

The three standards, the sign convention, and the frame math - on one page.

## Read the sign first

**Positive (+)** Audio AHEAD of picture (early). Unnatural - noticed sooner.

**Negative (-)** Audio BEHIND picture (late). Natural - tolerated longer.

## The three standards (positive = audio ahead of picture)

Standard	Scope	Ahead	Behind	Notes
ITU-R BT.1359-1	viewer perception - detectability	+45 ms	-125 ms	Inside this, almost all viewers perceive sync. Aim here.
ITU-R BT.1359-1	viewer perception - acceptability	+90 ms	-185 ms	Beyond this it is distracting; quality degrades.
EBU R37	per single production stage	+5 ms	-15 ms	Error-accumulation limit, not a perceptual one.
EBU R37	overall, output to transmitter	+40 ms	-60 ms	Sum of stages; still inside BT.1359.
ATSC IS-191	DTV encoder input (North America)	+15 ms	-45 ms	Lock sync before it is baked into the stream.

## Turning milliseconds into frames

24 fps	1 / 24 = 41.67 ms per frame	+45 ms ahead ~ 1 frame; -125 ms behind ~ 3 frames
25 fps	1 / 25 = 40.00 ms per frame	+45 ms ahead ~ 1.1 frames; -125 ms behind ~ 3.1 frames
30 fps	1 / 30 = 33.33 ms per frame	+45 ms ahead ~ 1.4 frames; -125 ms behind ~ 3.75 frames

## Remember

- Perfect sync does not exist - perception accepts a window, not a point.
- The window is asymmetric: late audio is tolerated ~2.7x more than early audio.
- Target band (BT.1359 detectability): +45 ms ahead to -125 ms behind.
- Aim to land slightly LATE - the widest, most forgiving part of the window.
- Production rules (EBU R37, ATSC IS-191) are tighter to control accumulation.
- Sharp transients (claps, drums) are policed far more strictly than speech.
- Always state the sign: '+' = audio ahead; '-' = audio behind. Never just a number.
- The 'within one frame' rule is symmetric and wrong - the real window is lopsided.