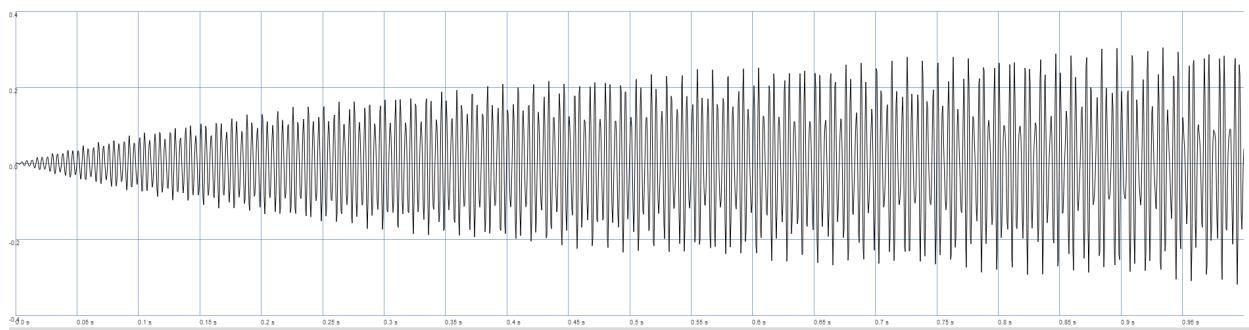


How do computers make sound?

We use computers so often to watch videos and make music that it's easy to slip by without asking yourself "how does that even work?" but in this article we'll be trying to do exactly that: we'll talk a little about what sound even is, how a computer can make it, and the ways music is stored digitally!

The first thing to address is what *is* sound, anyway? The short answer is that "sound" is how we, through our ears and the rest of our bodies, experience *vibration*. We feel these *patterns*, *waves*, of high and low pressure that cause our ear drums to move in and out repeatedly.

When we think of *musical notes* what we're describing are particular regular frequencies of vibration, often measured in Hz---which is short for hertz, pronounced "hertz", and is a measure of how many times per second something happens. For example, we've defined the frequency 440Hz to be the A above middle-C on a piano. This means that a vibration happening smoothly four-hundred and forty times per second *will be* the note A4 to our ears. Of course, real sounds are far more complicated than just a single tone. If you were to hit the A4 key on a piano you'd hear much, much more than just a simple 440Hz. This complicated mess of vibrations is what makes every instrument different and rich, because a pure tone is actually kind of boring!



Now, the complication here is that describing these messes of vibrations isn't obviously something a computer could do well. Sound moving through the air---like most things in physics---is inherently *analog*, which means that a precise mathematical description involves numbers *that can't be written down with a finite number of digits*. But wait, you might ask, why would that be true?

Hold your thumb and forefinger so that they're a bit apart. Look at the space between them. From what we know of the universe so far, there is an infinite number of points between your two fingers. Hold your fingers closer together and there's still an infinite number of points between them. Heck, press your thumb and finger *together* and there's still an infinite number of points between them because you can't *actually* ever press the molecules of any two things together. This property of there always being infinitely more *stuff* no matter how small you try to

zoom in is what causes you to need *an infinite number of digits after the decimal point* in order to name that exact spot in the infinite mess.

Similarly, if you were to try and describe the shape of the vibrations perfectly you'd have to do it at an infinite number of points! This is absolutely not doable on a computer.

But, okay, so given that we obviously *do* listen to music on a computer how exactly does it work? The first step is that since you can't include data on the wave shape for *all* the infinite points you need to pick a finite number of them. This means that you take a certain number of *samples* per second; this is called the *sampling rate*. These samples tell you the shape of the vibration at that point in time. Now, sampling rate is also measured in Hz just like the frequency of the sound. In fact, there's a really cool result called the Nyquist theorem that tells us a simple relationship between the range of vibrations we want to encode and how many times we need to sample per second. The punchline is that if you take the highest frequency you want to accurately encode then you need a sampling rate *twice* that at minimum. So since human hearing caps out at around 20000Hz (20kHz) then you need a sampling rate of at least 40kHz, which is how most audio is encoded!

The magic then happens with something called the digital-analog converter (DAC). Every computer, every phone, every digital device that can play sound has one. It's the dedicated device that takes all these *samples* that are the digital approximation of the real physical sound and converts it into a signal that can drive the speakers. The cool thing here is that while *digital* devices like computers can't create real analog signals it turns out that you can create pure tones out of electronics components. The DAC takes the samples, creates a signal that's like a jerky version of the real audio with sudden changes that would sound terrible if it were played, and then smooths it out into transitions that are much more natural.

The end result is an electrical current that looks just like the vibrations of the sound that then will cause the speakers to vibrate in and out, finally creating sound!