Sound and Perception

Sound is time-varying pressure in the molecules of the air. We could use a function $p(t)$ to represent this pressure as a function of time, and graph it.

A constant pressure $p(t) = c$ is silence. Atmospheric pressure is, relatively speaking, constant.

A pressure which varies periodically with a period of roughly between 20 and 20,000 Hertz can be perceived by humans as sound; we say these frequencies are in our auditory range.

Sound Reproduction

When we record sound using microphones, we are using a transducer to convert the pressure function from a mechanical wave to an electronic wave. The electronic wave is then encoded somehow (using analog or digital techniques).

Record albums were made by encoding the pressure function in grooves in the vinyl, cassettes by encoding the function magnetically. CDs and other forms of digital audio encode samples of the pressure function.
Digital Audio

The function $p(t)$ is a continuous function. To represent a sound wave digitally we have to approximate this continuous function by taking samples of it, a process called Analog-to-Digital conversion (ADC).

Then, we can play these samples fast enough to produce the illusion of continuous sound, similarly to the way motion pictures display a sequence of camera shots fast enough to fool the brain into seeing motion. This is Digital-to-Analog conversion (DAC).

Analog->Digital Conversion

Sampling is the process of taking 'snapshots' of the continuous function $p(t)$ at even time intervals, then storing each value for later playback or analysis.

How do we decide how fast to sample the function? That is, how small should the sampling time interval be, and how many total samples do we need?

Another way of asking this: how can we get the best possible approximation for our hearing apparatus?

The Sampling Theorem

Luckily, the field of information theory was being developed alongside the development of computers by Claude Shannon. While information was going digital, Shannon and others were ferreting out the theory of reliably encoding and transmitting information.

The sampling theorem (sometimes the Shannon-Nyquist theorem or just Nyquist theorem) says any function $p(t)$ is completely determined by sampling at twice the highest frequency to be reproduced. Most of us can detect frequencies of around 20,000 Hz max, so at minimum we should sample audio at least 40,000 times per second, or once every 0.025 milliseconds. Professional-quality recording is often done at much higher sample rates.

Digital->Analog Conversion (DAC)

The spec for CD quality audio is to playback at 44,100 samples taken per second. This reproduces audio up to a frequency of 22,050 Hz, which is fine for younger ears not yet decimated by playing loud music through headphones.

What does this mean for your dog?

In principle, any sound can be produced by generating its pressure function. So, any mechanism (hardware or software) that generates this function and can convert it to the electrical signals appropriate for driving a loudspeaker can in principle reproduce any sound, including those never before heard.
Milestones in Electro-Acoustic Music

- Musique Concrete - using everyday objects or sounds as instruments
- Use of reel-to-reel tape techniques (splicing, reversing)
- Electronic components - oscillators, envelope generators, filters, etc.
- Computer algorithms run on mainframes (beginning with Max Mathews)
- Integrated circuits (Digital Signal processors as hardware devices)
- Microprocessors become powerful enough for real-time synthesis

Tape Techniques

Raymond Scott's studio (1950's)

Tape Techniques

Old School Audio Editing

Electronic Modules

Used in analog synthesis
Music on Mainframes
Max Mathews at Bell Labs

The Moog Synthesizer
Walter Carlos’ Switched-on Bach

Going Digital
Switched-on Bach 2000

Music on Microprocessors
Max Mathews at 80