What does it mean, exactly, that sequence \((x_n)\) converges to limit \(a\)? In most proofs-oriented analysis books, you will find an \(\epsilon\)-definition of a limit, which is hard to digest when you first see it. On the other hand, you probably remember from calculus that convergence means that the terms \(x_n\) of the sequence get closer and closer to \(A\). The calculus description is not precise and gives rise to a lot of questions. How close is “close”? If one term is close to \(A\), should the next be even closer, or can it be further out? In class, we studied a lot of examples to make these notions precise. To pin down the idea of closedness, we considered various neighborhoods of the limit point \(A\).

**Definition 1.** Let \(A\) be a number (thought of as a point on the real line). A neighborhood of \(A\) is any open interval centered at \(A\).

A neighborhood can be large or small; the intervals \((-100, 100)\) and \((-0.001, 0.001)\) are both neighborhoods of 0, the intervals \((-2, 0)\) and \((-1.01, -0.99)\) are neighborhoods of \(-1\).

Now, for a sequence \((x_n)\) to converge to \(A\), we require that the terms of the sequence concentrate in any chosen neighborhood of \(A\), no matter how small the neighborhood is. More precisely, for any neighborhood that someone might choose, we must have that eventually, starting from a certain moment in the sequence, all terms \(x_n\) are contained in that chosen neighborhood. Note the word **eventually**: it means that in the beginning of the sequence, there may be a lot of “junk” terms that are very far from \(A\) and do not fit into the given neighborhood. However, if you move forward along the sequence, you must be able to find a moment in a sequence where the junk stops, and all the terms coming after that moment fit into the chosen neighborhood.

The idea of the previous paragraph is easier to articulate if we introduce a bit of jargon.

**Definition 2.** A tail of the sequence \((x_n)\) includes all its terms after a certain moment (index). For example, \((x_5, x_6, x_7, x_8, \ldots)\) and \((x_{1000}, x_{1001}, x_{1002}, x_{1003}, \ldots)\) are two different tails of \((x_n)\). Sometimes we will use the term \(N\)-tail to refer to all terms with indices greater than \(N\), i.e. the \(N\)-tail is \(x_{N+1}, x_{N+2}, \ldots\). For example, 100-tail is \((x_{101}, x_{102}, x_{103}, \ldots)\).

Note that while a tail is missing some terms in the beginning of the sequence, it must include all subsequent terms once it starts. For instance, a sequence of terms

\[
(x_{100}, x_{102}, x_{104}, x_{106}, \ldots)
\]
OLGA PLAMENEVSKAYA

is not a tail of \((x_n)\) because it misses odd-numbered terms \(x_{101}, x_{103}, \ldots\). A sequence \((x_{100}, x_{102}, x_{103}, x_{104}, \ldots)\) is not a tail because it misses \(x_{101}\), even though it includes all terms from \(x_{102}\) on.

Now the requirements for a convergent sequence can be stated neatly:

**Definition 3.** \((x_n)\) converges to \(A\) if every neighborhood of \(A\) contains some tail of the sequence \((x_n)\).

In other words, \(\lim x_n = A\) means that every neighborhood captures a tail of our sequence — but of course different neighborhoods capture different tails. A large neighborhood of \(A\) might contain the whole sequence; for a very small neighborhood, you’ll probably have to go pretty far out in the sequence before the required tail starts. Once the tail is in the neighborhood, there’s no requirement on how the terms should behave — they do not have to get closer and closer to the limit, increase or decrease, or follow any pattern.

Note that the above definition sounds unofficial, but it is mathematically rigorous: we have specified the precise meaning of all the words we used.

**Example.** Using this definition, let’s show that the sequence \((\frac{1}{n})\) converges to 0. We have to examine all possible neighborhoods of 0 and show that each of them captures a tail of our sequence. For instance, if we take the neighborhood \((-\frac{1}{10}, \frac{1}{10})\), we notice that all terms \(\frac{1}{11}, \frac{1}{12}, \frac{1}{13}, \ldots\) are positive numbers that are less than \(\frac{1}{10}\), and thus are contained in the interval \((-\frac{1}{10}, \frac{1}{10})\). Therefore, \((\frac{1}{11}, \frac{1}{12}, \frac{1}{13}, \ldots)\) is a tail captured by this neighborhood.

This is not the end of the proof though: we have to study all possible neighborhoods of 0, not just one neighborhood that we like. A general neighborhood of 0 is an open interval symmetric about 0; we can write it as \((-d, d)\) if it extends distance \(d\) on each side of 0. The point \(\frac{1}{n}\) will be contained in this interval as long as \(\frac{1}{n} < d\), or, in other words, whenever \(n > \frac{1}{d}\). Thus, the tail \((x_N, x_{N+1}, x_{N+2}, \ldots)\) will be entirely contained in \((-d, d)\) if we take \(N\) to be any integer that is greater than \(\frac{1}{d}\).

A very similar argument shows that the sequence \((-\frac{1}{n})\) also converges to 0. (Notice that in this sequence, the terms alternate between positive and negative jumping around 0. It is always helpful to write out a few first terms, \((-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4})\). For an (arbitrary!) neighborhood \((-d, d)\), we notice that the (positive or negative) number \((-\frac{1}{n}) = \pm \frac{1}{n}\) is contained in \((-d, d)\) whenever \(n > 1/d\). As before, we now pick any integer \(N > d\), and see that \((x_N, x_{N+1}, x_{N+2}, \ldots)\) will be entirely contained in \((-d, d)\).

Most textbooks, including Ross, use an epsilon-definition of the limit. Though less visual, the more “formal” definition is often useful for calculations, when the sequence is given by a complicated formula which is harder to plot and visualize. We will arrive at that definition by translating neighborhoods-and-tails into formulas.

The definition of the limit says that an arbitrary neighborhood of the limit \(A\) must capture some tail of the sequence \((x_n)\). An arbitrary neighborhood of \(A\) is an open interval centered at \(A\); if it extends distance \(d\) on each side, we can write it as \((A - d, A + d)\). Typically, the Greek letter \(\epsilon\) is used instead of \(d\), so an arbitrary neighborhood can be written as \((A - \epsilon, A + \epsilon)\). (Since it stands for a distance, \(\epsilon\) is a positive number.) We can also replace the phrase “the neighborhood contains some tail” by the phrase “we can find a number \(N\)
such that the $N$-tail is contained in the neighborhood”. The statement
\emph{every neighborhood of $A$ contains some tail of $(x_n)$}

can now be written as
\emph{for every $\epsilon > 0$, the interval $(A - \epsilon, A + \epsilon)$ contains some tail of $(x_n)$}

and then as
\emph{for every $\epsilon > 0$, we can find $N$ such that the interval $(A - \epsilon, A + \epsilon)$ contains the $N$-tail of $(x_n)$}.

If we recall that $N$-tail consists of all terms $x_n$ with $n > N$, we can rewrite the last statement once again:
\emph{for every $\epsilon > 0$, we can find $N$ such that the interval $(A - \epsilon, A + \epsilon)$ contains $x_n$ whenever $n > N$}.

Finally, we give a closer look to the statement “\textit{the interval $(A - \epsilon, A + \epsilon)$ contains $x_n$}”. It means
\begin{align*}
A - \epsilon &< x_n < A + \epsilon, \\
-\epsilon &< x_n - A < \epsilon, \\
|x_n - A| &< \epsilon.
\end{align*}

The last inequality can also be obtained in a more geometric way: “\textit{the interval $(A - \epsilon, A + \epsilon)$ contains $x_n$}” iff the distance between $x_n$ and $A$ is less than $\epsilon$. But the distance between $x_n$ and $A$ is exactly $|x_n - A|$. Indeed, if $a$ and $b$ are two points on the real line, the distance between them equals to $b - a$ (if $b > a$, i.e. $b$ to the right of $a$) or $a - b$ (if $b < a$, i.e. $b$ to the left of $a$), or 0 if $b = a$, which gives that the distance is $|a - b|$ in all cases.

Therefore, the final translation of our statement is
\emph{for every $\epsilon > 0$, we can find $N$ such that $|x_n - A| < \epsilon$ whenever $n > N$}.

Replacing “we can find” by more formal wording, “there exists”, we recover the definition of the limit found in most books (compare Ross Definition 7.1).