

音声情報の表現

アナログ・デジタル

- アナログ情報: 実数によって表現される**連続**的な情報
 - 自然界で観測された物理的な量: 重さ、長さ、時間、空間の位置、電流電圧など
 - 特徴: 連続で切れ目がない。たとえば0と1の間には0.5が、0.0と0.5の間には0.25が、0と0.25の間には0.125が...という具合に、どれだけ近い二つの数(数値)の間にも別の数(数値)が存在
 - このような数値を連続値とよぶ。対して、切れ目のある、連続しないものを離散値とよぶ

アナログ・デジタル

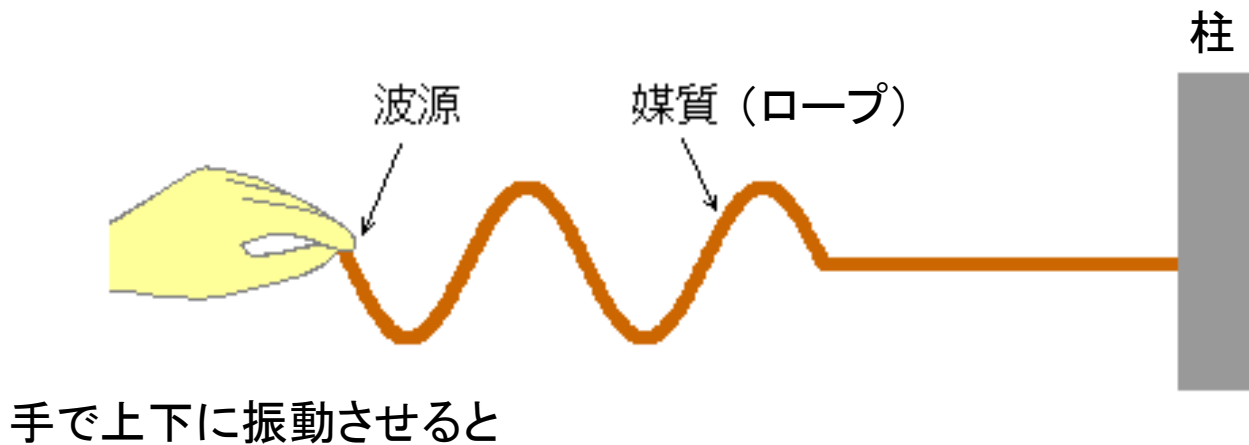
- デジタル情報: 自然数(0以上の整数。これが離散値)で表すことの可能な、**離散**的な情報
 - 個数、番号、名前、文字など
 - 自然数は2進数に変換可能なので、すべてのデジタル情報はコンピュータで扱うことが可能
- 離散化: 連続値を離散値で表現すること

Question

- アナログ情報・デジタル情報の例を挙げてみよう。

音

- 音は空気の振動で、波である
- 波（波動）とは、波源（音であれば音源）での振動が媒質（音であれば空気）を介して伝わっていく現象

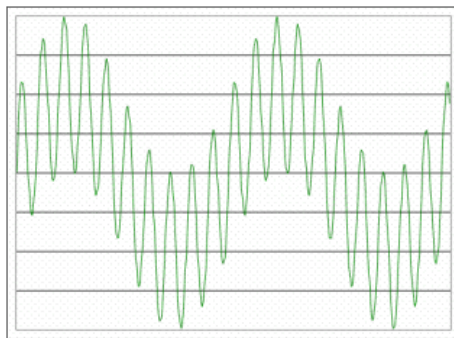
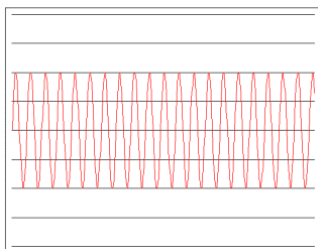
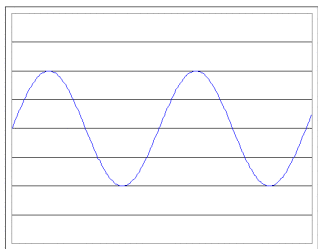


- 観察点を固定して、その観察点での振動の時間変化の記録を取っておくと、同じ形の波形が得られる
- 波源が音声であればこのような波形が音声波形となる

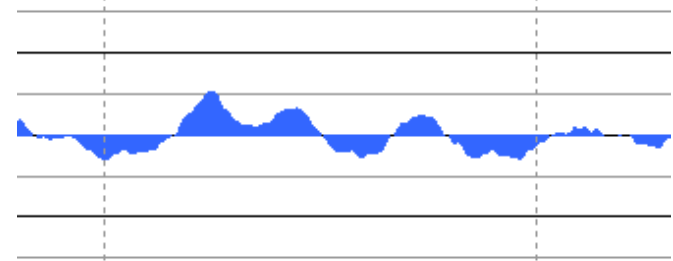
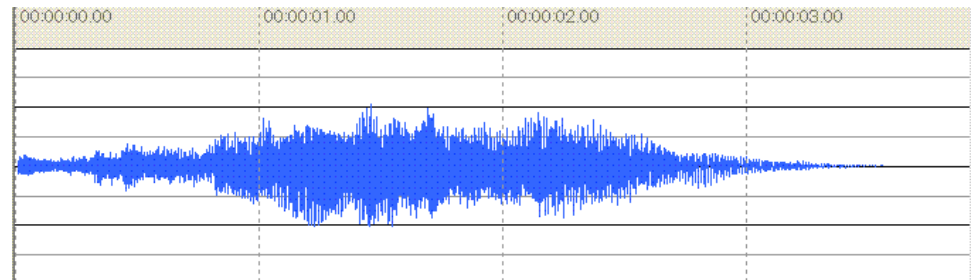
音

- 音声波形は音を特徴づける三要素を含む。その三要素とは
 - 音の大きさ: 振幅
 - 音の高さ: 周期あるいは周波数
 - 音色: 波の形 (正弦波・矩形波・三角形波・のこぎり波など)

- 音声波形の例



二つ純音の
合成波形

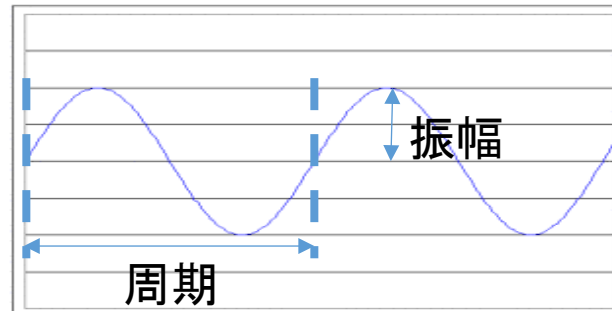


Windowsの起動音: 多数の周波数の合成波形 (下図は一部拡大)

純音

周期・周波数・振幅

- 周期: 波の繰り返し1つ分の所要時間
- 周波数: 波の単位時間あたりの繰り返しの回数
 - 周期の逆数。単位はヘルツ(Hz)
 - Hz(ヘルツ): 周波数・振動数の単位。「1Hz=1秒間に1回の周波数・振動数」で定義されている。つまり、たとえば50Hzは1秒間に50回波が繰り返されることを意味
 - これは国際単位系の単位で、キロヘルツ(kHz)やメガヘルツ(MHz)、ギガヘルツ(GHz)などもある
- 振幅: 波の振れ幅のことで波の一番高い地点から低い地点の差の半分

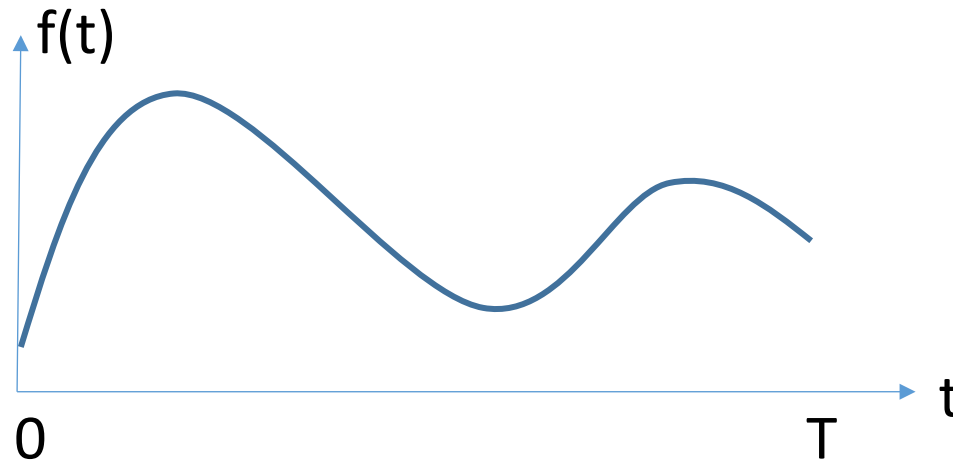


音声のデジタル化

- 音声波形(信号)はアナログ情報である。アナログ波形ともいう
- 横軸(時間)も縦軸(振幅)も連続値なので、その両方を離散化しなければならない
- 一般的に、横軸が表す量を**変数**、縦軸を表す量を**値**といい、変数の離散化を**標本化**(サンプリング)、値の離散化を**量子化**という
- 標本化・量子化後に音声波形をデジタル化(符号化)する

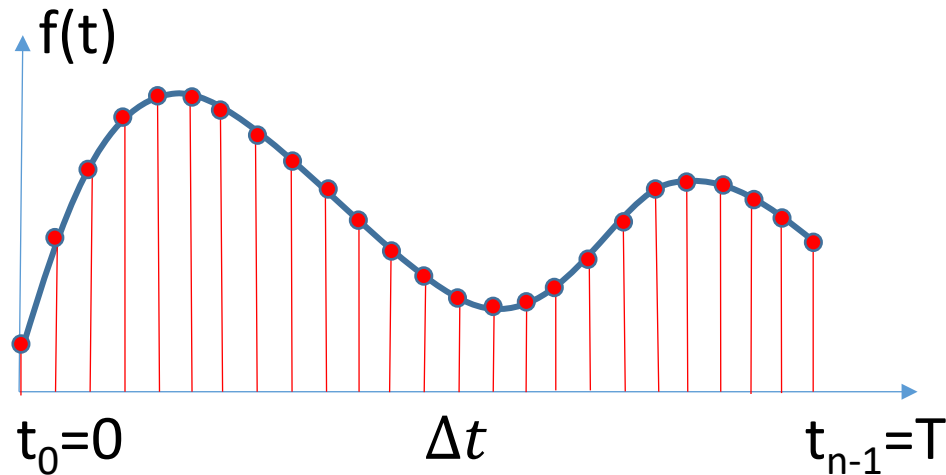
標本化(サンプリング)

- アナログ波形 $y=f(t)$ は、 $0 \leq t \leq T$ を満たすすべての実数 t に対して値をとる連続な関数であるため、波形値 $f(t)$ が無限に存在し、コンピュータでは扱えない



標本化(サンプリング)

- そこで、横軸に対して n 個の点 t_0, \dots, t_{n-1} を等間隔に取り、点上の波形値 $f(t_0), \dots, f(t_{n-1})$ で波形を表現する。この操作を**標本化**と呼ぶ。これらの点を**標本点**、間隔 (Δt) を**標本間隔**と呼ぶ



標本化(サンプリング)

- 横軸が時間の場合、標本間隔を**標本化周期**という。たとえば、標本化周期が10ms(ミリ秒)などである
- また、その逆数が**標本化周波数**である。これは「**秒あたりに標本を何個取得するか**」を表す。例えば上記の例では $1/10\text{ms}=1/(10\times 10^{-3})=100$ で、つまり、標本化周波数が100Hz(ヘルツ)である(1秒間に100個の標本を取る)

演習04-1

- x 秒間の音声データを y kHzの標本化周波数でサンプリングするとして、計何個の標本データ(標本点)が得られるかをもとめる問題。また、標本化周期も求める
- manaba, 8分

標本化定理

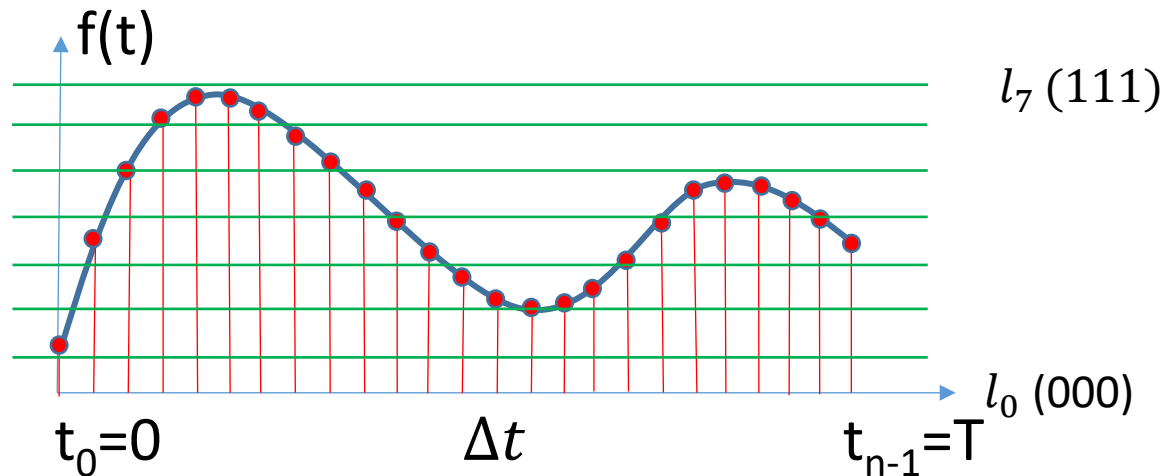
- 標本化周期 Δt を小さくするほど元の波形を忠実に表現することができるが、その分標本点の数が増えデジタル後のデータ量が増大する
- シャノンの**標本化定理**: 入力波形に含まれる最大の周波数成分が W (Hz)としたとき、標本化周波数が $2W$ 以上 (つまり標本化周期が $1/(2W)$ 以下)であれば、標本化されたデータから元の入力波形を完全に復元できる
 - これが適切な標本化周波数・標本化周期を決める際の目安となる

演習04-2

- 標本化定理を用いて最大周波数を求める問題
- manaba, 5分

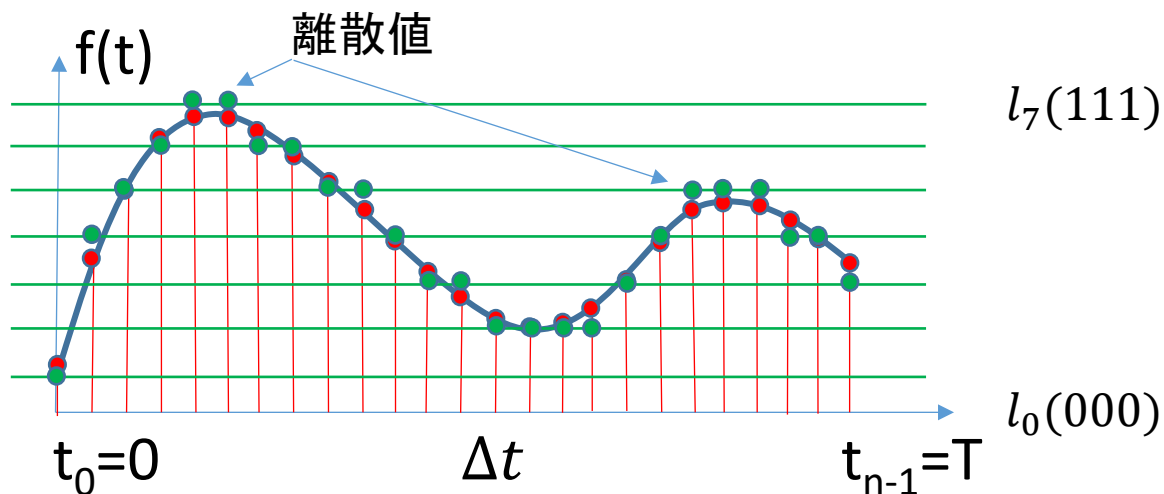
量子化

- 標本化で得られた波形値 $f(t_0), \dots, f(t_{n-1})$ は一般に実数値を取るが、コンピュータでは無限の精度を持つ実数値を扱うことができない
- このため、離散化处理、つまり、各波形値を有限精度で近似的に表現する必要がある
- たとえば3ビットで波形値を表現する場合は、まず波形値の取りえる値を8レベルに分割し、各レベルの代表値 l_0, l_1, \dots, l_7 を定める



量子化

- 波形値 $f(t_0), \dots, f(t_{n-1})$ をおのおのが最も近い代表値に置き換える。
これを(波形値の)丸め処理という



- このような離散化処理を量子化という。波形値を分割するレベル数を量子化レベルという
- 一般に n ビット精度で量子化を行う場合、量子化レベルは 2^n となる。この n を量子化ビット数ともいう

量子化の例

- 波形値の取りうる範囲が0.0-15.0(実数)の波形を2ビットで量子化を行うとする。波形値(1.2, 8.0, 14.8, 6.5, 3.6)の量子化後の離散値(10進数)をもとめなさい

量子化の例

- $n=2$ ビットなので、量子化レベルは $2^2=4$
- 実数範囲が $a-b$ で、量子化レベルが 2^n であれば、レベル間の間隔の数が 2^n-1 で、量子化幅が $(b-a)/(2^n-1)$ の切り上げ $\rightarrow 15/3=5$
 - $0, 0+5=5, 5+5=10, 10+5=15$
- 代表値: $0, 5, 10, 15$ 。つまり、すべての連続値をこの4つの代表値に置き換え(量子化)、この4つの代表値を $00, 01, 10, 11$ の4レベルで表現(符号化)



- $(1.2, 8.0, 14.8, 6.5, 3.6)$



量子化

符号化

- $(0, 10, 15, 5, 5)$



$(00, 10, 11, 01, 01)$

離散値

演習04-3

- n ビット量子化についての問題
- manaba, 10分

量子化

- 量子化は、代表値への丸め処理で誤差が生じる
- 量子化レベルを増やすと元波形に近づくがデータ量が増える
- 量子化は標本化とは異なり不可逆な処理で量子化誤差によって失った情報は元に戻せない
- したがって、デジタル化した波形をどのような用途に使用するかに応じてそれぞれ適切な量子化レベルを設定する
- 一般に、音声(オーディオ)には16ビット精度の量子化が行われている

音のデジタル化の実例: CDの規格

CD-DA(compact disc digital audio)

- チャンネル数: 2
- 標本化周波数: 44.1kHz
- 量子化ビット数: 16bit
- **Question**
 1. 1秒間に記録した音声のデータ量は何Bか
 2. 5分間の曲を記録したデータは何MBか
 3. 容量8GBのフラッシュメモリには2.と同じ長さの曲は最大何曲記録できるか

2チャンネルとは、1つの音声に対して左右2つの独立した音声信号(音声波形)が含まれている形式を指す。この2つの音声信号により、ステレオ効果が生じ、音の位置や広がりを感じることができる