Winning the loudness war

Harlan Connor

November 11, 2023

About me

PhD maths student at the University of Liverpool.
Outside of my work, I'm a amateur radio operator.
Also do some music, and there is a surprisingly large intersection between them.

Backstory

A YouTuber (Dan Worrall) told people not to purely optimise for loudness.
A commenter told him that he was just unable to make loud songs.
Dan drops a song with a peak 2x the maximum storable value.

Sampling

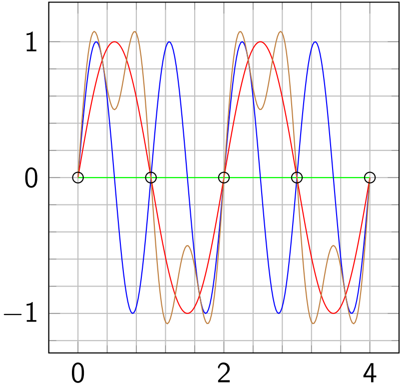
How do we store a continuous-time function (i.e. sound) in finite space?

The standard approach is sampling it at fixed intervals, and joining it up.

Digital systems usually store samples within $[-1, 1]$, with volume knob scaling these values.

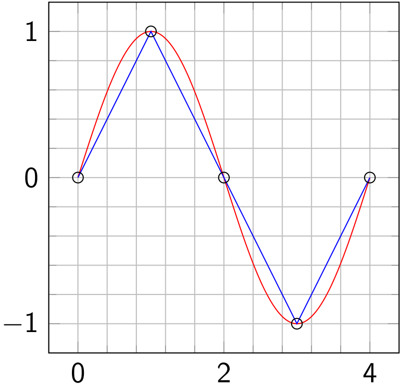
Aliasing

Many functions may sample the same
What are the ‘nicest’ conditions to ensure uniqueness?



Interpolation

Just joining the dots doesn't give us the wave back, and it will sound different.
How can we interpolate correctly?



The sampling theorem

Aliasing is solved by the following theorem:

Theorem

Any real function with frequency component all $< f$ can be perfectly reconstructed by sampling with frequency $\geq 2f$.

The sample rate $2f$ is called the ‘Nyquist limit’.
But how do we reconstruct it?

The interpolation formula

For samples a_n taken at an interval of T, our reconstruction is:

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \cdot \text{sinc}\left(\frac{t - nT}{T}\right)$$

With the normalised sinc function defined as:

$$\text{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$$

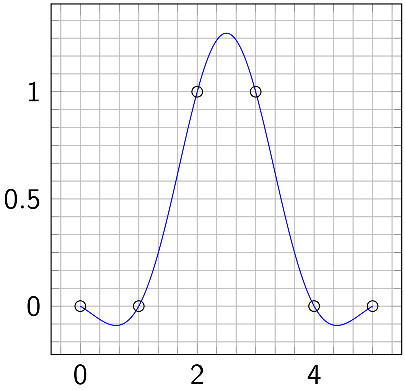
Inter-sample peaks

$$x(t) = \sum_{n=-\infty}^{\infty} a_n \cdot \mathrm{sinc}\left(\frac{t - nT}{T}\right)$$

At $t = kT$, only the sample at t is taken into account.
Can we have $x(t)$ bigger than the two samples either side?

Inter-sample peaks

Yes!
Inter-sample peaks mean that sounds louder than any individual sample can be produced.
How much louder?



Supremal loudness

Let's put our spike at $t = 1/2$, and sample interval $T = 1$. This gives us, with some rearranging:

$$x(1/2) = \sum_{n=-\infty}^{\infty} a_n \cdot \text{sinc}\left(n - \frac{1}{2}\right)$$

$$= \frac{1}{\pi} \sum_{n=-\infty}^{\infty} a_n \cdot \frac{\sin\left(\left(n - \frac{1}{2}\right) \cdot \pi\right)}{n - \frac{1}{2}}$$

$$= \frac{1}{\pi} \sum_{n=-\infty}^{\infty} a_n \cdot \frac{(-1)^n}{n - \frac{1}{2}}$$

Some of you should see where this is going...

Supremal loudness

It is clear that a_n should make the nth term positive, and take the largest value (i.e. ± 1):

$$a_n := \operatorname{sgn}\left(\frac{(-1)^n}{n-\frac{1}{2}}\right)$$

$$x(1/2) = \frac{1}{\pi}\sum_{n=-\infty}^{\infty}\frac{1}{\left|n-\frac{1}{2}\right|} \geq \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{1}{n-\frac{1}{2}} \geq \frac{1}{\pi}\sum_{n=0}^{\infty}\frac{1}{n}$$

Uh-oh...

Supremal loudness

We have that $x(1/2) \geq \frac{1}{\pi} \sum_{n=0}^{\infty} \frac{1}{n}$, which is a non-zero multiple of the harmonic series, and so tends to infinity!

We can make *arbitrarily loud* sounds, and we have the formula for it!

Your computer either should be able to output arbitrarily large voltage spikes, or it is infinitely inaccurate.

Most computers do a mix: this can actually damage audio equipment!

Thanks for listening!