Sampling and analysing sound

Algorithm for sampling and spectral analysis of sound-data

The algorithm for spectral analysis is called FFT, Fast Fourier Transformation.

Parameters:

SF		Sampling frequency, e.g. 20000 Hz (typical 22050 Hz).
N		Total number of data points, here 4000 (typical 4096).
T	= N/SF	Total sampling time, e.g. 0,2 s.
dt	= 1/SF:	Time interval between two data points, here 5 10^{-5} s.
df	= 1/T = SF/N:	Fundamental frequency, interval between succeeding frequencies in the spectral analysis, here 5 Hz.
f_{max}	= SF/2	Maximum possible frequency in the spectral analysis, here 10000 Hz.

Properties of the algorithm for spectral analysis, FFT, Fast Fourier Transformation,
and considerations of the choices of number of datapoints N and sampling frequency SF:

- FFT delivers the spectral analysis as ”strengths” (amplitudes or amounts of energy) of different frequencies given as whole number multiples of the fundamental frequency df ranging from 1*df to ½N*df = ½ SF = f_{max} , ½ N different frequencies.
  Here: 2000 different frequencies: 5, 10, 15, 20, … , 9990, 9995, 10000 Hz.

- The choice of SF determines the maximum frequency f_{max} = SF/2 in the spectral analysis. Here SF = 20000 Hz, so harmonics with higher frequencies than 10000 Hz will not be included in the analysis. If you want to see higher harmonics, you must choose a higher sampling frequency, e.g. SF = 40000 Hz (typical 44100 Hz) implies f_{max} = 20000 Hz.

- If you keep the same number of data points N, then the choice of a higher sampling frequency SF implies a higher fundamental frequency df = SF/N. This again results in a more coarse spectral analysis having larger jumbs df between the frequencies,
  e.g. chosing SF = 40000 implies df = 10 Hz with analysis presenting 2000 different frequencies ranging from 10 Hz to 20000 Hz in steps of 10 Hz.

- To avoid the more coarse spectral analysis given the new sampling frequency you then maybe choose to increase the number of data points N. Then steps between frequencies df = SF/N becomes smaller, the peaks of the spectral analysis becomes more narrow and well defined. The cost of this will be more processing time and more memory requirements for the PC treating the software (and a longer total sampling time T = N/SF, which could be a problem sampling on a very non-uniform sound), e.g. with SF = 40000, doubling N so N = 8000 implies back again with df = 5 Hz and T = 0,2 s.

Fourier series

Any periodic function with period T can be represented by a trigonometric series of a sum of sine- and cosine expressions. This so-called Fourier series of the function can be given in two forms:

(1)

$$f(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(\frac{2n\pi}{T} t) + b_n \sin(\frac{2n\pi}{T} t))$$

(2)

$$f(t) = a_0 + \sum_{n=1}^{\infty} (c_n \cos(\frac{2n\pi}{T} t - \phi_n))$$

The fourier series of f(t) in the form (2) shows the amplitude c_n as a function of the whole number n or as indirectly as a function of the frequency $f_n = n/T$. The frequency df = $f_1 = 1/T$ is the so-called fundamental frequency of the fourier series.

The relationship between c_n and a_n, b_n is:: $c_n^2 = a_n^2 + b_n^2$.

In praktice it is not possible to represent the fourier series with all infinite many frequencies, it must be restricted to include a finite number of frequencies starting with the fundamental frequency.

Assuming the number of data points to be N = 4000 measurements of f(t) during the time period T = 0.2 s, the samplingfrequency SF becomes 4000/0,2 = 20000 Hz.

The 4000 data points result in the same number of equations with a_n and b_n as unknowns.
If 4000 simultaneous equations contain exactly 4000 unknowns then the equations have exactly one unique solution for each of the unknowns.

In the fourier series of form (1) each part of the sum contains 2 unknowns, a_n og b_n (disregarding the lonely unknown a_0 which can be solved in a special way).
To find an unique solution to the 4000 equations we can only allow 4000 unknowns, so having 2 unknowns in each part the sum must only contain 4000/2 = 2000 parts. We are then allowed only to sum up from n = 1 til 2000 corresponding to the frequencies 1/T, 2/T, ..., 2000/T.
The fourier series then only must contain a finite number of frequencies from the fundamental frequency f_1 = 1/T up to the maximum frequency f_{2000} = 2000/T = 10000 Hz, and the frequencies comes in steps of df = 1/T = f1.

In general sampling with frequency SF in the total time T implies:
N = SF*T data points, N equations each with a sum of ½N parts
and frequencies from 1/T to max. ½N/T = SF/2 in steps of 1/T.

After solving for all the unknowns a_n og b_n the amplitudes c_n (and the so-called phase shift φ_n) can be calculated. The resulting spectrum then can be described either by c_n as a function of the whole number n, or by c_n as a function of the frequency $f_n = n/T$, with n from 1 til ½N.
Sometimes instead of c_n the square c_n^2 will be applied this being proportional with the energy.