

# CONSEQUENCES OF NYQUIST THEOREM FOR ACOUSTIC SIGNALS STORED IN DIGITAL FORMAT

Marek Roland - Mieszkowski and Wayne R. Young  
School of Human Communication Disorders, Dalhousie University  
5599 Fenwick Street, Halifax, Nova Scotia, B3H-1R2, Canada.

The calculation of functions in digital domain from analogue acoustic signals involves a two-step process which includes analog to digital (A/D) conversion and digital calculations performed on digitized acoustical signal. Function  $V(t)$  will be represented digitally without any loss of information as long as sampling occurs in accordance with the Nyquist criteria [1-9]. How can we determine the values of digitized function for points between samples when we have only  $N$  samples available? The Nyquist formula requires an infinite number of samples to accomplish this task. In situations when digital samples are sufficiently dense, one can approximate many continuous functions with their discrete formulations. Errors generated in these cases will be small, since they depend on spacing between samples. The situation will be different, however, when samples are coarsely spaced. For example, a sinusoidal tone of frequency  $f=20,000$  Hz sampled at  $f_{sampling}=44,100$  Hz is represented by only 2.205 samples per period. Calculation of many functions (for example RMS values) may lead in this case to some errors. Finite duration sampling of continuous signal results in errors caused by our limited knowledge of the function for all points in time. It turns out that the more samples we have around the region of interest, the more accurately we are able to reproduce the function there. This paper investigates the error caused by truncation of the Nyquist sampling formula with the aim of quantifying it and establishing ways to minimize its effect.

## Nyquist Theorem

According to the Nyquist theorem [1-9] the discrete time sequence of a sampled continuous function  $\{ V(t=n \cdot T_s) \}$  contains enough information to reproduce the function  $V=V(t)$  exactly provided that the sampling rate ( $f_s=1/T_s$ ) is at least twice that of the highest frequency contained in the original signal  $V(t)$ :

$$V(t) = \sum_{n=-\infty}^{+\infty} V[n] \cdot \frac{\sin [\pi \cdot f_s \cdot (t - n \cdot T_s)]}{\pi \cdot f_s \cdot (t - n \cdot T_s)} \quad (1)$$

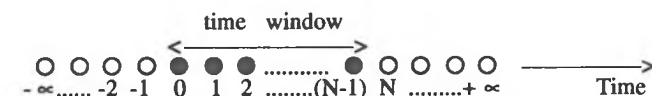
where:

$$\begin{aligned} f_s &= 1 / T_s - \text{sampling frequency} \\ V(t) &- \text{value of signal (voltage) at arbitrary time } t \\ V[n] &= V(n \cdot T_s) - \text{value of signal at time } t = n \cdot T_s \end{aligned}$$

## Nyquist Theorem's consequences

It is worth noting that information about the signal  $V=V(t)$  at any given moment in time  $t \neq n \cdot T_s$  is distributed among all discrete samples  $\{ V[n] \}$  with appropriate weights ( see eq. (1) ). Realistically, we are never presented with an infinite discrete time sequence and are therefore forced to perform the summation over a finite range. This is equivalent to a loss of information about the function  $V=V(t)$  not only before and after our time window (which is understandable), but also at the time points between the sampling points. This can introduce errors into the process of reconstructing the function. Let assume that we have available to us  $N$  digital samples of function  $V=V(t)$  (this is illustrated in fig. 1):

$$V[n] = V(t=n \cdot T_s) \quad \text{where: } n = 0, 1, 2, \dots, (N-1) \quad (2)$$



● - known and ○ - unknown values of function  $V[n]$

Fig.1. Illustration of available values  $V[n]$  in the time window  $[0, (N-1) \cdot T_s]$

Values of the function  $V=V(t)$  for the times  $t \in [0 \cdot T_s ; (N-1) \cdot T_s]$ , can be estimated by a truncated version of formula (1):

$$V(t) = \sum_{n=0}^{N-1} V[n] \cdot \frac{\sin [\pi \cdot f_s \cdot (t - n \cdot T_s)]}{\pi \cdot f_s \cdot (t - n \cdot T_s)} \quad (3)$$

The errors resulting from truncation are  $\epsilon_{LEFT}(t)$  and  $\epsilon_{RIGHT}(t)$ . They represent the "LEFT" and "RIGHT" portions of the sum (with respect to the time axis) in eq. (1) that are omitted in eq.(3), and can be represented mathematically by the following formulas:

$$\epsilon_{LEFT}(t) = \sum_{n=-\infty}^{-1} V[n] \cdot \frac{\sin [\pi \cdot f_s \cdot (t - n \cdot T_s)]}{\pi \cdot f_s \cdot (t - n \cdot T_s)} \quad (4 a)$$

$$\epsilon_{RIGHT}(t) = \sum_{n=N}^{+\infty} V[n] \cdot \frac{\sin [\pi \cdot f_s \cdot (t - n \cdot T_s)]}{\pi \cdot f_s \cdot (t - n \cdot T_s)} \quad (4 b)$$

The sum truncation error is generated when eq.(3) is used instead of eq.(1), and is given by the following formula:

$$\epsilon_{TOTAL}(t) = \epsilon_{LEFT}(t) + \epsilon_{RIGHT}(t) \quad (5)$$

where:

$\epsilon_{LEFT}(t)$  - "LEFT" error ( due to truncation to the left where values of  $V[n]$  are unknown)  
 $\epsilon_{RIGHT}(t)$  - "RIGHT" error ( due to truncation to the right where values of  $V[n]$  are unknown)

In the next section we will try to evaluate the sum truncation error for different cases.

## Evaluation of sum truncation error

● If a priori information is given, that  $V(t) = 0$  for  $t \notin [0; (N-1)T_s]$  then  $\epsilon_{TOTAL} = 0$  (since  $\epsilon_{LEFT} = 0$  and  $\epsilon_{RIGHT} = 0$ ) and we can use with full confidence the truncated version of the Nyquist sum ( eq. (3) ). Otherwise using the sum from eq. (3) is equivalent to the zero-extension (or zero-padding) method [3,5].

● If a priori information is given, that function  $V = V(t)$  is periodic ( ie.  $V(t) = V(t+T)$  ) then this information can be used in formula (1). In the special case when the function period  $T = N_0 \cdot T_s$ , where  $N_0 \leq N$ , one can use formula (1) directly since all values of function  $V[n]$  are known for  $n \in (-\infty ; +\infty)$ . In this case one can also use the Discrete Fourier Transform ( DFT) on  $N_0$  consecutive data points (from the available  $N$  data points  $V[n]$  where  $n=0, 1, 2, \dots, (N-1)$  ) to calculate amplitudes  $A_i$ , and phases  $\phi_i$ , of the periodic signal and then use the formula:

$$V(t) = A_0 + \sum_{i=1}^M A_i \cdot \sin (2 \pi f_i \cdot t + \phi_i) \quad (6)$$

where:

$A_0$  - DC component of the periodic signal  $V = V(t)$

$A_i$  - amplitudes calculated from DFT

$\phi_i$  - phases calculated from DFT

$f_i = (i / N_0) \cdot f_s$  - frequencies available in periodic signal  $V = V(t)$

$M = \text{INT} ( N_0 / 2 )$  - number of the highest possible harmonic (since  $f_i < f_s / 2$ )

● When a priori information is not available about the function  $V=V(t)$  then direct use of the truncated Nyquist sum ( eq. (3) ) is going to lead to truncation error  $\epsilon_{TOTAL}$  given by eq. (4) and (5). Values  $V[n]$  in eq. (4 a) and (4 b) can be any arbitrary numbers, since we do not have any a priori information about the signal. In the next section we will investigate this in greater detail to estimate values of possible errors.

## Estimation of truncation error for general case

Let's consider for simplicity the "LEFT" error given by eq. (4 a) (estimation of "RIGHT" error is performed in identical way). Total error given by eq. (4 a) is a sum of contributions from data points  $V[n]$  where  $n=-1, -2, -3, \dots, -\infty$ .

Error contribution from the  $n$ -th point is given by:

$$\epsilon_n(t) = V[n] \cdot \frac{\sin[\pi \cdot f_s \cdot (t - n \cdot T_s)]}{\pi \cdot f_s \cdot (t - n \cdot T_s)} \quad (7)$$

where:  
 $t \in [0; (N-1) \cdot T_s]$   
 $n = -1, -2, -3, \dots, -\infty$  for "LEFT" errors  
 $n = N, (N+1), \dots, +\infty$  for "RIGHT" errors

The function  $\sin(x)/x$  (where  $x = \pi \cdot f_s \cdot (t - n \cdot T_s)$ ) is equal to 0 at time points  $t = m \cdot T_s$ ;  $m = 0, 1, 2, 3, \dots, (N-1)$  (which are the sampling points in our time window). Therefore there is no contribution to the error in  $V(t)$  at sampling points due to truncation. Also the function  $\sin(x)/x$  has local max. and min. approximately at the middle points between adjacent sampling points in our time window. This can be proved easily by taking first derivative of function  $\sin(x)/x$ . The approximation gets better for larger values of  $x$ . Therefore from now on we will consider the error at the middle points between samples  $V[n]$  in our time window (see fig.1):

$$t = (m + 1/2) \cdot T_s \quad (8)$$

Substituting time  $t$  from eq (8) into formula (7) we get:

$$\epsilon_n[m + 1/2] = V[n] \cdot \frac{(-1)^{(m-n)}}{\pi \cdot [(m + 1/2) - n]} \quad (9)$$

where:  
 $m = 0, 1, 2, 3, \dots, (N-2)$  - indexing of middle points between adjacent samples in the time window  
 $n = -1, -2, -3, \dots, -\infty$  - indexing of points on the left of the time window (see fig.1)

● An interesting question to ask is how large the value of index "m" in eq.(9) must be in order for the absolute value of error  $\epsilon_n[m+1/2]$  to be equal or smaller than 1/2 of the quantization step  $\Delta$ , which is defined as a difference between quantization levels [1,3,8,9]. This is a reasonable comparison since all samples  $V[n]$  have quantization error inherent to the process of A/D conversion [1,3,8,9]. Quantization error is uniformly distributed in the range  $[-\Delta/2; \Delta/2]$ , where  $\Delta$  is a step size in the A/D converter [1,3,8,9]. From eq. (9) we have for  $n=-1$  and  $V[-1] = V_{MAX} = 2^{B-1} \cdot \Delta$  (where  $V_{MAX}$  is the max. possible signal amplitude in the A/D converter and  $B$  is the No. of bits in the A/D converter):

$$m \geq 2^B / \pi - 3/2 \quad (10)$$

From inequality (10) we get (times are calculated for  $f_s = 44,100$  Hz):

for B = 8 bit : $m \geq$	80	( 0.001 sec	inside sound file)
for B = 10 bit : $m \geq$	325	( 0.007 sec	inside sound file)
for B = 12 bit : $m \geq$	1,303	( 0.030 sec	inside sound file)
for B = 14 bit : $m \geq$	5,214	( 0.118 sec	inside sound file)
for B = 16 bit : $m \geq$	20,860	( 0.473 sec	inside sound file)
for B = 18 bit : $m \geq$	83,442	( 1.892 sec	inside sound file)
for B = 20 bit : $m \geq$	333,771	( 7.569 sec	inside sound file)

Results in (11) show that if we have no information about the signal before our time window, then in order to avoid errors associated with the unknown values of function  $V[-1], V[-2], V[-3], \dots$ , one has to be "m" samples deep inside of the time window. We can then use the formula (3) for any time  $t$  as long as we stay away from the ends of the time window by "m" samples ( $t \in [m \cdot T_s; (N-1-m) \cdot T_s]$ ).

● Another interesting question to ask is which sequence of samples  $V[-1], V[-2], V[-3], \dots$  will generate the largest error at the middle points in our time window? Taking the summation of eq. (9) from  $n=-1$  to  $m=-\infty$  we get:

$$\epsilon_{LEFT}[m + 1/2] = 1/\pi \cdot (-1)^m \cdot \sum_{n=-1}^{-\infty} V[n] \cdot \frac{(-1)^{-n}}{(m + 1/2) - n} \quad (12)$$

If we take the sequence of samples  $V[n] = V_{MAX} \cdot (-1)^n$  then we get from eq (12):

$$\epsilon_{LEFT}[m + 1/2] = 1/\pi \cdot V_{MAX} \cdot (-1)^m \cdot \sum_{n=-1}^{-\infty} \frac{1}{(m + 1/2) - n} \quad (13)$$

Unfortunately our choice of  $\{V[n]\}$  in eq. (13) was inappropriate because the sum of this series diverges to  $\infty$ . As a matter of fact this happened since this particular series corresponds to the digital representation of a sine wave with the Nyquist frequency  $f = f_s/2$ . Such a frequency can't exist in the digital domain, since it can't be

recorded via A/D conversion. A realizable choice of samples would be, for example, one which represents sinusoidal wave with frequency  $f < f_s/2$ :

$$V[n] = V_{MAX} \cdot \sin[2 \cdot \pi \cdot f \cdot n \cdot T_s + \phi] \quad (14)$$

where:  
 $\phi = \pi/2 + 2 \cdot \pi \cdot f \cdot T_s$  - phase angle chosen to maximize error for  $n = -1$  ( $V[n=-1] = V_{MAX}$ )

Substituting (14) in eq. (12) we get:

$$\epsilon_{LEFT}[m+1/2] = 1/\pi \cdot V_{MAX} \cdot (-1)^m \sum_{n=-1}^{-\infty} \sin[2 \cdot \pi \cdot f \cdot n \cdot T_s + \phi] \cdot \frac{(-1)^{-n}}{(m+1/2) - n} \quad (15)$$

Computer calculations of error were performed using formula (15) for  $m=1000$  (calculations of error in middle point between 1000 and 1001 sample in time window). Results were as follows:

for $f = 0.250 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 1.45 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.300 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 1.45 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.350 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 1.45 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.400 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 1.45 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.450 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 1.48 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.490 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 2.24 \cdot 10^{-4} \cdot V_{MAX}$
for $f = 0.499 \cdot f_s$ :	$S(1) = 3.18 \cdot 10^{-4} \cdot V_{MAX}$ ;	$S(10000) = 7.23 \cdot 10^{-3} \cdot V_{MAX}$

where:  
 $S(1)$  represents first term in eq.(15) (contribution from the  $V[-1]$ )  
 $S(10000)$  represents summation of first 10000 terms in eq. (15)

Results (16) show, that max. error is obtained when frequency  $f$  approaches the Nyquist frequency  $f_s/2$ . The value of the error from the first term ( $V[-1]$ ) is larger than from the sinusoidal wave as long as frequency  $f$  is smaller than about  $0.490 \cdot f_s$  (which is 98% of Nyquist frequency =  $f_s/2$ ). This seems reasonable since the input antialiasing filter would usually eliminate all frequencies above  $0.46 f_s$  (see for example data for PCM or DAT recorders). The errors resulting from sinusoidal waves are similar to the error contribution from  $V[-1]$ , therefore we expect that the recommendation given in (11) should be valid for any arbitrary sequence  $V[-1], V[-2], V[-3], \dots$ , since such sequence is a linear combination of pure tones according to the Fourier Theorem. However, further investigation and computer simulations are required to substantiate this claim.

## Conclusions

In this paper we investigated the errors due to finite duration sampling of continuous signal and determined that this error can be considerable at the beginning and near the end of the sampling time window. These errors had a tendency to get larger at higher frequencies as they approach the Nyquist frequency ( $f_s/2$ ) for signal near the inside boundaries of the time window. At this time however, we don't know which physically realizable sequence of samples  $V[-1], V[-2], V[-3], \dots$  will produce the largest error inside of the time window. Further tests and computer simulations are required.

## References :

- [1] Nakajima et al., "The Sony Book of Digital Audio Technology", published by TAB Books Inc., 1983.
- [2] Papoulis Athanasios, "Circuits and Systems, A Modern Approach", Holt, published by Rinehart and Winston Inc., 1980.
- [3] Oppenheim Alan V., Schafer Ronald W., "Discrete-Time Signal Processing", published by Prentice Hall Inc., 1989.
- [4] Jayant N. S., Noll Peter, "Digital Coding of Waveforms, Principles and Applications to Speech and Video", published by Prentice-Hall Inc., 1984.
- [5] Jackson Leland B., "Digital Filters and Signal Processing", published by Kluwer Academic Publishers, 1989.
- [6] Press William et al., "Numerical Recipes", published by Cambridge University Press, 1987.
- [7] Weaver H. Joseph, "Applications of Discrete and Continuous Fourier Analysis", published by John Wiley & Sons, 1983.
- [8] Marek Roland-Mieszkowski, "Introduction to Digital Recording Techniques", Proceedings from "Acoustic Week in Canada 1989" - CAA Conference, Halifax, N.S., Canada, October 16-19, 1989, 73-77.
- [9] Robert Wannamaker, Stanley Lipshitz and John Vanderkooy, "Dithering to eliminate quantization distortion", Proceedings from "Acoustic Week in Canada 1989" - CAA Conference, Halifax, N.S., Canada, October 16-19, 1989, pp. 78-86.