How to Work Out Proofs in Analysis I

by DiBeos

1. The let move
2. The naming move
3. Substitution into a target
4. Expansion
5. Substitution into a hypothesis
6. Modus ponens

Let's say I wanted to ask you to memorize this sequence 765432154321.

Would you have to memorize the entire string of 12 numbers? Well you could do that, but the easier and probably better thing to do is to remember that you're just counting down from 7 up to 1 and then from 5 to 1.

7 6 5 4 3 2 1 5 4 3 2 1

What you want when it comes to proofs is the same kind of intuition – you just keep on doing the obvious thing. It just so happens that sometimes the next step isn't obvious, and you need to remember it. Once you do that though, you can now better understand how come this step, which wasn't obvious to you before, was actually a very reasonable option, and the more you develop this feeling about the step the easier it will become to memorize it.

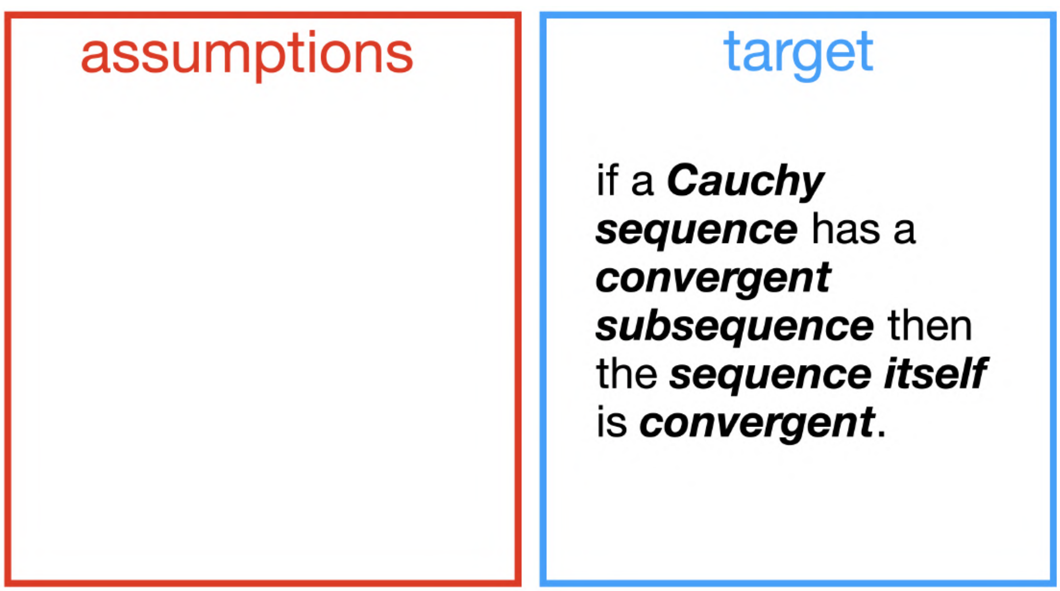
Though the proofs may seem hard at first glance, they're actually pretty easy, in the sense that you just do the obvious thing repeatedly, and you end up solving them.

The basic idea of this method is that you need to teach yourself a set of specific moves, and apply them at an appropriate opportunity. This is easier said than done, but it gives you a rough idea of what the process should look like.

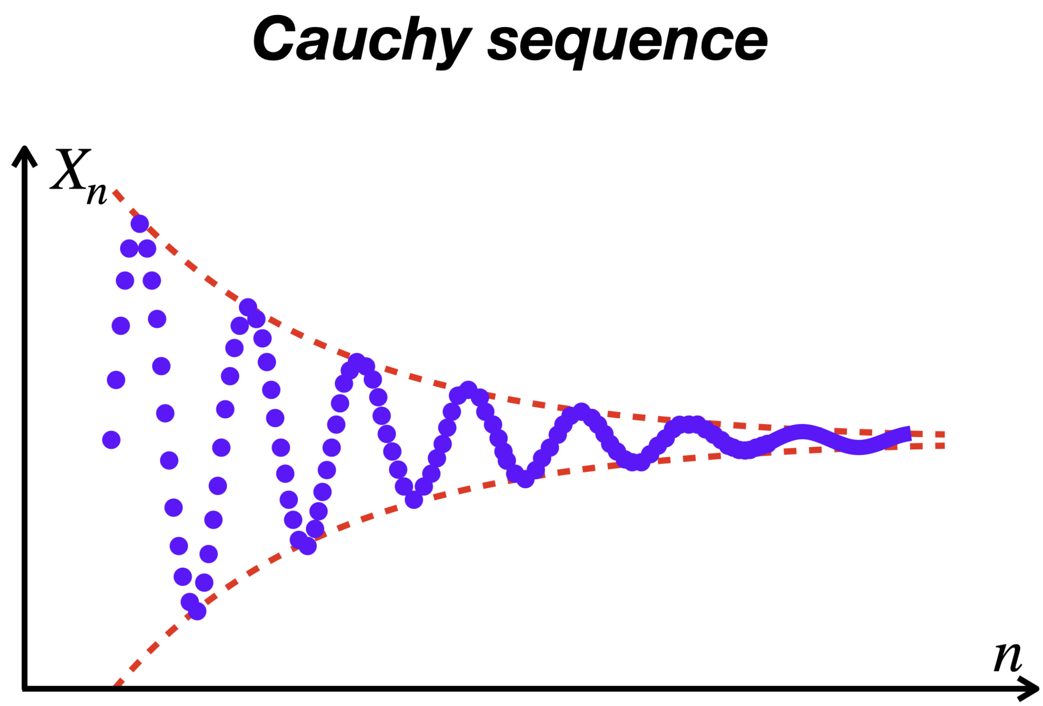
This technique was put together by mathematician *Timothy Gower*. It simply asks you to at any moment during your work, write down what you know (like your assumptions or likely tools) and what you are trying to prove (so your target). These aren't fixed, they change as you get new insights.

if a ***Cauchy sequence*** has a ***convergent subsequence*** then the ***sequence itself*** is ***convergent***.

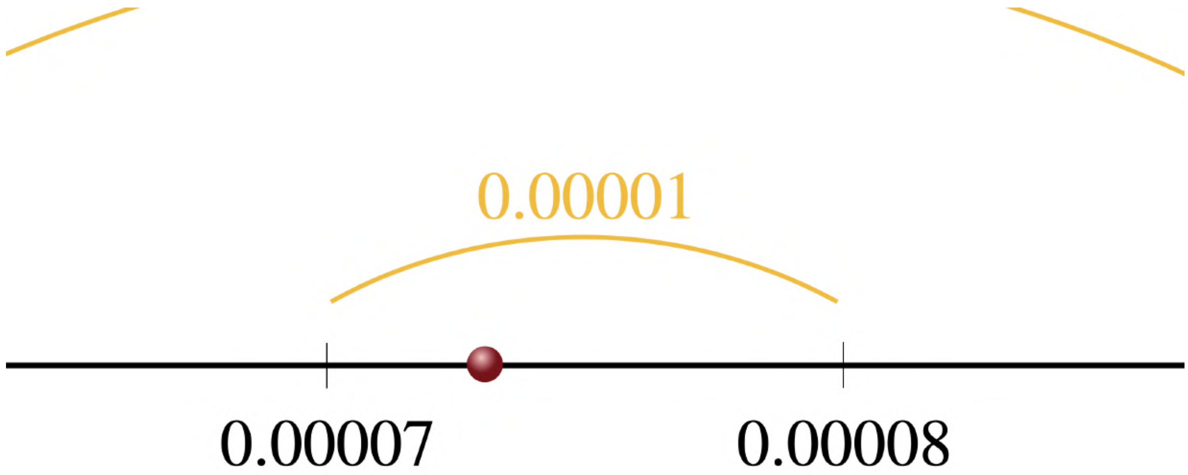
Let's prove that if a Cauchy sequence has a convergent subsequence then the sequence itself is convergent. As of this moment right now, we have nothing that we might want to use, but we do have our target, the statement we want to prove.



Before delving into the technique of the proof, you obviously need to make sure that you know what it is that you're proving. A Cauchy sequence is a sequence where the numbers get closer and closer to each other as the sequence goes on.



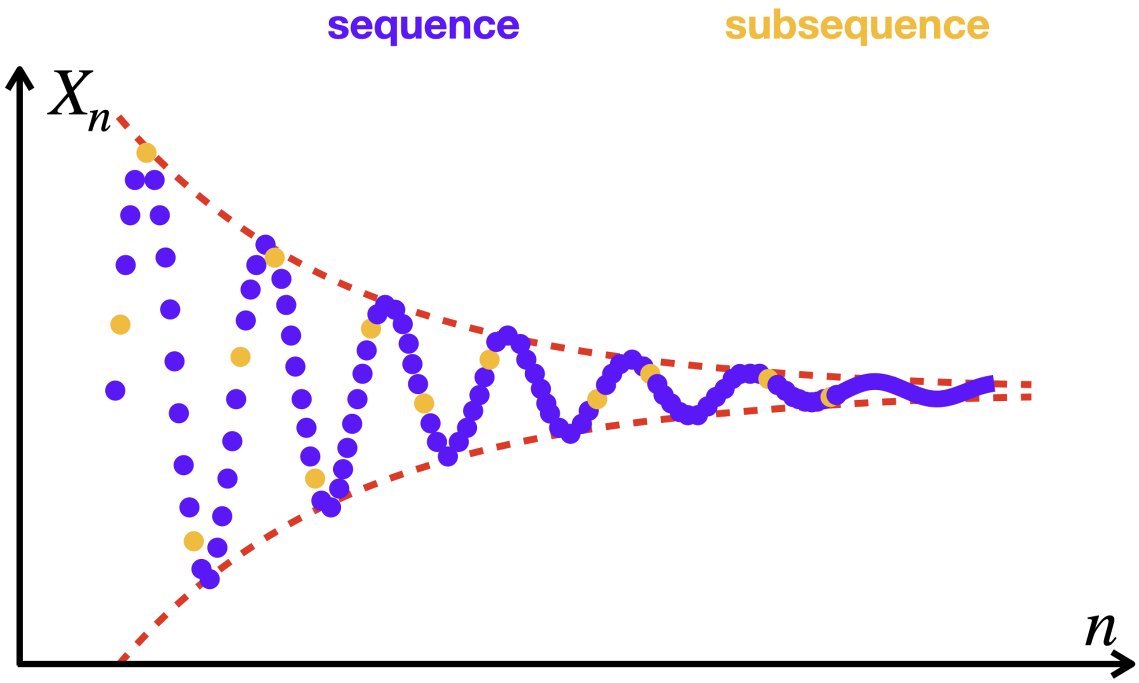
More precisely, no matter how small a distance you choose, there comes a point in the sequence after which all the numbers are within that distance of each other. So, eventually, the sequence settles down in a tight cluster, even if we don't know exactly where it's going.



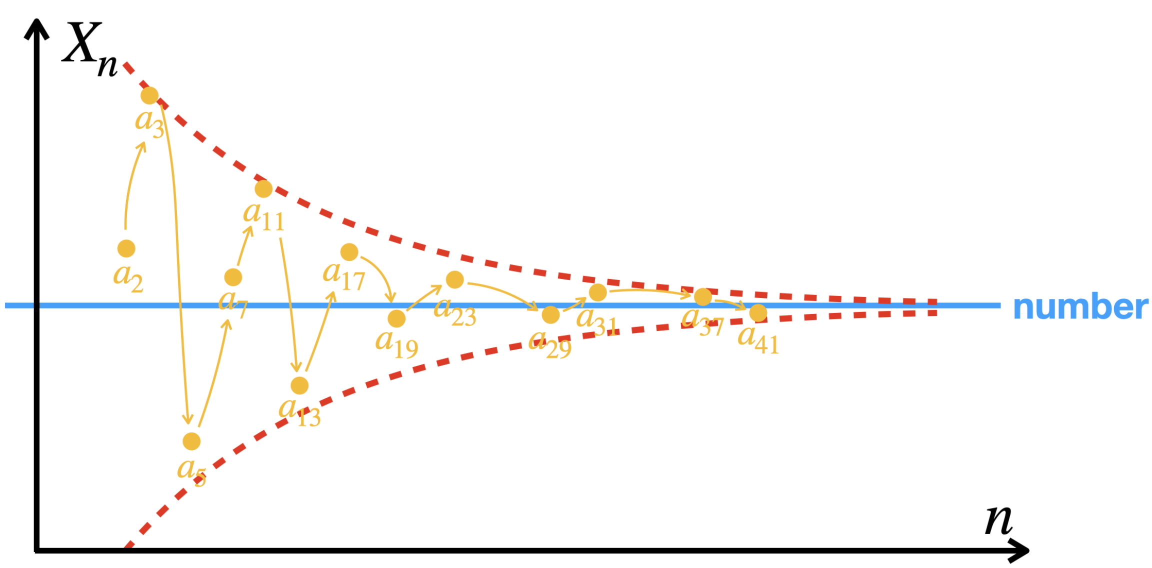
Rigorously, the definition of a Cauchy sequence is expressed as this

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \text{ such that } |a_n - a_m| < \varepsilon \text{ for all } n, m > N$$

You can think of it like that: the sequence may not yet reach a final value, but past a certain point, a sort of point of no return, all the values stick close together.

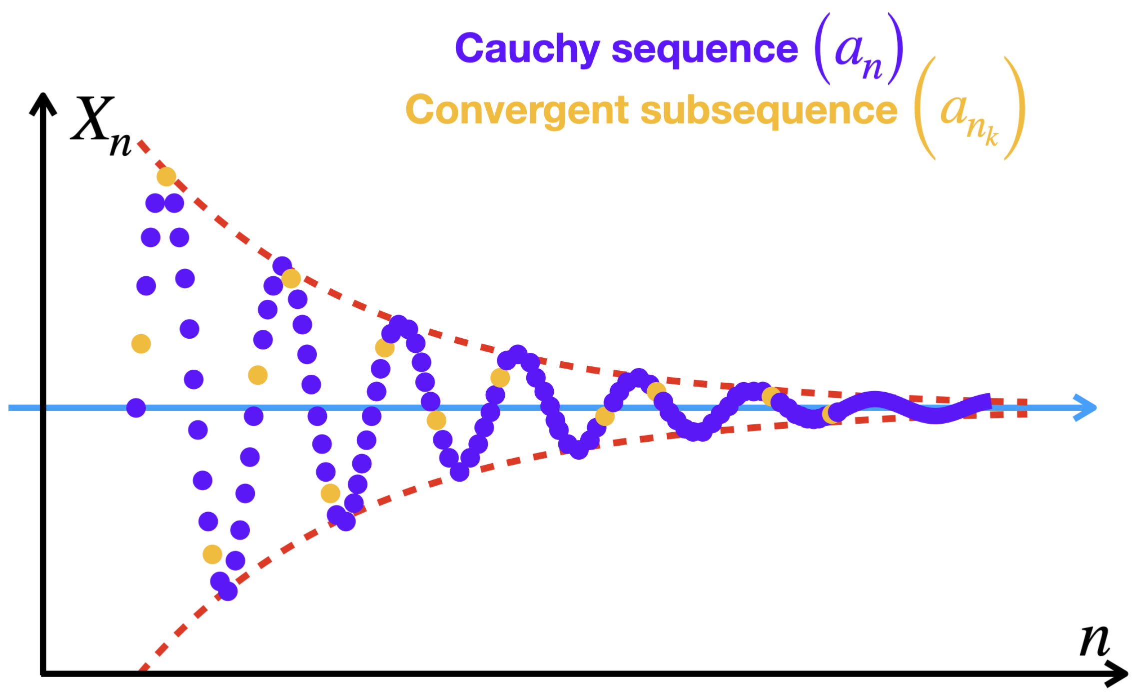


Now, referring back to our initial statement, what does it mean for a Cauchy sequence to have a convergent subsequence?



A subsequence is just picking out some of the terms in the sequence – like every second term, or every term where the index is prime, for example. A convergent subsequence is when this smaller, picked-out list does settle down and approaches a definite number.

So in other words, some part of the sequence is not just bunched together, but is heading to a specific destination.



Rigorously speaking, a subsequence is a selection of terms in the sequence an, say a_{n_k}, where $n_1 < n_2 < n_3 < \ ...$

It converges if it approaches some limit L:

$$\exists L \in \mathbb{R} \text{ such that } a_{n_k} \to L$$

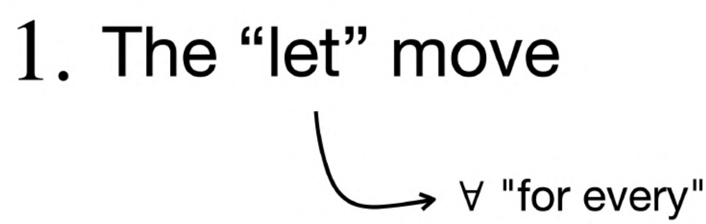
So, in the case of what we are trying to prove here, namely that "Every Cauchy sequence with a convergent subsequence converges", we have that:

If the whole sequence is internally stable (Cauchy), and some part of it knows where it's going (a convergent subsequence), then the whole sequence must actually be heading to that same place.

If we were to write this statement that we want to prove in more mathematical language we would write this:

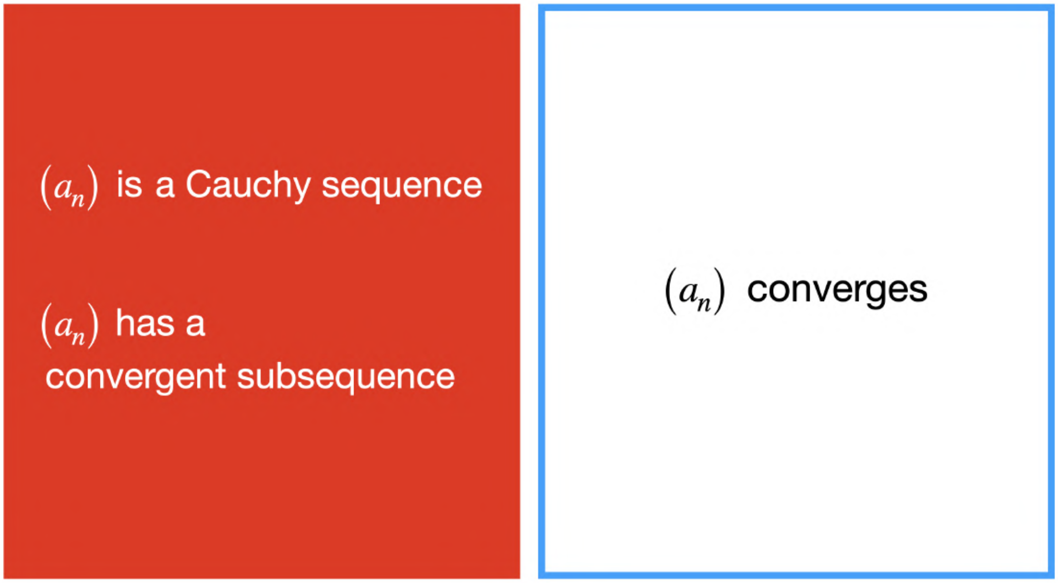
$\forall (a_n), (a_n)$ is Cauchy and (a_n) has a convergent subsequence $\implies$ (a_n) converges.

Remember, we need to distinguish what we can assume, and what we are trying to prove. This helps you keep track of where you are in the proof and encourages you to ask: Given what I know, what can I logically deduce?



Now we're going to apply something called the "let" move, which is a classic move for kind of "for every" (∀) statements.

The idea is: instead of reasoning about all sequences at once, just focus on one arbitrary sequence that satisfies the assumptions. So, we instead change it to: let a_n be a Cauchy sequence that has a convergent subsequence. This is now a working assumption – something we're allowed to use.

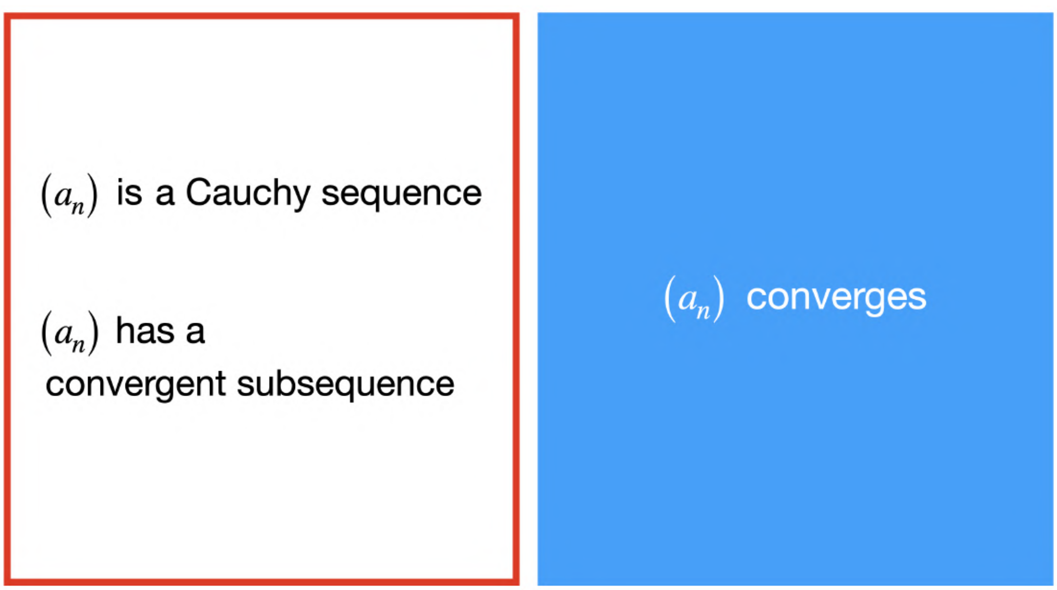


We therefore change our target to that of proving that a_n converges. We now represent what we've got as this:

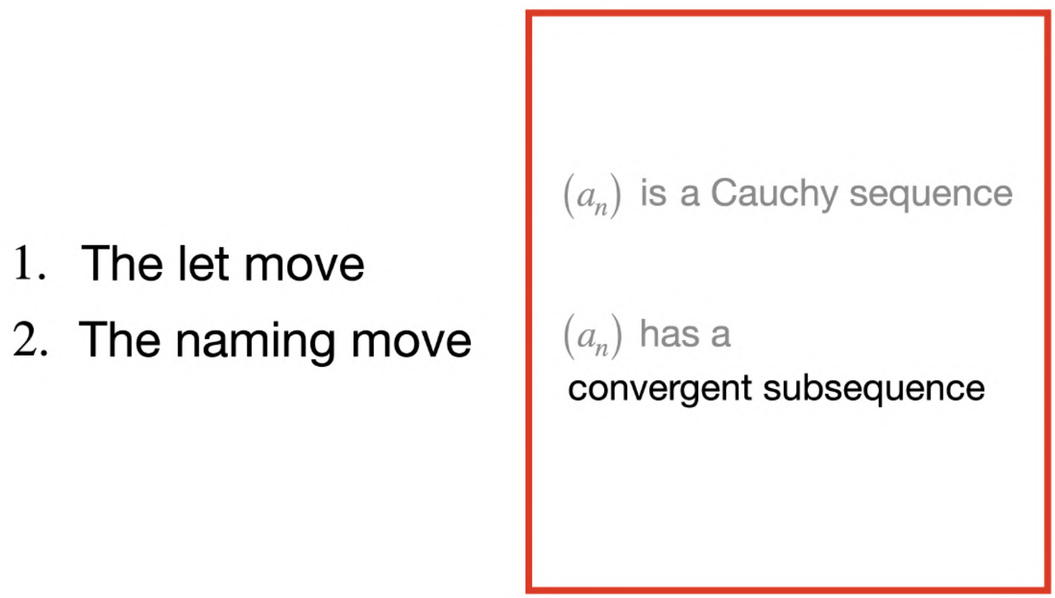
(a_n) is a Cauchy sequence.

(a_n) has a convergent subsequence.

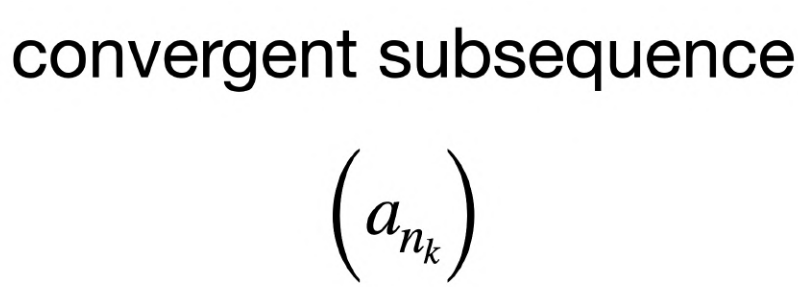
These are our assumptions, however what we don't yet know is whether a_n converges.



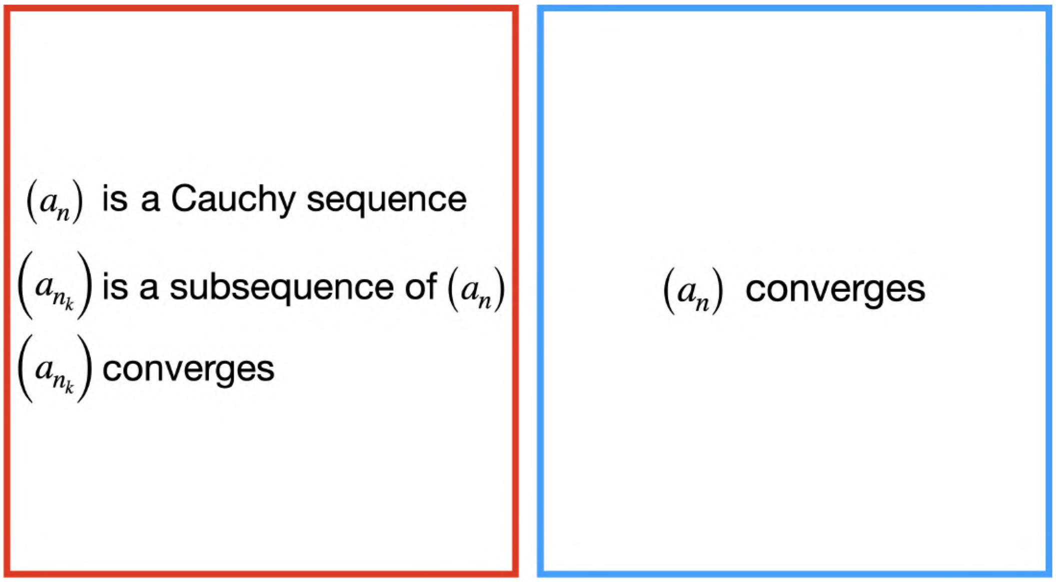
The statements that we can assume are all on the left side, while the statements we are trying to prove are on the right side. Basically, just ask yourself: given what's on the left side, what can I do to get what's on the right side?



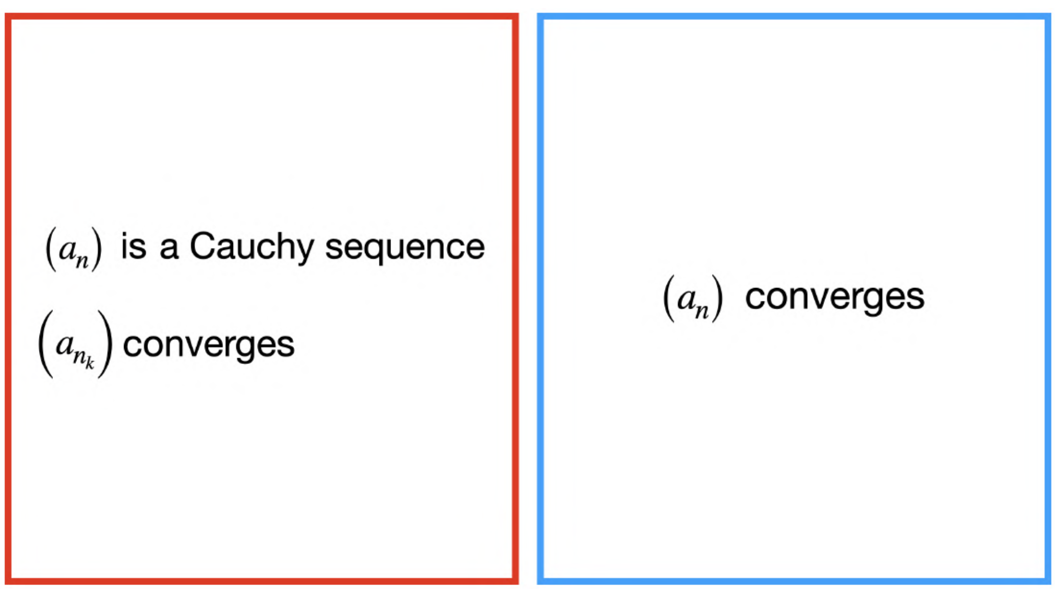
Now, it would be appropriate to give a name to the convergent subsequence, because if you're told that you've got something, you need to name it. We're going to call it a_{n_k}, because we want to remember that it's a subsequence of an and that it converges.



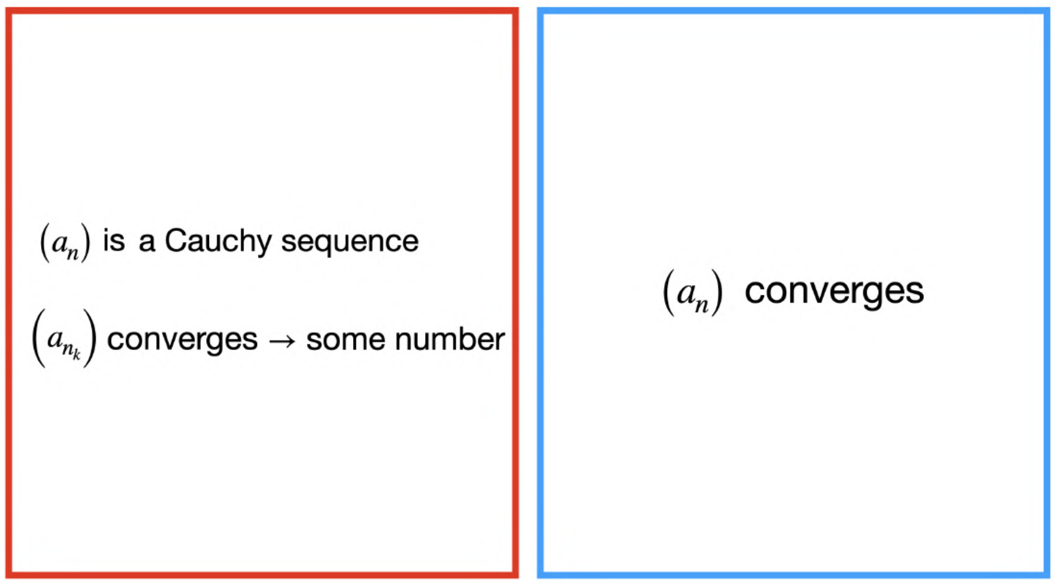
We're going to list these two properties separately for the sake of clarification.



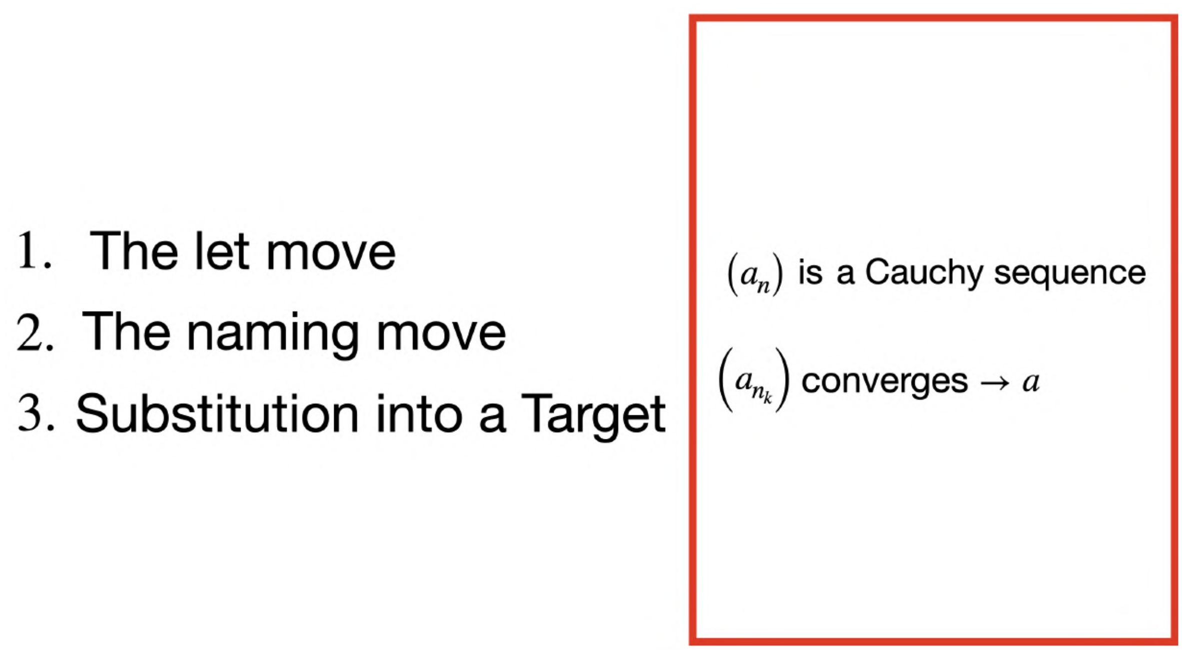
Now that it's clear though, let's remove the second hypothesis, because that's already obvious from the notation.



Great! Now we know that the sequence a_{n_k} converges to some number.

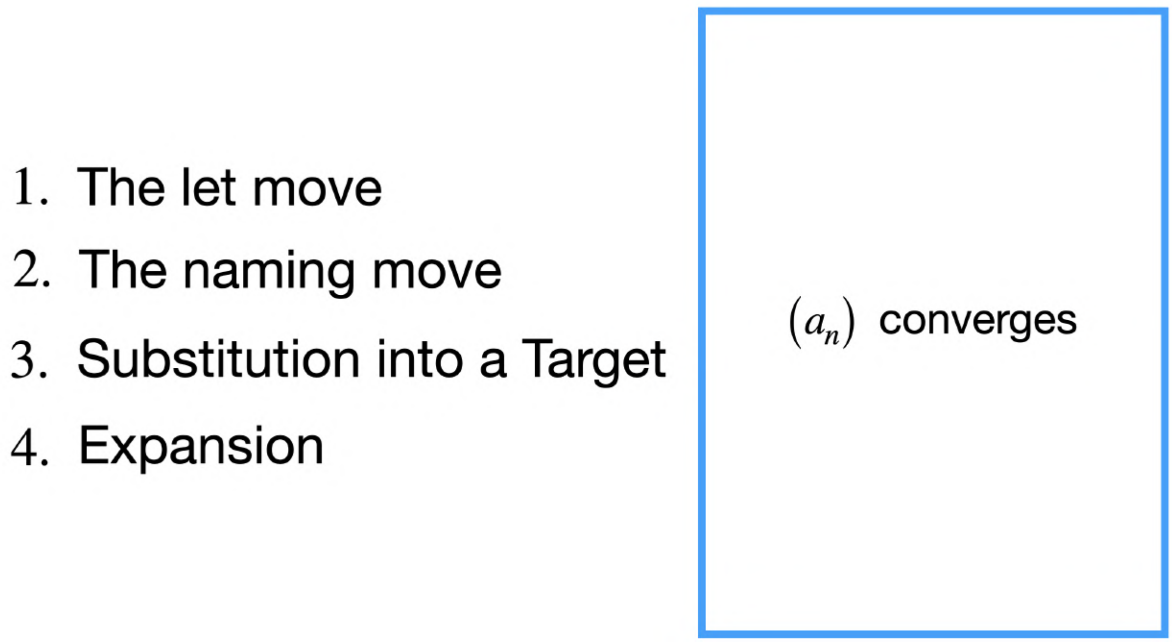


So instead of calling it some number let's call it a.



This move is known as **Substitution into a Target**.

Now that we've had enough of a look at our assumptions on the left side, let's try and take a closer look at what it is that we are actually trying to prove. In order to do that, we are going to use a process called *expansion*.



To expand means to take the definition you've got and write it out in more detail. Usually, it'd be good to avoid expanding any definitions you may have unless you're stuck. But here we really need to go back to first principles because we're proving everything from scratch.

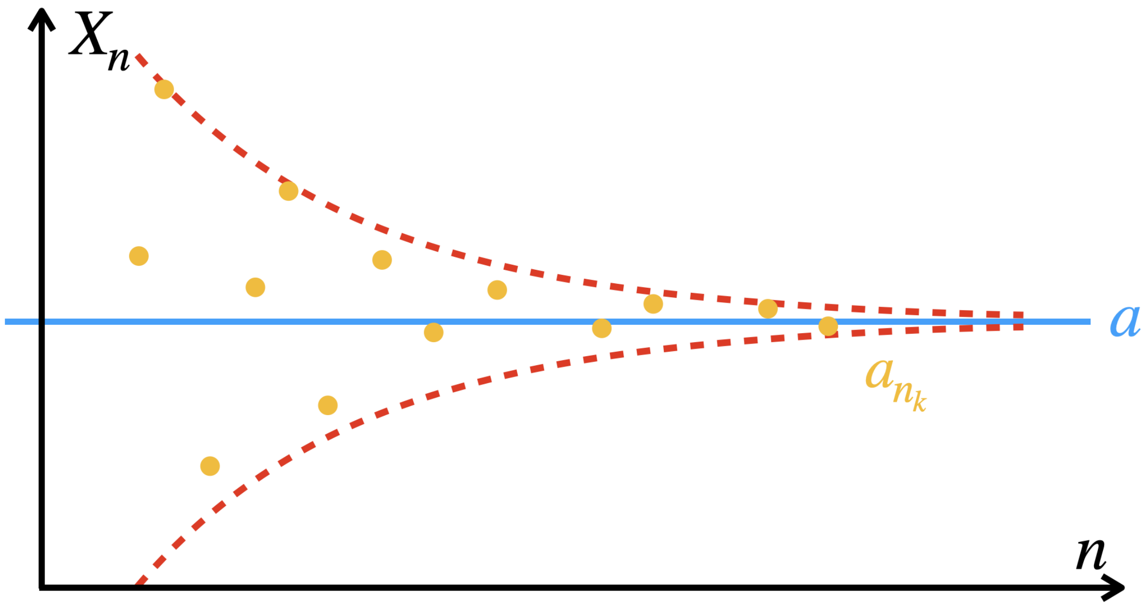
So let's remember what it means for a sequence to converge:

A sequence (a_n) converges if there exists some number x such that $a_n \to x$

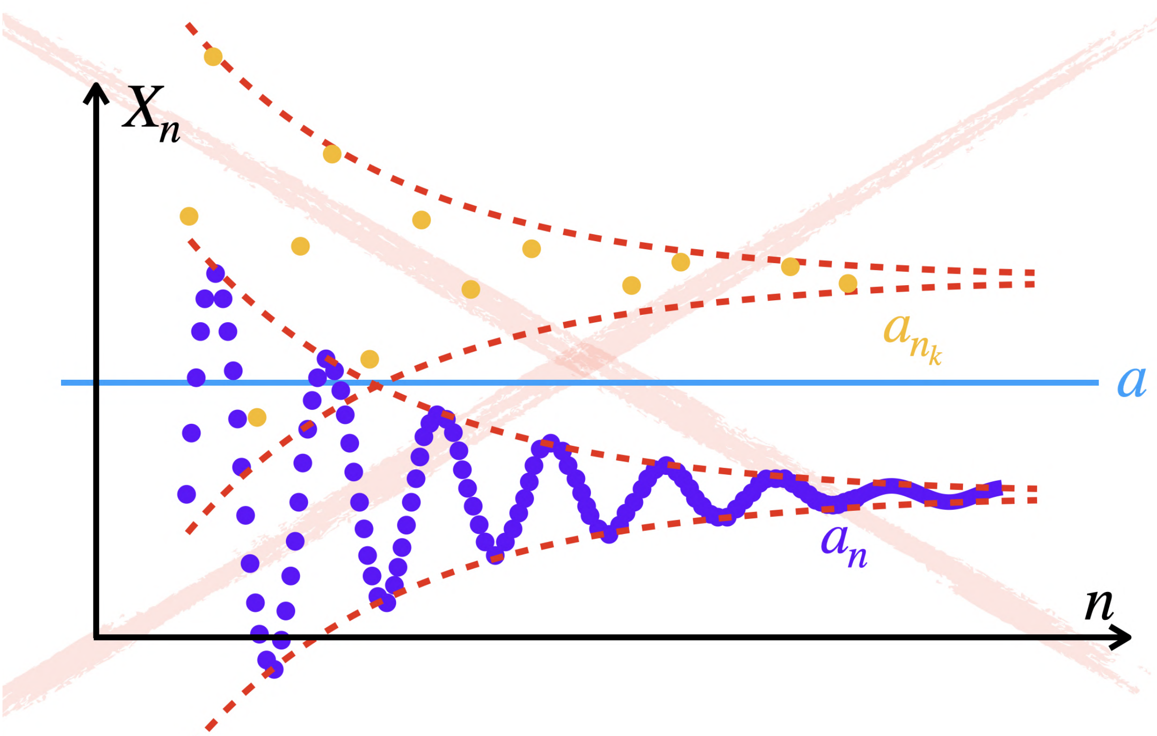
Now, our table becomes this:

(a_n) is a Cauchy sequence $a_{n_k} \to a$	~~(a_n) converges~~ $\exists x$ (a_n) converges to x

Now, our goal changed to an existential statement. So how will we find the x that the sequence converges to? Proving the existence of something in math can sometimes be extremely hard. That is unless you already have a good guess for what it should be. And in this case, we do.



We already have a subsequence a_{n_k} that converges, and it converges to some number a. Since a_n is a Cauchy sequence, its terms all bunch together, meaning they can't split and go to different limits.



The most reasonable candidate for the full sequence's limit is the same as the a. So now we write this:

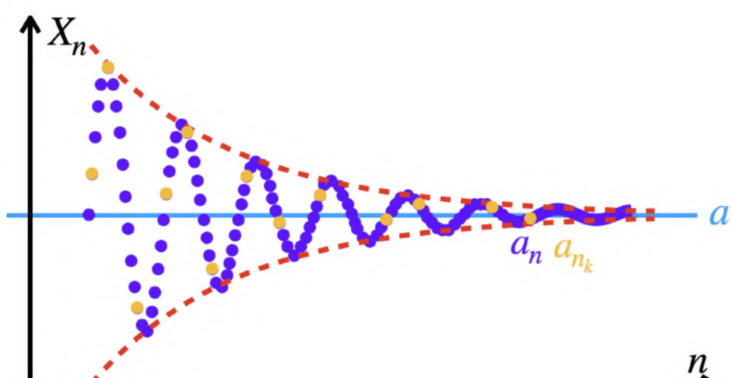
(a_n) is a Cauchy sequence

$a_{n_k} \to a$

$a_n \to a$

We've changed the problem from asking "does some limit exist?" to "let's show it converges to this particular number".

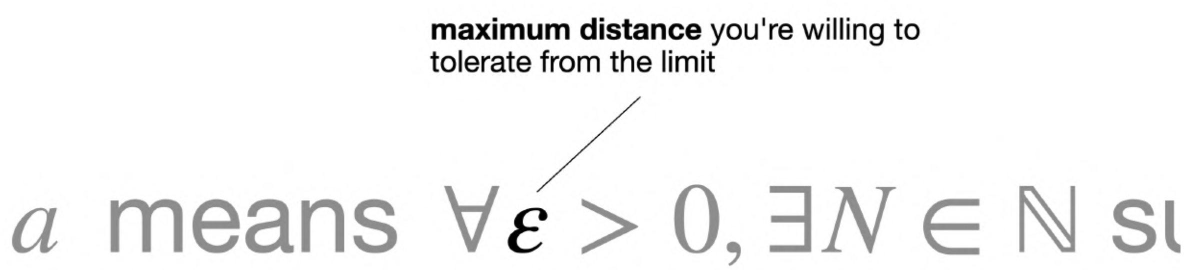
$a_n \to a$



Now we have to continue to expand, and show that $a_n \to a$, or in other words that the terms of the sequence get closer and closer to a. Here we need to familiarize ourselves with this definition:

$$a_n \to a \text{ means } \forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } \forall n \geq N, |a_n - a| < \varepsilon$$

We choose a small number ($\varepsilon > 0$). Think of this as the maximum distance you're willing to tolerate from the limit a.



($\exists N$) There exists a point in the sequence – call it position N – after which all terms are close to a.

($\forall n \geq N$) Once you go past that N, all future terms of the sequence stay close to a.

($|a_n - a| < \varepsilon$) Close to a means: the distance between a_n and a is less than ε.

Therefore our table becomes this:

(a_n) is a Cauchy sequence $a_{n_k} \to a$	$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N \; |a_n - a| < \varepsilon$

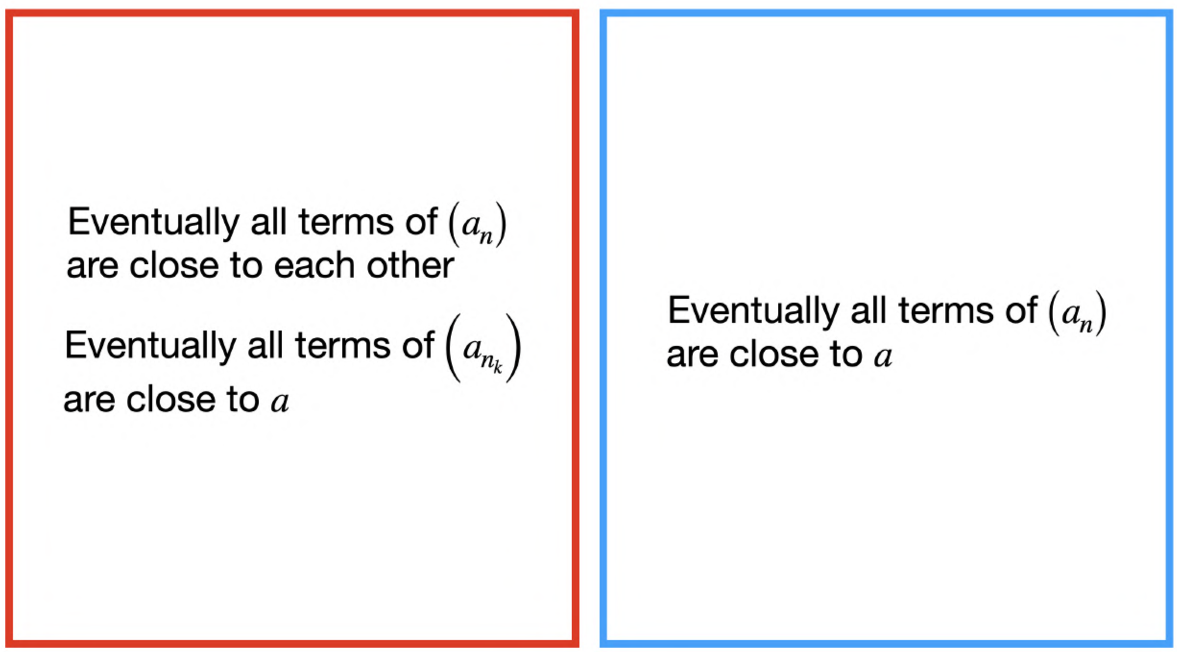
Expanding it like this gives us a way to build the proof from scratch.

Our target now begins with a universal quantifier ($\varepsilon > 0$) and we apply the "let" move again, just as we did last time.

(a_n) is a Cauchy sequence $a_{n_k} \to a$ $\varepsilon > 0$	$\forall \varepsilon > 0 \; \exists N \; \forall n \geq N \; |a_n - a| < \varepsilon$

So, whenever your goal is of the form "for every x something happens" we transform it into "Let x be arbitrary" and then try to prove the rest of the arbitrary value.

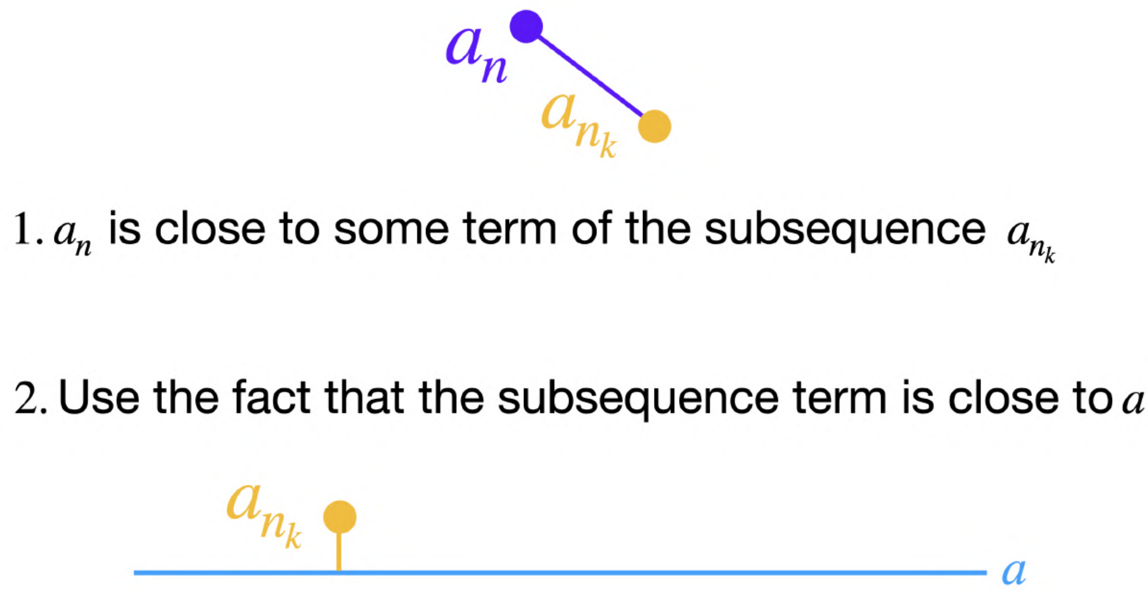
Okay, here is when things get a little harder, because this time we don't have an obvious answer for the thing that we are trying to find, which is N. We can try and write out the meaning of each statement in general terms:



At this point, the overall strategy of the proof should make sense, read carefully:

If all the terms in the subsequence get close to the number a, and if all the terms in the full sequence get close to each other (because it's Cauchy), then that means any term in the full sequence will eventually be close to some term in the subsequence – and since that subsequence term is already close to a, the original term must be close to a too.

Looking at the big picture, we want to show that eventually, every term of the sequence is within ε of the limit a.



But we don't know how to jump straight from a_n to a. Instead, we use a two-step detour:

1. First, we show an is close to some term of the subsequence a_{n_k}.

2. Then, we use the fact that that subsequence term is close to a.

If both are within $\varepsilon/2$, then by the *triangle inequality*:

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

So let's apply the two main hypothesis with $\varepsilon/2$, and see what we get:

$$\exists N_1 \ \forall p, q \geq N_1 \left| a_p - a_q \right| < \varepsilon/2$$

$$\exists N_2 \ \forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

$$\exists N \ \forall n \geq N \left| a_n - a \right| < \varepsilon$$

Here we are in a position where we've been kind of given something, which in this particular case is N_1 and N_2. We can go ahead and drop the there exists and just use N_1 and N_2.

$$\exists N_1 \ \forall p, q \geq N_1 \left| a_p - a_q \right| < \varepsilon/2$$

$$\exists N_2 \ \forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

Even though some people might object, this is a pretty common thing to do because once you've acknowledged that something exists, you choose it and give it a label – this makes the reasoning flow better.

$$\forall p, q \geq N_1 \left| a_p - a_q \right| < \varepsilon/2$$

$$\forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

$$\exists N \ \forall n \geq N \left| a_n - a \right| < \varepsilon$$

Okay, now we want to make sure that the absolute difference between a_n and a is less than epsilon: $|a_n - a| < \varepsilon$.

$$\exists N \ \forall n \geq N \left| a_n - a \right| < \varepsilon$$

How can we force it to do that? Or in other words, how can we make sure that the distance between a_n and a becomes smaller than any positive number we choose?

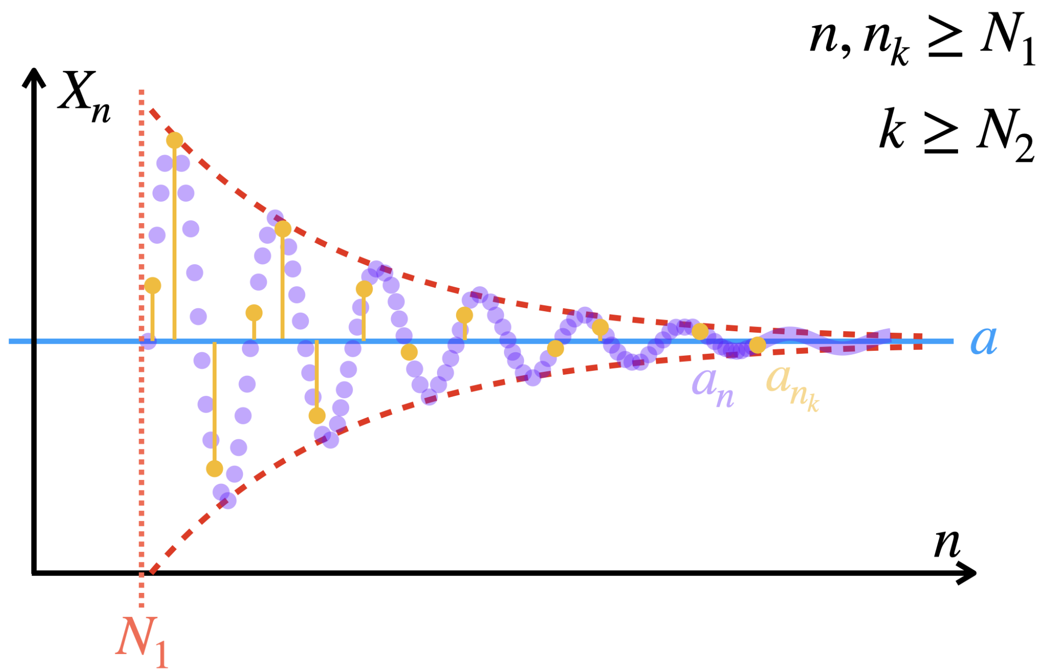
$$\exists N \ \forall n \geq N \left| a_n - a \right| < \varepsilon$$

Well, through trying and picking a value k so that two things happen: a_n is close to the subsequence term a_{n_k}, and that subsequence term a_{n_k} is close to the limit a. Both within half of ε:

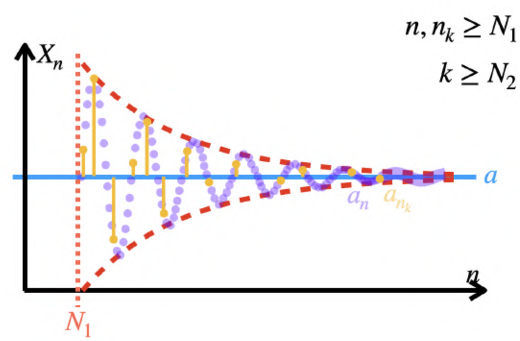
$$a_n \text{ is close to } a_{n_k} : |a_n - a_{n_k}| < \varepsilon/2$$

$$a_{n_k} \text{ is close to } a : |a_{n_k} - a| < \varepsilon/2$$

The fact that the sequence is Cauchy guarantees that a_n and a_{n_k} are close – as long as both of their indices are large enough (specifically, $\geq N_1$).



The fact that the subsequence converges to a guarantees that a_{n_k} is close to a – again, as long as k is large enough (specifically, $k \geq N_2$).



1. The let move
2. The naming move
3. Substitution into a Target
4. Expansion
5. Substitution into a Hypothesis

This move is known as **Substitution Into a Hypothesis**.

So our plan is to label the index of our main sequence term as p, and the index of the matching subsequence term as q, so we can apply the Cauchy property between them.

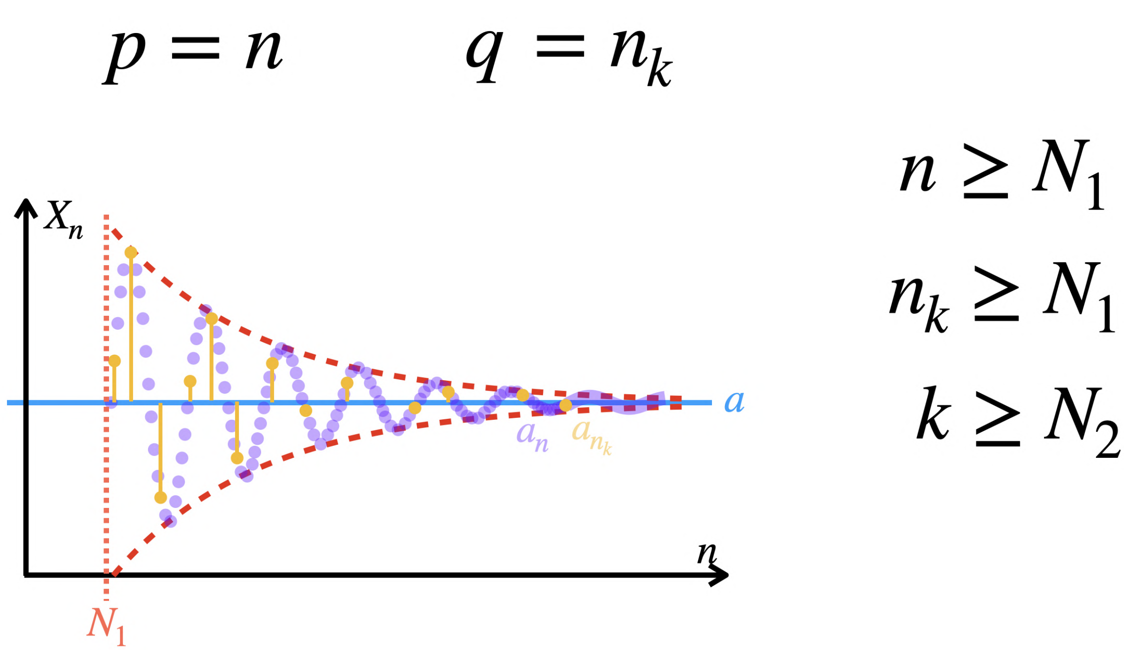
$$p = n \quad \text{and} \quad q = nk$$

To make all this happen, we need:

$(*)$ The index n to be big enough for the Cauchy property;

$(*)$ The subsequence index n_k to also be big enough for the Cauchy property, and

$(*)$ The position k in the subsequence to be big enough for the convergence part.



We've gathered everything we need, so now it's time to actually pick a value for N.

$$\forall p, q \geq N_1 \left| a_p - a_q \right| < \varepsilon/2$$

$$\forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

$$\forall n \geq N_1 \left| a_n - a \right| < \varepsilon$$

Since our Cauchy condition becomes valid when n is at least N_1, we might as well define N to be exactly that. That way, our conclusion $|a_n - a| < \varepsilon$ will be guaranteed to hold when $n \geq N_1$.

By picking a specific value for N (namely, N_1), we no longer need to say "there exists some N", because we've already found one!

So we remove the existential quantifier and now write the conclusion directly as this:

$$\forall n \geq N_1, \ |a_n - a| < \varepsilon$$

Let's now apply the "let" move again as we did previously, which lets us drop the "for all" and we can just focus on a single, arbitrary n.

$$\forall p, q \geq N_1 \left| a_p - a_q \right| < \varepsilon/2$$

$$\forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

$$n \geq N_1$$

$$\left| a_n - a \right| < \varepsilon$$

Earlier we said we would eventually choose $p = n$ and $q = n_k$. Now we're locking that in – we can choose $p = n$ because we previously assumed that $n \geq N_1$, which is the condition needed to use the first hypothesis. We've already used the assumption $n \geq N_1$ to justify applying the first inequality, so we don't need to keep it listed anymore.

1. The let move
2. The naming move
3. Substitution into a target
4. Expansion
5. Substitution into a hypothesis
6. Modus ponens

$$\forall q \geq N_1 \left| a_n - a_q \right| < \varepsilon/2$$

$$\forall k \geq N_2 \left| a_{n_k} - a \right| < \varepsilon/2$$

$$n \geq N_1$$

This is called the **Modus Ponens** move.

Anyway, we need to pick specific values for the indices q and k that satisfy:

$$k \geq N_2$$

$$n_k \geq N_1$$

This is so that:

a_n is close to a_{n_k}

a_{n_k} is close to a

We're going to use r to denote the k, and write down the conditions r should satisfy.

Okay, let's move n_r into the first hypothesis and r into the second hypothesis:

$$\left|a_n - a_{n_r}\right| < \varepsilon/2$$

$$\left|a_{n_r} - a\right| < \varepsilon/2$$

$$\left|a_n - a\right| < \varepsilon$$

Finally, we finish with the triangle inequality (which is a classic trick):

$$|a_n - a| \leq |a_n - a_{n_r}| + |a_{n_r} - a| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

So, $a_n \to a$, which proves that the original sequence converges. $\square$

Now, let's recap all the moves we've used to get here.

1. The let move
2. The naming move
3. Substitution into a target
4. Expansion
5. Substitution into a hypothesis
6. Modus ponens

First, the "let" move. When you're trying to prove something like "for every x such that $P(x)$, we have $Q(x)$," you "let" x be some object satisfying $P(x)$, and then shift your target to proving $Q(x)$.

This simplifies the proof step-by-step.

Second, the Naming move. Basically if you are told something exists, give it a name.

Third, Expansion. Try and make abstract definitions more usable in proofs.

Instead of relying on the word "converges" or "Cauchy", rewrite the definition explicitly with quantifiers and inequalities $\forall \epsilon > 0 \ldots$

Fourth, Substitution into a Hypothesis – apply a general statement to a specific case.

Given something like $\forall \eta > 0, \exists N, \forall p, q \geq N, |a_p - a_q| < \eta$, you can choose a specific value (e.g. $\eta = \epsilon/2$) and substitute it in.

Fifth, Modus Ponens. If you have $P(u) \implies Q(u)$, and you've already shown $P(u)$, then you can conclude $Q(u)$.

This is a logical deduction based on satisfying conditions.

And sixth, Substitution into a Target. If your goal is to prove a statement of the form "$\exists u$ such that $P(u)$", and you have a candidate x that you think will satisfy P, then you can change your target to that of proving $P(x)$ directly.

In the proof, this happened when we moved from proving "(a_n) converges to something" to "$(a_n) \to a$".

This doesn't mean that these are all the moves that exist in proofs, but it's a gateway to show you that proofs can become fairly automatic and straightforward when you've had enough practice.

This PDF file was based on an article by *Timothy Gowers* (link below).

How to work out proofs in analysis I.

If you found this document useful let us know. If you found typos and things to improve, let us know as well. Your feedback is very important to us. We're working hard to deliver the best material possible. Contact us at: dibeos.contact@gmail.com