for variable; upright symmetric cross for operation
$0, O$	zero / capital O	slash zero when both occur; keep O round
$1, l$	one / lowercase ell	serif or top hook on 1; loop the ell
$z, 2$	z / two	cross the z when both occur
$\subset, \in$	subset / membership	open wide $\subset$; smaller three-stroke $\in$
∂, d	partial / d	curled ∂; upright open-loop d

Choose the distinctions once at the start of a course. Changing your Greek alphabet halfway through a notebook creates the same problem as changing notation halfway through a paper.

20.2 Size carries grammar

Some handwritten marks depend on size:

- a summation sign $\sum$ must be tall enough that its limits read as limits rather than exponents;
- set braces {} must be visibly different from parentheses;
- vector bars and hats must cover the whole symbol they decorate;
- a radical must extend over its full argument;
- fraction bars must be long enough to reveal the numerator and denominator as groups.

When space becomes tight, do not shrink until the grouping disappears. Move the expression to a new line. The page is allowed to use paper.

MISREAD

$$\sqrt{x}+1 \qquad \frac{a}{+}bc+d \qquad \sum i = 1ni$$

Each expression may be recoverable from context, but the handwriting has stopped carrying its grammar.

READ CORRECTLY

$$\sqrt{x+1}, \qquad \frac{a+b}{c+d}, \qquad \sum_{i=1}^{n} i.$$

The repair is spacing and enclosure, not prettier letters.

20.3 Decorations need attachment points

A prime, dot, bar, arrow, or hat belongs to one base symbol. Put it where that ownership is visible.

- Keep primes high and close: f' should not resemble $f\ \prime$ as a second object.
- Centre a dot over the variable: $\dot{x}$, not between x and the next symbol.
- Draw a vector arrow long enough to cover the letter but not its subscript.
- Use parentheses before decorating a compound expression: $\overline{z+w}$ rather than a bar whose endpoint is debatable.
- Leave horizontal air between adjacent decorated variables such as $\bar{x}\hat{y}$.

If the decoration cannot be attached clearly, write the meaning in a less compressed form. A derivative can be dx/dt instead of $\dot{x}$; a vector can be $\boldsymbol{v}$ or $\vec{v}$ depending on the course.

20.4 The five-line legibility test

Before a timed course settles into bad habits, write these five lines:

$$\begin{array}{l} u,\ \upsilon,\ \nu,\ \mu,\ p,\ \rho, \\ 0,\ O,\ 1,\ l,\ z,\ 2, \\ x \times y, \quad A \subset B, \quad x \in A, \\ \phi,\ \varphi,\ \emptyset,\ \partial,\ d, \\ \bar{x},\ \hat{x},\ \vec{x},\ x',\ \dot{x}. \end{array}$$

Leave the page for a day, then read it without the prompt. Circle every pair that requires context. Redesign only those pairs and repeat.

This is faster than discovering during revision that three months of lecture notes use one shape for both υ and ν.

20.5 Handwriting in a proof

Legibility is not limited to glyphs. The page must reveal the argument.

- Start a new line when the logical direction changes.
- Align equals signs only when every line is genuinely equal.
- Put reasons in the right margin: “def.,” “IH,” “triangle ineq.,” “by T3.4.”
- Box a final result, not an intermediate expression.
- Never let an arrow stand in for a sentence unless the course has declared that convention.

The reader of your exam script has no time to reverse-engineer a dense diagonal calculation. White space is part of the proof.

20.6 When accessibility changes the recommendation

Some readers have dysgraphia, tremor, low vision, or another reason that handwriting carries a higher cost. The goal remains recoverable work, not performing penmanship.

Use the accommodations available: typed assignments, a tablet with zoom, graph paper for alignment, a thicker pen, dictated scratch work, or extra time. For your own notes, a mixed system is legitimate: diagrams and short calculations by hand, long definitions and summaries typed.

The right medium is the one that preserves the mathematics and lets you retrieve it later.

PRACTISE THIS

Run the five-line legibility test and wait until tomorrow to mark it. Choose one permanent distinction for every pair you misread. Write those choices inside the cover of your current notebook, then use them for the rest of the course.

21

Meeting Notation You Have Never Seen

At some point a symbol appears that isn't in this book, isn't in your notes, and isn't in any glossary you can find.

This chapter is the procedure for that, because it will happen every year for the rest of your working life.

21.1 First: it is probably defined nearby

Most unfamiliar notation in a paper or textbook is defined in that document. People forget to look.

Within a page	Most local notation is introduced within a paragraph or two of first use. Search backwards first
The notation section	Many textbooks have one in the front matter. Many papers have one after the introduction
The first use	Search the document for the symbol and go to the earliest occurrence. That is nearly always the definition
An appendix	Longer papers put conventions there
A cited source	"Following the notation of [12]" is a real sentence that appears in real papers

> *Search backwards for the first occurrence before searching outwards. The definition is in the document about three times out of four, and finding it takes thirty seconds.*

21.2 Second: identify it by shape

If you can't name a symbol, you can't search for it. This is the specific trap, and it has specific solutions.

Appendix B	The look-up-by-shape index in this book: symbols grouped by what they look like rather than what they mean
Handwriting search	Detexify and similar tools let you draw a symbol and return the LaTeX command, which gives you the name
The Unicode name	Most symbols have an official name. Copy the character into a Unicode lookup and you get "N-ARY CIRCLED PLUS OPERATOR" or similar, which is searchable
Ask	Two minutes in a tutorial. Nobody thinks less of you; everybody has had to ask

That last row matters more than it looks. Students routinely spend an hour on something a supervisor would answer in ten seconds, because asking feels like an admission. It isn't. Notation is arbitrary, and not knowing an arbitrary convention says nothing about you.

21.3 Third: infer from use

If it genuinely isn't defined, you can usually recover the meaning from how it behaves. Four questions.

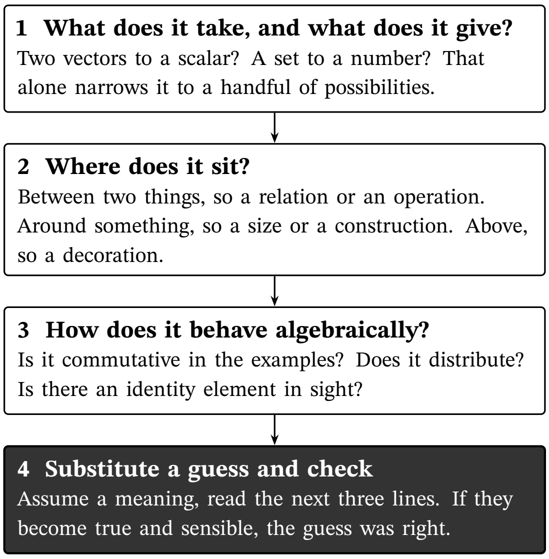

Figure 21.1: Reverse-engineering an undefined symbol from its behaviour. Type-checking the operands gets you most of the way in practice.

Step 4 is how experienced readers actually do it, and it's fast because published mathematics is usually correct. A wrong guess makes the following paragraph nonsense almost immediately.

21.4 Reading a paper in a field you don't know

You'll do this for a project or a dissertation, and the notation is usually the barrier rather than the mathematics.

1. **Read the notation section first**, before the abstract. Two minutes, and it changes how much of the rest lands.
2. **Build a glossary as you go.** One sheet, symbol and meaning. Ten entries covers most papers.
3. **Accept partial understanding.** You do not need every symbol to follow the argument. Skipping a term you can't identify is legitimate if the sentence still makes sense.
4. **Find the field's standard textbook** and read its notation chapter. One source will explain conventions that fifty papers assume.
5. **Check whether the field has a conventions document.** Many do: style guides from the main journal or society, and they exist because even insiders get confused.

21.5 When the notation is genuinely bad

Sometimes it isn't you. Notation in a paper can be inconsistent, undefined, or reused, and authors are human.

Signs it's the document rather than you: the same symbol used two ways within a page, a symbol that appears once and never returns, a definition that contradicts a later use, or subscripts that change convention between sections.

When you're confident it's the source, note it and move on with your best reading. If it's a paper you're refereeing one day, say so in the report, because it's a real defect and a fixable one.

HOW IT IS ACTUALLY USED

Nobody knows all of it. A working mathematician moving into a neighbouring field spends the first fortnight mostly decoding notation, and they know that's what they're doing. The difference between them and a student is not vocabulary, it's that they expect the decoding to take a fortnight and have a procedure for it. This chapter is that procedure.

21.6 Building your own reference

By the end of a degree you'll have met a few hundred symbols. Keeping them costs almost nothing if you do it as you go.

One page per module	At the front of each notebook. Symbol, meaning, where you first met it
A collision page	Symbols that mean different things in different courses. Chapter 16
A read-aloud column	How you say it. This is the column people skip and the one that pays off

Three columns, added to whenever something stops you. In four years it becomes the most valuable document you own, and it's the reason you'll be able to read a paper in an unfamiliar field faster than the person next to you.

PRACTISE THIS

Find one symbol in your current reading that you've been skipping over. Run the procedure: search backwards for its first use, then the shape index in Appendix B, then infer from behaviour. Whatever you find, write it on your notation page with how to say it aloud.

22

Decoding the First Page of a Paper

A notation list is useful until twelve symbols arrive in one paragraph and depend on one another. Then alphabetical lookup is too slow. You need to reconstruct the local system in dependency order.

The extract below is constructed for this exercise, but it has the shape of a modern combinatorics paper: a graph, a family of subsets, a weighted sum, a probability distribution, and a theorem with a parameter range.

22.1 The page as printed

> Let $G = (V, E)$ be a finite graph with maximum degree $\Delta \geq 2$. Write $\mathcal{I}(G)$ for the family of independent sets of G. For $\lambda > 0$, define
>
> $$Z_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|} \quad \text{and} \quad \mu_{G,\lambda}(I) = \frac{\lambda^{|I|}}{Z_G(\lambda)}.$$
>
> For $v \in V$, let $X_v(I) = \mathbf{1}_{\{v \in I\}}$, and put $X = \sum_{v \in V} X_v$. We show that if $0 < \lambda < 1/(\Delta - 1)$, then
>
> $$\mathbb{E}_{\mu_{G,\lambda}} X \leq \frac{\lambda}{1 + \lambda} |V|.$$

Reading left to right produces a queue of questions. What is $\mathcal{I}$? Why a calligraphic I? What is Z? Is μ a number or a measure? What does the bold 1 mean? Which variable is random?

Do not answer in arrival order. Build the dependency chain.

22.2 Layer 1: type every object

Notation	Type	Meaning
$G = (V, E)$	graph	vertices V, edges E
Δ	integer	maximum vertex degree
$\mathcal{I}(G)$	set of sets	all independent vertex sets
λ	positive real	weight parameter
$Z_G(\lambda)$	positive real	total weight of all independent sets
$\mu_{G,\lambda}$	probability distribution	assigns probability to each $I \in \mathcal{I}(G)$
X_v	function / indicator	1 when $v \in I$, otherwise 0
X	random variable	size of the chosen independent set

The type column resolves half the notation before any calculation. $\mathcal{I}(G)$ is a collection whose elements I are themselves sets. $Z_G(\lambda)$ is a number, so it can normalize weights. And $\mu_{G,\lambda}(I)$ is a number between 0 and 1 for each eligible I.

22.3 Layer 2: read the definitions aloud

The partition function is

$$Z_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|}.$$

Say it: "Add λ to the size of I over every independent set I of G."

The subscript G says the graph is fixed. The argument (λ) says the value changes with the weight. The exponent $|I|$ rewards or penalizes larger independent sets depending on whether λ is greater or less than 1.

Now the distribution:

$$\mu_{G,\lambda}(I) = \frac{\lambda^{|I|}}{Z_G(\lambda)}.$$

Say it: "The probability of choosing I is its weight divided by the total weight." Summing this over all I gives 1 because the numerator sums to $Z_G(\lambda)$. The normalization is visible rather than magic.

22.4 Layer 3: make the smallest graph

Take G to be one edge with vertices a and b. Its independent sets are

$$\mathcal{I}(G) = \{\emptyset, \{a\}, \{b\}\}.$$

The set $\{a, b\}$ is excluded because its vertices are adjacent.

So

$$Z_G(\lambda) = 1 + \lambda + \lambda = 1 + 2\lambda.$$

The probabilities are

$$\mu(\emptyset) = \frac{1}{1+2\lambda}, \qquad \mu(\{a\}) = \mu(\{b\}) = \frac{\lambda}{1+2\lambda}.$$

The indicator $X_a(I)$ equals 1 on the outcome $\{a\}$ and 0 on the other two. Likewise for X_b. Hence

$$X = X_a + X_b = |I|.$$

The displayed theorem is bounding the expected size of a randomly chosen independent set under this weighted distribution.

Nothing on the first page said "randomly chosen." The μ definition and expectation symbol supplied it. The toy graph makes that hidden story concrete.

22.5 Layer 4: parse the theorem

Split hypotheses and conclusion.

> **H1** G is finite.
> **H2** Its maximum degree satisfies $\Delta \geq 2$.
> **H3** $0 < \lambda < 1/(\Delta - 1)$.

> **C** The expected independent-set size is at most $\lambda|V|/(1+\lambda)$.

The theorem's scale is now readable. The right side is a fraction of the number of vertices. When λ is small, large independent sets receive little weight, so a small expected size is plausible.

The parameter restriction is not decorative. As Δ grows, the allowed λ shrinks. Dense local neighbourhoods require a smaller weight for the bound claimed by this theorem.

Do not infer more. The theorem gives an upper bound, not an exact formula. It applies in the stated parameter range, not for every positive λ. And it says expectation, not that every independent set has at most that size.

22.6 The notation sheet after one page

The useful sheet is ordered by dependence:

1. graph $G = (V, E)$ and degree bound Δ;
2. independent-set family $\mathcal{I}(G)$;
3. weight $\lambda^{|I|}$;
4. total weight $Z_G(\lambda)$;
5. distribution $\mu_{G,\lambda}$;
6. indicators X_v and total X;
7. expectation under $\mu_{G,\lambda}$.

An alphabetical list would put X before Z and λ before μ. That is fine for lookup and poor for understanding. Dependency order shows what must be known before what.

> *When a paper introduces a local notation system, build it twice: alphabetically for later lookup and in dependency order for the first read. One helps you find a symbol. The other helps you understand the sentence containing it.*

PRACTISE THIS

Take the first notation-heavy page of a paper you need. Type every symbol, then build the smallest object on which all the definitions can be evaluated. Finish by splitting the main displayed claim into hypotheses and conclusion. Do not read page two until page one has a dependency chain.

Part IV

Toolkit

A

The Greek Alphabet, Said Aloud

Every letter, how to say it, and what it conventionally denotes.

The pronunciations given are the ones you'll hear in an English-speaking mathematics department. Modern Greek speakers say several of them differently, and are right, and it doesn't matter.

A.1 Lowercase

Letter	Said	Usually means
α	AL-fa	angle, coefficient, significance level, a root
β	BEE-ta (UK), BAY-ta (US)	angle, regression coefficient, type II error rate
γ	GAM-a	angle, Euler's constant, the Lorentz factor, a curve
δ	DEL-ta	a small quantity, Kronecker delta, Dirac delta, a variation
ϵ, ε	EP-si-lon	a small positive quantity, an error term. The two forms are interchangeable; pick one and stay with it
ζ	ZAY-ta, ZEE-ta	the Riemann zeta function, a complex variable
η	AY-ta, EE-ta	efficiency, viscosity, a learning rate
θ, ϑ	THAY-ta	an angle, a parameter to be estimated
ι	eye-OH-ta	rare. An inclusion map
κ	KAP-a	curvature, a condition number
λ	LAM-da	eigenvalue, wavelength, Poisson rate, Lagrange multiplier
μ	MEW	population mean, a measure, permeability, the prefix micro
ν	NEW	frequency, degrees of freedom, a measure
ξ	KSY, ZY	a dummy variable, a random variable
π	PIE	the constant, a projection, a permutation, a prior, the prime-counting function
ρ, ϱ	ROH	density, correlation, spectral radius
σ, ς	SIG-ma	standard deviation, a permutation, a sigma-algebra, stress

continued on next page

Letter	Said	Usually means
τ	TAW, TOR	a time constant, torque, a topology, Kendall's tau
υ	UP-si-lon	rare, and easily confused with v
ϕ, φ	FY, FEE	an angle, the golden ratio, Euler's totient, a homomorphism, a potential
χ	KY	the chi-squared statistic, a character, Euler characteristic
ψ	SY, PSY	a wavefunction, an angle
ω	oh-MAY-ga	angular frequency, a root of unity, an outcome, a differential form

A.2 Uppercase

Only the letters that don't look like Roman capitals are used, which is why this list is short.

Γ	GAM-a	the gamma function, a curve, a graph
Δ	DEL-ta	a change, the Laplacian, a discriminant, a triangle
Θ	THAY-ta	the tight asymptotic bound
Λ	LAM-da	a diagonal matrix of eigenvalues, the cosmological constant
Ξ	KSY	rare
Π	PIE	product notation
Σ	SIG-ma	summation, a covariance matrix, an alphabet
Υ	UP-si-lon	rare
Φ	FY	the normal cumulative distribution function, a potential
Ψ	SY	a wavefunction
Ω	oh-MAY-ga	sample space, growth lower bound, ohm, solid angle, a region

A.3 The pairs that get confused

ν and v	nu and vee. The Greek has a sharper left stroke. Avoid using both in one document
υ and u	upsilon and you. Almost indistinguishable in print. Do not use upsilon
$\varkappa$ and k	kappa and kay
ρ and p	rho and pee. The descender differs
ς and ζ	final sigma and zeta
ϵ and $\in$	epsilon and "is an element of". **These are different symbols.** $\in$ is a relation and sits between two things; ϵ is a variable
ϕ and $\emptyset$	phi and the empty set. The empty set is a Scandinavian slashed O, not a Greek letter, and calling it "phi" is wrong and extremely common

Those last two are worth internalising. Reading $\in$ as a variable, or calling $\emptyset$ "phi", both mark you as new in a way that's trivially avoidable.

A.4 Two variants of the same letter

Several letters have two printed forms, and the difference carries no meaning:

$$\epsilon\,\varepsilon \qquad \theta\,\vartheta \qquad \phi\,\varphi \qquad \rho\,\varrho \qquad \pi\,\varpi \qquad \sigma\,\varsigma$$

Some authors use both forms in one document for two different quantities. This is legal and unkind. If you do it, say so explicitly.

A.5 Hebrew and other alphabets

$\aleph$	AH-lef	cardinality of infinite sets: $\aleph_0$ is the cardinality of $\mathbb{N}$
$\beth$	BAYT	the beth numbers, in set theory
$\hbar$	aitch-BAR	reduced Planck constant. A modified Roman h, not a Greek letter
ℓ	ELL	a length, a line, a linear functional. Script l, used to avoid confusion with the digit 1
$\wp$	WY-er-strass P	the Weierstrass elliptic function
∂	PAR-shal, DEL, DEE	partial derivative
∇	DEL, NAB-la	the gradient operator

Note that ∂ and ∇ are both commonly called "del" in speech, which is unhelpful and universal. Say "partial" for ∂ if you want to be unambiguous.

B

Look-Up by Shape

For when you have a symbol and no name for it.

Everything else in this book is organised by meaning, which is useless when you can't name the thing you're looking at. This appendix is organised by what the symbol looks like.

Find the shape, get the name, then look the name up in the index or in Part II.

B.1 Arrows

$\rightarrow$	rightarrow	function, tends to
$\mapsto$	mapsto	element sent to
$\Rightarrow$	Rightarrow	implies
$\Leftrightarrow$	Leftrightarrow	if and only if
$\longrightarrow$	longrightarrow	tends to
$\hookrightarrow$	hookrightarrow	injects into
$\twoheadrightarrow$	two-headed arrow	surjects onto
$\rightsquigarrow$	rightsquigarrow	leads to, weak convergence
$\nearrow \searrow$	nearrow, searrow	increases to, decreases to
$\uparrow \downarrow$	uparrow, downarrow	increases to, decreases to
$\rightrightarrows$	rightrightarrows	converges uniformly
$\leftarrow$	leftarrow	assignment, is implied by

B.2 Circles and rounded shapes

$\circ$	circ	composition, degrees
$\bullet$	bullet	a placeholder, an operation
$\oplus$	oplus	direct sum, exclusive or
$\otimes$	otimes	tensor product
$\odot$	odot	element-wise product
$\ominus$	ominus	symmetric difference
$\emptyset$ $\varnothing$	emptyset, varnothing	the empty set
∂	partial	partial derivative, boundary
∞	infty	infinity

B.3 Angular and pointed shapes

$\in$	in	is an element of
$\notin$	notin	is not an element of
$\subset \subseteq$	subset, subseteq	is contained in
$\supset \supseteq$	supset, supseteq	contains
$\cup$	cup	union
$\cap$	cap	intersection
$\land \lor$	land, lor	and, or
$\langle \rangle$	langle, rangle	inner product, average, generated by
∇	nabla	gradient
$\triangle$	triangle	a triangle, a change
$\wedge$	wedge	and, wedge product
$\prec \preceq$	prec, preceq	precedes

B.4 Bars and lines

$\mid$	mid	divides, given, such that
$\lvert x \rvert$	lvert, rvert	absolute value, size, determinant
$\lVert x \rVert$	lVert, rVert	norm
$\parallel$	parallel	is parallel to
$\perp$	perp	perpendicular, orthogonal, independent
$\vdash$	vdash	proves
$\vDash$	vDash	models, satisfies
$\top \bot$	top, bot	true, false
$\dagger$	dagger	conjugate transpose, pseudoinverse
$\prime$	prime	derivative, a different object

B.5 Boxes and blackboard shapes

$\square$	square, Box	end of proof, necessarily
$\blacksquare$	blacksquare	end of proof
$\Diamond$	Diamond	possibly, in modal logic
$\mathbb{R}$	mathbb{R}	blackboard bold: a number system
$\mathcal{F}$	mathcal{F}	script: a family, a sigma-algebra, a functional
$\mathfrak{g}$	mathfrak{g}	fraktur: a Lie algebra, an ideal
tr	mathrm	upright: a named operator

B.6 Large operators

$\sum$	sum	sum
$\prod$	prod	product
$\coprod$	coprod	coproduct, disjoint union
$\int$	int	integral
$\oint$	oint	contour integral
$\bigcup \bigcap$	bigcup, bigcap	union, intersection over a family
$\bigoplus$	bigoplus	direct sum over a family
$\sqrt{}$	sqrt	root

B.7 Relations that look like equals

$=$	=	equals
$\neq$	neq	not equal
$\equiv$	equiv	identically equal, congruent, defined as
$\approx$	approx	approximately
$\simeq$	simeq	isomorphic, approximately
$\cong$	cong	isomorphic, congruent
$\sim$	sim	distributed as, asymptotic, similar
$\propto$	propto	proportional to
$:=$	coloneqq	is defined to be
$\triangleq$	triangleq	is defined to be
$\doteq$	doteq	is defined to be, approaches
$\asymp$	asymp	of the same order
$\leq \geq$	leq, geq	at most, at least
$\ll \gg$	ll, gg	much less, much greater
$\lesssim$	lesssim	at most, up to a constant

B.8 Decorations, by position

$\hat{x}$	hat	unit vector, estimate, operator
$\bar{x}$	bar	mean, conjugate, closure, complement
$\tilde{x}$	tilde	median, transformed version
$\dot{x}$	dot	time derivative
$\vec{x}$	vec	vector
$\underline{x}$	underline	vector, or matrix
$\boldsymbol{x}$	boldsymbol	vector, matrix
$\mathring{x}$	mathring	interior

B.9 If it isn't here

Three tools, in order of speed:

1. **Draw it.** Detexify and similar handwriting-recognition tools take a sketch and return the LaTeX command, which gives you a searchable name.

2. **Copy the character** into any Unicode lookup. Official names like "N-ARY CIRCLED PLUS OPERATOR" are ugly and searchable.
3. **Search the document** for its first occurrence. Chapter 21 covers the full procedure.

C

The Read-Aloud Card

Photocopy this and keep it in the front of a notebook.

Say each line out loud once. That's the whole exercise, and it works.

C.1 Powers, indices, and decorations

x^2, x^3	x squared, x cubed
x^n, x^{n+1}	x to the n; x to the n plus one
x^{-1}	x inverse, or one over x
x_n, x_{n+1}	x sub n; x sub n plus one
a_{ij}	a i j
$f^{(3)}$	the third derivative of f
$\bar{x}$, $\hat{x}$, $\tilde{x}$	x bar, x hat, x tilde
$\dot{x}$, $\ddot{x}$	x dot, x double dot
$\vec{v}$, $\boldsymbol{v}$	vector v
$A^\top$, A^*, $A^\dagger$	A transpose, A star, A dagger

C.2 Sizes and brackets

$\lvert x\rvert$	mod x, the absolute value of x
$\lVert v\rVert$	the norm of v
$\lfloor x\rfloor$, $\lceil x\rceil$	the floor of x, the ceiling of x
$\lvert A\rvert$	the size of A, the cardinality of A
$\det A$	the determinant of A
$n!$	n factorial
$\binom{n}{k}$	n choose k
$\langle u, v\rangle$	the inner product of u and v
$[a, b]$, (a, b)	the closed interval a to b; the open interval

C.3 Sets

$x \in A$	x is in A, x is an element of A
$x \notin A$	x is not in A
$A \subseteq B$	A is a subset of B, A is contained in B
$A \subsetneq B$	A is a proper subset of B
$A \cup B$, $A \cap B$	A union B; A intersect B
$A \setminus B$	A minus B, A without B
$A \times B$	A cross B, the Cartesian product
$\mathcal{P}(A)$	the power set of A
$\emptyset$	the empty set
$\{x \in A : P(x)\}$	the set of x in A such that P of x
$\bigcup_{i\in I} A_i$	the union over i in I of A sub i

C.4 Logic

$\neg P$	not P
$P \wedge Q, P \vee Q$	P and Q; P or Q
$P \Rightarrow Q$	P implies Q; if P then Q
$P \Leftrightarrow Q$	P if and only if Q
$\forall x$	for all x, for every x
$\exists x$	there exists an x
$\exists! x$	there exists a unique x
$x := y$	x is defined to be y
$\therefore$	therefore
$\blacksquare$	which completes the proof

C.5 Functions

$f(x)$	f of x
$f : A \to B$	f from A to B, f maps A into B
$x \longmapsto x^2$	x maps to x squared
$g \circ f$	g composed with f, g after f
f^{-1}	f inverse
$f\|_S$	f restricted to S
$\ker f$, $\operatorname{im} f$	the kernel of f, the image of f
$A \hookrightarrow B$	A embeds in B
$A \twoheadrightarrow B$	A surjects onto B

C.6 Calculus

$f'(x), f''(x)$	f prime of x; f double prime of x
$\frac{dy}{dx}$	dy by dx
$\frac{\partial f}{\partial x}$	partial f by partial x
$\int_a^b f(x)\,dx$	the integral from a to b of f of x, dx
$\oint$	the contour integral
$\sum_{k=1}^n a_k$	the sum from k equals one to n of a sub k
$\prod_{k=1}^n a_k$	the product from k equals one to n
$\lim_{x\to 0}$	the limit as x tends to zero
$\lim_{x\to a^+}$	the limit as x tends to a from above
$\limsup$, $\liminf$	lim sup, lim inf
$\sup S$, $\inf S$	the supremum of S, the infimum of S
∇f	grad f
$\nabla \cdot v$, $\nabla \times v$	div v, curl v
$\nabla^2 f$	the Laplacian of f

C.7 Growth

$O(n^2)$	big oh of n squared
$o(n)$	little oh of n
$\Theta(n \log n)$	big theta of $n \log n$
$\Omega(n)$	big omega of n
$f \sim g$	f is asymptotically equal to g
$f \asymp g$	f is of the same order as g

C.8 Probability and statistics

$\Pr(A)$	the probability of A
$\Pr(A \mid B)$	the probability of A given B
$\mathbb{E}[X]$	the expectation of X, the expected value of X
$\mathrm{Var}(X)$	the variance of X
$X \sim N(\mu, \sigma^2)$	X is distributed as normal, mu, sigma squared
$X \perp\!\!\!\perp Y$	X is independent of Y
$\hat{\theta}$	theta hat, the estimate of theta
H_0, H_1	H nought, H one

C.9 Number theory

$a \mid b$	a divides b
$a \nmid b$	a does not divide b
$a \equiv b \pmod{n}$	a is congruent to b mod n
$\gcd(a, b)$	the gcd of a and b
$\mathbb{N}, \mathbb{Z}$	the naturals, the integers
$\mathbb{Q}, \mathbb{R}, \mathbb{C}$	the rationals, the reals, the complex numbers
$\aleph_0$	aleph nought

C.10 Whole expressions, for practice

Say these five out loud, from cold, without decoding first. If you can, you're fluent.

1. $\forall \varepsilon > 0\ \exists N\ \forall n > N : |a_n - L| < \varepsilon$
2. $\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$
3. $\{ x \in \mathbb{R} : f(x) > 0 \}$
4. $\mathbb{E}[X] = \sum_k k \Pr(X = k)$
5. $\bigcup_{n=1}^{\infty} \left[0, 1 - \frac{1}{n}\right] = [0, 1)$

D

The Collision Table

Every symbol in this book that means more than one thing, sorted, with the field that picks the meaning.

Use it when you have a symbol, a context, and a suspicion that it isn't saying what you think.

D.1 Symbols

Symbol	Meaning	Field
$\mid$	divides	number theory
	given	probability
	such that	set-builder
	restricted to	analysis
$\lvert\ \rvert$	absolute value, modulus	everywhere
	cardinality	set theory
	determinant	linear algebra
	order of a group	algebra
$\sim$	is distributed as	statistics
	asymptotically equal	analysis
	same order of magnitude	physics
	is similar to	geometry
	an equivalence relation	algebra
	negation	older logic
$\equiv$	is defined as	physics, logic
	identically equal	analysis
	congruent mod n	number theory
	logically equivalent	logic
$\subset$	subset, equality allowed	most modern texts
	proper subset	older, continental texts
log	base e	pure maths, statistics
	base 10	engineering, chemistry
	base 2	computer science
(a, b)	ordered pair	everywhere
	open interval	analysis
	gcd	number theory

continued on next page

Symbol	Meaning	Field
	inner product	some linear algebra
$\times$	multiplication	arithmetic
	cross product	vector calculus
	Cartesian product	set theory
$\cdot$	multiplication	arithmetic
	dot product	vector calculus
	composition	algebra
	a placeholder argument	analysis
$\oplus$	direct sum	linear algebra
	exclusive or	logic, computing
$\otimes$	tensor product	algebra, physics
	Kronecker product	numerical work
$*$	convolution	analysis, signals
	adjoint, conjugate transpose	linear algebra
	complex conjugate	physics
	dual space	functional analysis
	Kleene star	formal languages
	the optimal value	optimisation
$\dagger$	conjugate transpose	physics
	pseudoinverse	numerical linear algebra
	adjoint	functional analysis
Ω	sample space	probability
	growth lower bound	algorithms
	ohm	electronics
	solid angle	physics
	a spatial region	PDEs
Δ	a finite change	everywhere
	the Laplacian	analysis, PDEs
	a discriminant	algebra
	a triangle	geometry
δ	Kronecker delta	linear algebra
	Dirac delta	physics, analysis
	a small change	calculus
	a variation	calculus of variations
λ	eigenvalue	linear algebra
	Poisson rate	probability
	wavelength	physics
	Lagrange multiplier	optimisation
	lambda abstraction	logic, CS
σ	standard deviation	statistics
	a permutation	algebra
	a sigma-algebra	measure theory
	stress	mechanics
ϕ, φ	the golden ratio	number theory
	Euler's totient	number theory

continued on next page

Symbol	Meaning	Field
	an angle	geometry
	a homomorphism	algebra
	a potential	physics
e	Euler's number	analysis
	the group identity	algebra
	a basis vector	linear algebra
	electron charge	physics
i, j	an index	everywhere
	$\sqrt{-1}$	maths uses i, EE uses j
β	regression coefficient	statistics
	type II error rate	statistics
	an angle	geometry
$'$	derivative	calculus
	a different object	everywhere
	transpose	older statistics
	complement	set theory
	minutes of arc	astronomy
$\hat{}$	unit vector	physics, geometry
	an estimate	statistics
	an operator	quantum mechanics
	Fourier transform	analysis
$\bar{}$	the mean	statistics
	complex conjugate	complex analysis
	closure	topology
	complement	set theory
	negation	logic, circuits
$\square$	end of proof	everywhere
	necessarily	modal logic
	the d'Alembertian	relativity
$\perp$	perpendicular	geometry
	orthogonal	linear algebra
	independent	statistics
	the least element	order theory
	contradiction	logic
f^n	the nth power	analysis
	the n-fold composition	dynamical systems
$\Rightarrow$	implies	logic
	converges in distribution	probability

D.2 The resolution procedure

Repeated from Chapter 16 because this is the page you'll have open.

1. **What subject is this?** Settles most cases on its own.
2. **What are the operands?** A bar around a set is cardinality; around a matrix, a determinant.

3. **Check the notation section.** Front matter or appendix.
4. **Try both and see which makes the theorem true.**

D.3 The four you should never write

Because a clear alternative exists and costs nothing.

Avoid	**Write instead**	**Reason**
$\subset$	$\subseteq$ or $\subsetneq$	two readings, no way to tell
$\lvert A\rvert$ for det	$\det A$	four other meanings
A' for transpose	A^{T}	primes mean derivatives
log where base matters	ln, $\log_2$, $\log_{10}$	three communities disagree

E

$\LaTeX$ for Every Symbol Here

You will have to write these as well as read them.

Commands are given without the backslash in the tables; add it. Anything marked amsmath needs `\usepackage{amsmath}`, which you should be loading anyway.

E.1 Logic

$\forall$	forall	$\exists$	exists
$\nexists$	nexists	$\neg$	neg, lnot
$\land$	land, wedge	$\lor$	lor, vee
$\Rightarrow$	Rightarrow	$\Leftrightarrow$	Leftrightarrow
$\implies$	implies	$\iff$	iff
$\vdash$	vdash	$\vDash$	vDash
$\therefore$	therefore	$\because$	because
$\blacksquare$	blacksquare	$\square$	square, Box
$\top$	top	$\bot$	bot
$\coloneqq$	coloneqq	$\triangleq$	triangleq

E.2 Sets

$\in$	in	$\notin$	notin
$\subset$	subset	$\subseteq$	subseteq
$\subsetneq$	subsetneq	$\supseteq$	supseteq
$\cup$	cup	$\cap$	cap
$\bigcup$	bigcup	$\bigcap$	bigcap
$\setminus$	setminus	$\sqcup$	sqcup
$\emptyset$	emptyset	$\varnothing$	varnothing
$\mathcal{P}$	mathcal{P}	$\complement$	complement
$\times$	times	$\mathbb{N}$	mathbb{N}

E.3 Relations

$\leq$	leq	$\geq$	geq
$\neq$	neq	$\equiv$	equiv
$\approx$	approx	$\simeq$	simeq
$\cong$	cong	$\sim$	sim
$\propto$	propto	$\asymp$	asymp
$\ll$	ll	$\gg$	gg
$\lesssim$	lesssim	$\gtrsim$	gtrsim
$\preceq$	preceq	$\sqsubseteq$	sqsubseteq
$\perp$	perp	$\parallel$	parallel
$\mid$	mid	$\nmid$	nmid

E.4 Arrows

$\to$	to, rightarrow	$\longmapsto$	mapsto
$\longrightarrow$	longrightarrow	$\hookrightarrow$	hookrightarrow
$\twoheadrightarrow$	twoheadrightarrow	$\rightsquigarrow$	rightsquigarrow
$\nearrow$	nearrow	$\searrow$	searrow
$\uparrow$	uparrow	$\downarrow$	downarrow
$\xrightarrow{d}$	xrightarrow{d}	amsmath	

E.5 Big operators and calculus

$\sum$	sum	$\prod$	prod
$\coprod$	coprod	$\int$	int
$\iint$	iint	$\oint$	oint
∂	partial	∇	nabla
$\lim$	lim	$\limsup$	limsup
$\sup$	sup	$\inf$	inf
∞	infty	$\sqrt{x}$	sqrt{x}
$\binom{n}{k}$	binom{n}{k}	amsmath	

E.6 Decorations

$\hat{x}$	hat{x}	$\bar{x}$	bar{x}
$\tilde{x}$	tilde{x}	$\dot{x}$	dot{x}
$\ddot{x}$	ddot{x}	$\vec{x}$	vec{x}
$\boldsymbol{x}$	boldsymbol{x}	$\underline{x}$	underline{x}
$\widehat{AB}$	widehat{AB}	$\overline{AB}$	overline{AB}
$\hat{\imath}$	hat{\imath}	dotless i, for hats	

Use `\widehat` and `\overline` for anything wider than one character; the short forms don't stretch.

E.7 Delimiters that grow

`\left( ...\right)`	Auto-sizing brackets
`\bigl( \Bigl( \biggl(`	Manual sizes, when auto gets it wrong
`\lvert ...\rvert`	Absolute value. Better than bare `|`, which does not size
`\lVert ...\rVert`	Norm
`\lfloor ...\rfloor`	Floor
`\langle ...\rangle`	Angle brackets. Not < and >

E.8 Named operators

Built in: `\sin`, `\cos`, `\tan`, `\log`, `\ln`, `\exp`, `\max`, `\min`, `\det`, `\dim`, `\gcd`, `\deg`, `\ker`, `\Pr`.

Not built in, and you will want them:

tr, rank, im, Var	`\operatorname{tr}` etc. amsmath
Declare once	`\DeclareMathOperator{\tr} {tr}` in the preamble, then use `\tr`

Never write `$tr(A)$`. It typesets as a product of three variables, in italics, with the wrong spacing.

E.9 Spacing in maths mode

`\,`	thin space. Use before dx in an integral
`\;`	medium space
`\quad`	one em
`\qquad`	two ems
`\!`	negative thin space

$\int f(x)\,dx$ with the thin space reads better than $\int f(x)dx$ without it. This is a small thing that makes a document look professionally set.

E.10 Alphabets

$\mathbb{R}$	mathbb{R}	blackboard bold
$\mathcal{F}$	mathcal{F}	calligraphic, capitals only
$\mathfrak{g}$	mathfrak{g}	fraktur
d	mathrm{d}	upright roman
T	mathsf{T}	sans serif, for transpose
$\boldsymbol{\alpha}$	boldsymbol{ \alpha}	bold, works for Greek

E.11 Four mistakes worth not making

1. **Bare `|` for absolute value.** It doesn't grow with its contents. Use `\lvert` and `\rvert`.
2. **< and > for angle brackets.** Use `\langle` and `\rangle`.
3. **Multi-letter names in italics.** `$Var$` is a product. Use `\operatorname`.
4. **`$$ ...$$` for displays.** It's plain TEX and breaks spacing. Use `\[ ...\]`.

F

Conventions by Field

What each subject does differently, on one page each.

Read the entry for whatever you're about to start. Ten minutes, and it saves the fortnight of decoding described in Chapter 21.

F.1 Pure mathematics

log	base e
$\subset$	usually allows equality
$\mathbb{N}$	varies; analysis often excludes 0, set theory includes it
Vectors	plain italic, since everything is a vector
$\sim$	an equivalence relation, or asymptotic equivalence
Proofs	prose, not symbols. $\Rightarrow$ is rare in print
i	$\sqrt{-1}$
Style	definitions before use, always; notation sections common

F.2 Analysis specifically

ε, δ	small positive quantities, essentially reserved
Δ	the Laplacian, in PDE work
$\bar{A}$	the closure
$\|\cdot\|$	norms everywhere, with subscripts naming which
L^p, C^k	function spaces. Superscripts are labels, not powers
O, o	error terms, usually as $x \to 0$

F.3 Statistics

$\sim$	**is distributed as**. Not "approximately"
Greek	population parameters, unknown
Roman and hats	sample estimates, computed
log	base e
β	regression coefficient, and also the type II error rate
$N(\mu, \sigma^2)$	second parameter usually variance, sometimes standard deviation
$\vert$	given
Style	notation is relaxed and inconsistent. Read for meaning

F.4 Computer science

log	base 2
$\mathbb{N}$	includes 0
O, Θ, Ω	running time; O is often said when Θ is meant
$\rightarrow$	often written for implication, instead of $\Rightarrow$
λ	lambda abstraction
$*$	Kleene star: zero or more
$:=$	assignment, as well as definition
$A[i][j]$	indexing, and beware: some libraries index column-first

F.5 Physics

$\hat{A}$	an operator, or a unit vector
A^*	complex conjugate only
$A^\dagger$	conjugate transpose
Δ	a change, almost never the Laplacian
∇^2	the Laplacian
$\langle\cdot\rangle$	a time or ensemble average
$\equiv$	a definition
$\sim$	order of magnitude
dy/dx	manipulated freely as a fraction
Einstein summation	assumed in relativity; sums are never written
Units	upright roman, with a space before them

F.6 Engineering

log	base 10
j	$\sqrt{-1}$, in electrical engineering, since i is current
Matrices	square brackets
$A^{\top}$	transpose, often written A^t
$\lVert\cdot\rVert$	norms, with 2-norm the default
$*$	convolution, heavily used in signals
Units	mandatory, and marked in reports
Significant figures	taken seriously; matching precision to measurement is expected

F.7 Probability specifically

Ω	the sample space
$\mathcal{F}$	a sigma-algebra
$\mathbb{P}$, Pr, P	all the same
$\Rightarrow$	sometimes convergence in distribution, which collides with implication
$X \perp\!\!\!\perp Y$	independence
a.s.	almost surely: with probability one
Capitals	random variables; lowercase for their values

That last row is a real convention and worth knowing: $\Pr(X = x)$ has an uppercase random variable and a lowercase value, and mixing them makes the expression meaningless.

F.8 Abstract algebra

e	the identity element
$\lvert G\rvert$	the order of a group
$H \leq G$	H is a subgroup, not a numeric inequality
$H \trianglelefteq G$	H is a normal subgroup
G/H	a quotient group
$\langle S\rangle$	the subgroup generated by S
$\cong$	isomorphic
$\mathfrak{g}$, $\mathfrak{p}$	fraktur for Lie algebras and ideals

F.9 Topology

$\bar{A}$	the closure
A°, $\mathring{A}$	the interior
∂A	the boundary, not a derivative
$f^{-1}(U)$	the preimage. Continuity is defined by preimages, and no inverse function is implied
$\simeq$	homotopy equivalent, weaker than homeomorphic
$\cong$	homeomorphic

F.10 The three questions to ask in week one

Whatever the course, these three answers save the most trouble.

1. **Does $\mathbb{N}$ include zero here?**
2. **Does $\subset$ allow equality here?**
3. **What base is** $\log$ **here?**

Ask them in the first tutorial. They take thirty seconds and no lecturer has ever thought worse of a student for asking.

G

Field-First Collision Table

Appendix D starts from a symbol. This table starts from the subject you are entering and names the inherited meanings most likely to mislead you.

G.1 Analysis and topology

Mark	Read here	Do not import blindly
$\bar{A}$	closure	complex conjugate or average
A°	interior	composition or a degree mark
∂A	boundary	partial derivative
$f^{-1}(U)$	preimage	inverse function
$\sim$	asymptotic equivalence	rough approximation
Δ	Laplacian	finite change

G.2 Algebra and category theory

Mark	Read here	Do not import blindly
e	identity element	Euler's number
$\lvert G\rvert$	group order	absolute value
$H \leq G$	subgroup	numeric inequality
G/H	quotient	ordinary division
$\mathcal{C}^{op}$	opposite category	inverse
$\eta : F \Rightarrow G$	natural transformation	implication
$F \dashv G$	adjunction	perpendicularity

G.3 Combinatorics and graph theory

Mark	Read here	Do not import blindly
$[n]$	$\{1, \dots, n\}$	interval
$[x^n]F(x)$	coefficient extractor	evaluation
$\lambda \vdash n$	integer partition	divisibility
$N(v)$	neighbourhood	number-valued function
$d(v)$	degree	derivative or distance
$G[S]$	induced subgraph	array lookup

G.4 Probability and statistics

Mark	Read here	Do not import blindly
$X \sim D$	distributed as	approximately equal
$\hat{\theta}$	estimator	unit vector or operator
$\bar{X}$	sample mean	closure or conjugate
$P(A \mid B)$	conditional probability	divisibility
Ω	sample space	ohms or solid angle
$\mathbf{1}_A$	indicator	identity matrix

G.5 Computer science and formal languages

Mark	Read here	Do not import blindly
Σ	alphabet	sum
ε	empty string	small positive real
Σ^*	all finite words	conjugate or dual
uv	concatenation	multiplication
$u \Rightarrow_G v$	derivation	logical implication
$\delta(q, a)$	transition function	small change
$O(n)$	growth bound	exact running time

G.6 Physics and engineering

Mark	Read here	Do not import blindly
$\hat{A}$	operator or unit vector	estimator
j	imaginary unit in electrical work	index
Δ	finite change	Laplacian
$\langle A \rangle$	average or expectation	span
$A^\dagger$	conjugate transpose	pseudoinverse
x^μ	tensor component	power
$\sim$	order of magnitude	ratio tends to one

G.7 Order theory

Mark	Read here	Do not import blindly
$x \preceq y$	partial-order relation	numeric comparison
$x \parallel y$	incomparable	parallel geometry
$x \wedge y$	meet	logical conjunction
$x \vee y$	join	logical disjunction
$\downarrow x$	principal down-set	convergence direction
$x \lessdot y$	cover relation	approximation

G.8 How to use the table

Read the six or seven rows for the field before the first chapter or paper. Circle any mark whose old meaning is automatic for you. Those are the collisions that need an explicit note the first three times they appear.

The table does not decide the meaning. The local definition does. Its job is to warn you which old answer is most likely to arrive before you look.

Index

A Note from the Author

Thank you for reading. Books like this one travel by word of mouth among students, so if it saved you an afternoon of staring at a symbol, two minutes of your help matter more than you'd think: leave a short, honest review wherever you bought it, or lend your copy to whoever in your tutorial group is too embarrassed to ask.

The appendices are working tools rather than reading. The read-aloud card, the look-up-by-shape index, and the collision table are meant to be photocopied and kept in the front of a notebook. You'll find printable companions and more of my writing for students at **gauravtiwari.org**, where corrections and questions are welcome.

And if you meet a symbol this book should have covered and doesn't, please tell me. A field guide gets better by being used against the field, and the next edition will owe its best entries to those messages.

Also by Gaurav Tiwari

How to Read Mathematics
Definitions, Theorems, Proofs, and Papers for Students

The natural next book after this one. A six-step drill for opening any definition, three ways to read a proof, a stuck protocol that always ends in an action, and honest reading speeds for research papers.

How to Write Mathematics
Proofs, Papers, and Problem Sets for Students

Writing solutions that collect the marks the work deserved: proof makeovers, full-credit problem sets, and the grader's-eye view of partial credit.

How to Study STEM
Problem Sets, Practicals, and Exams Without the All-Nighter

The week-by-week system underneath everything else: how to work a problem set, what to do when you're stuck, and how to keep a term from collapsing into the fortnight before finals.

The STEM Exam Handbook
Timing, Triage, and Partial Credit in Maths, Physics, and Engineering Exams

The four-minute triage, how method marks really work, salvage writing, and ten annotated scripts where every lost mark is traced to its cause.

How to Write Physics and Engineering
Lab Reports, Derivations, Design Documents, and Your Final-Year Project

The report side of a STEM degree, from the four-sentence abstract to the final-year thesis, including uncertainty written as prose rather than arithmetic.

Your First Research Paper
From Project to Submission, Peer Review, and Publication

What happens after the project works: choosing a venue, stating a claim, authorship, peer review, and what to do with a rejection.

Other titles

- *LaTeX for Students*
- *LaTeX for Theses and Dissertations*
- *The Student Research Workflow*
- *Indian Polity: Complete Notes for UPSC Prelims & Mains*

All titles are available on Amazon. Find the full, current catalog at **gauravtiwari.org**.