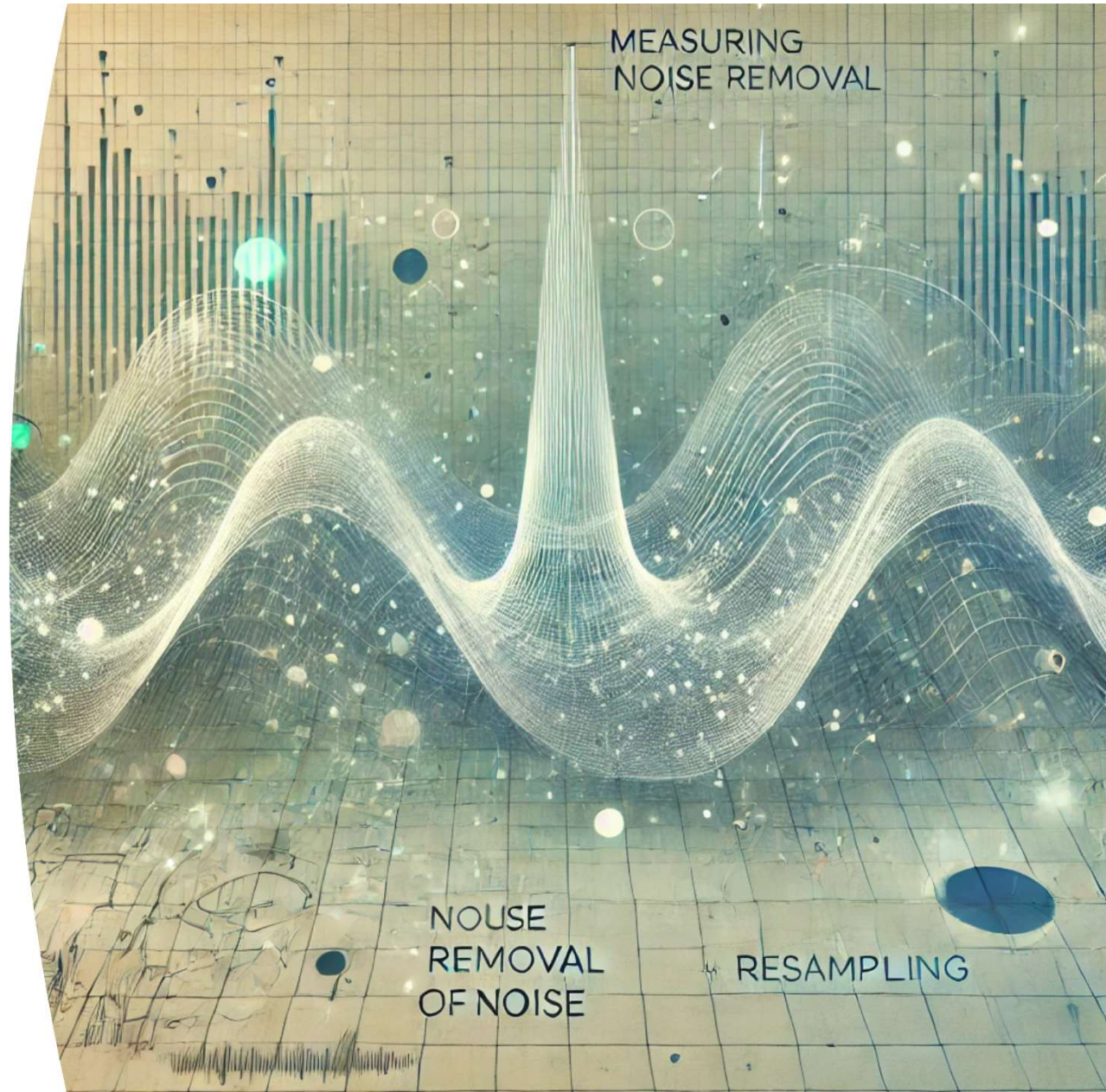


## 主専攻実験第 8 回

### 信号の前処理

- 信号の計測とサンプリング定理
- 周波数成分への変換
  - フーリエ
  - 短時間フーリエ
  - ウェーブレット
- 前処理の目的
  - ノイズ除去
  - リサンプリング
  - 特徴抽出



# 生体信号の計測システム ふたたび

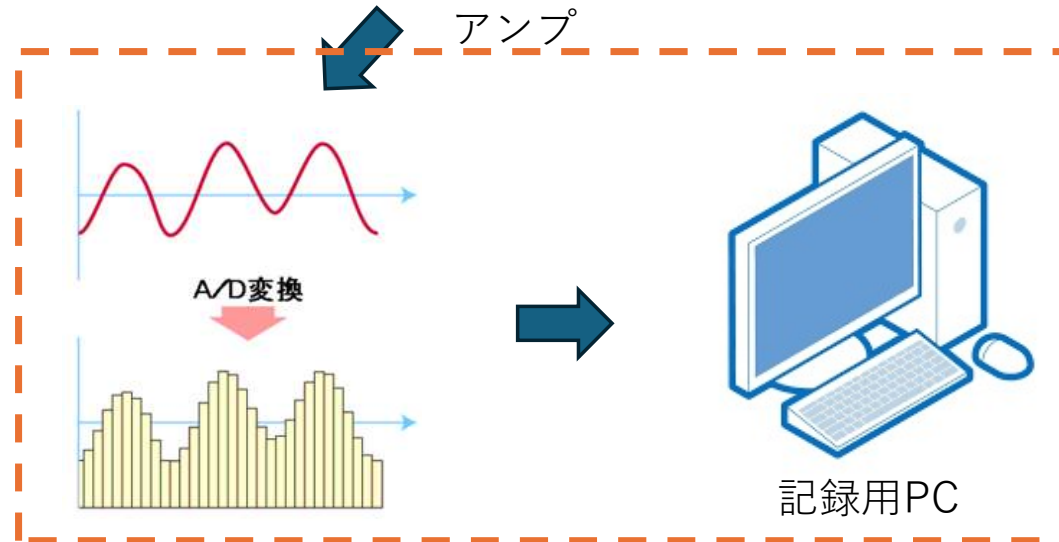
ArtiseBiomedical 



計測電極



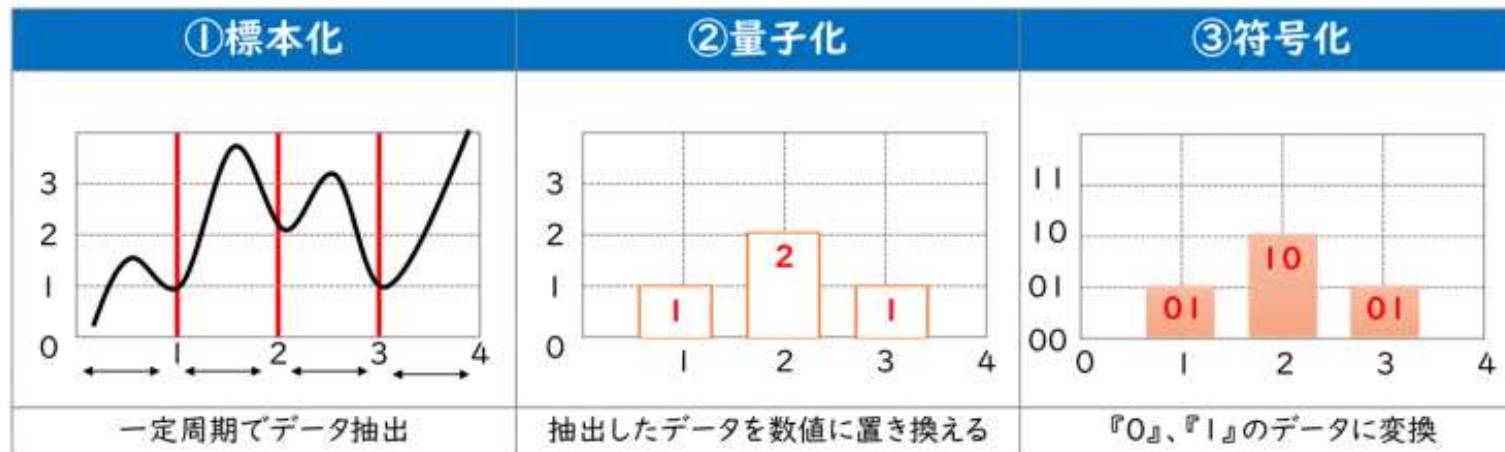
アンプ



医師との連携研究でも、  
AD変換・PCでの  
データ管理は  
情報側の方が詳しい  
→設定をちゃんと確認

# A/D変換

計測機器，アンプから出力される信号は連続値  
→デジタル信号に変換する必要がある。

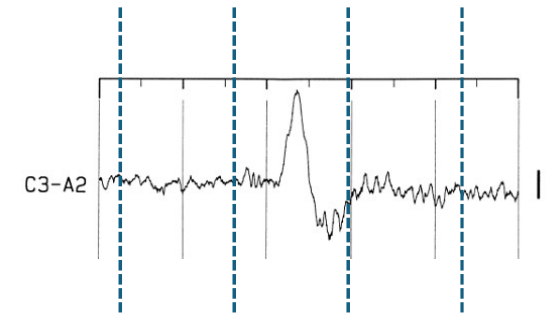


基本的には誤差が発生することになる。  
これ以降の診断・情報処理を行う上での「許容量」を把握，  
許容用以下の誤差になるように，パラメータを調整する必要あり

# 標本化と量子化による誤差

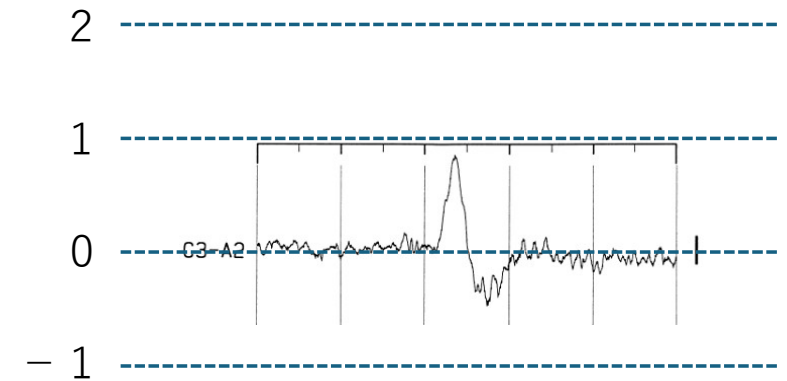
- 標本化による誤差

サンプリングの頻度によっては、  
重要な波形情報を取りこぼす可能性がある  
→サンプリング定理



- 量子化による誤差

アナログ値をデジタル値になるように  
切り捨て、切り上げ、四捨五入  
→「丸め誤差」が発生.



# 標本化定理と実際の運用

標本化定理 (From Wikipedia)

標本化定理は、元の信号をその最大周波数の**2倍**を超えた周波数で標本化すれば完全に元の波形に再構成されることを示す。

# 標本化定理と実際の運用

標本化定理 (From Wikipedia)

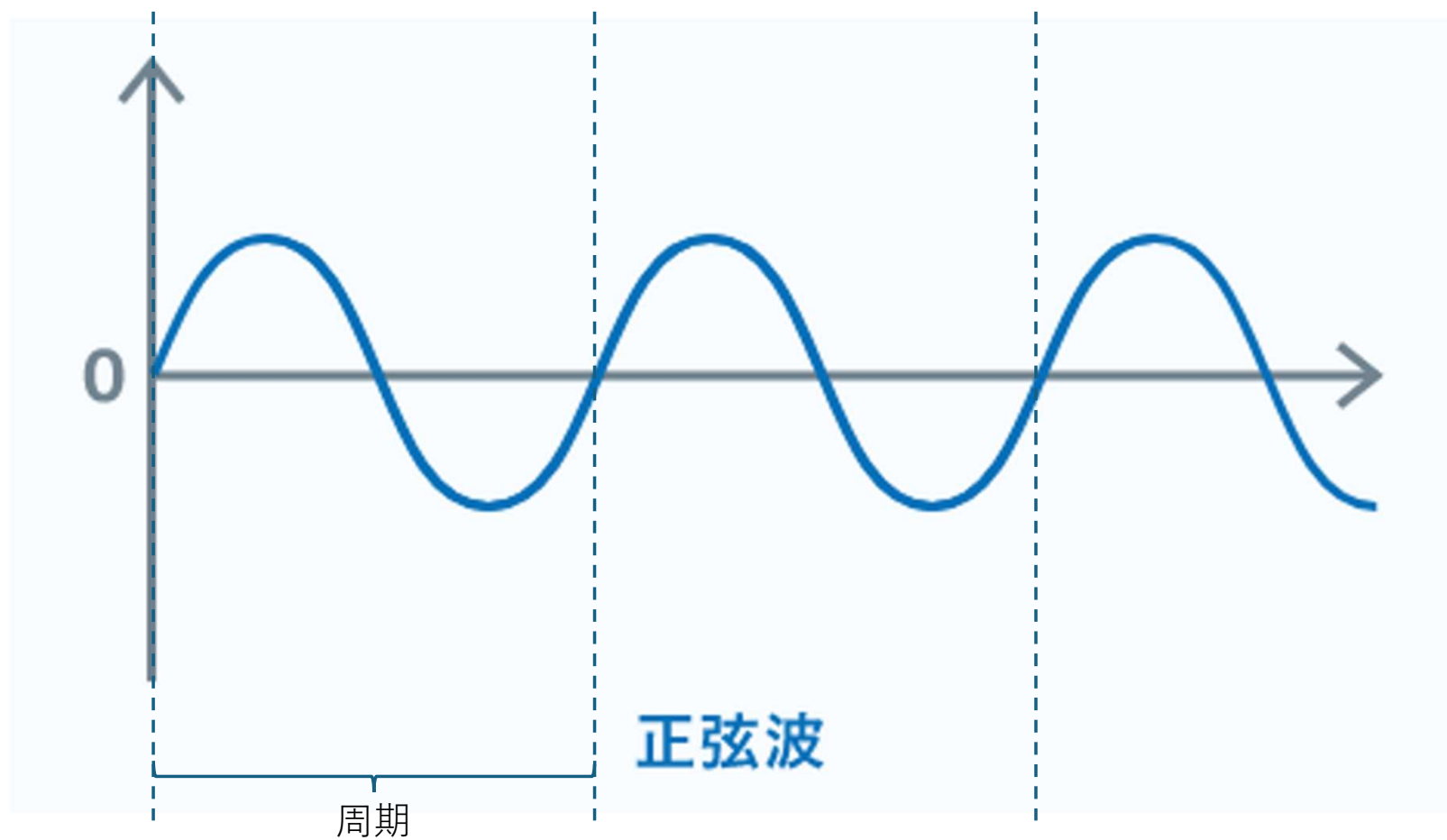
標本化定理は、元の信号をその最大周波数の**2倍**を超えた周波数で標本化すれば完全に元の波形に再構成されることを示す。



ある信号を標本化したとき、標本化周波数の**1/2**以上の周波数の成分は再構成できない  
(エイリアシングの原因となる)

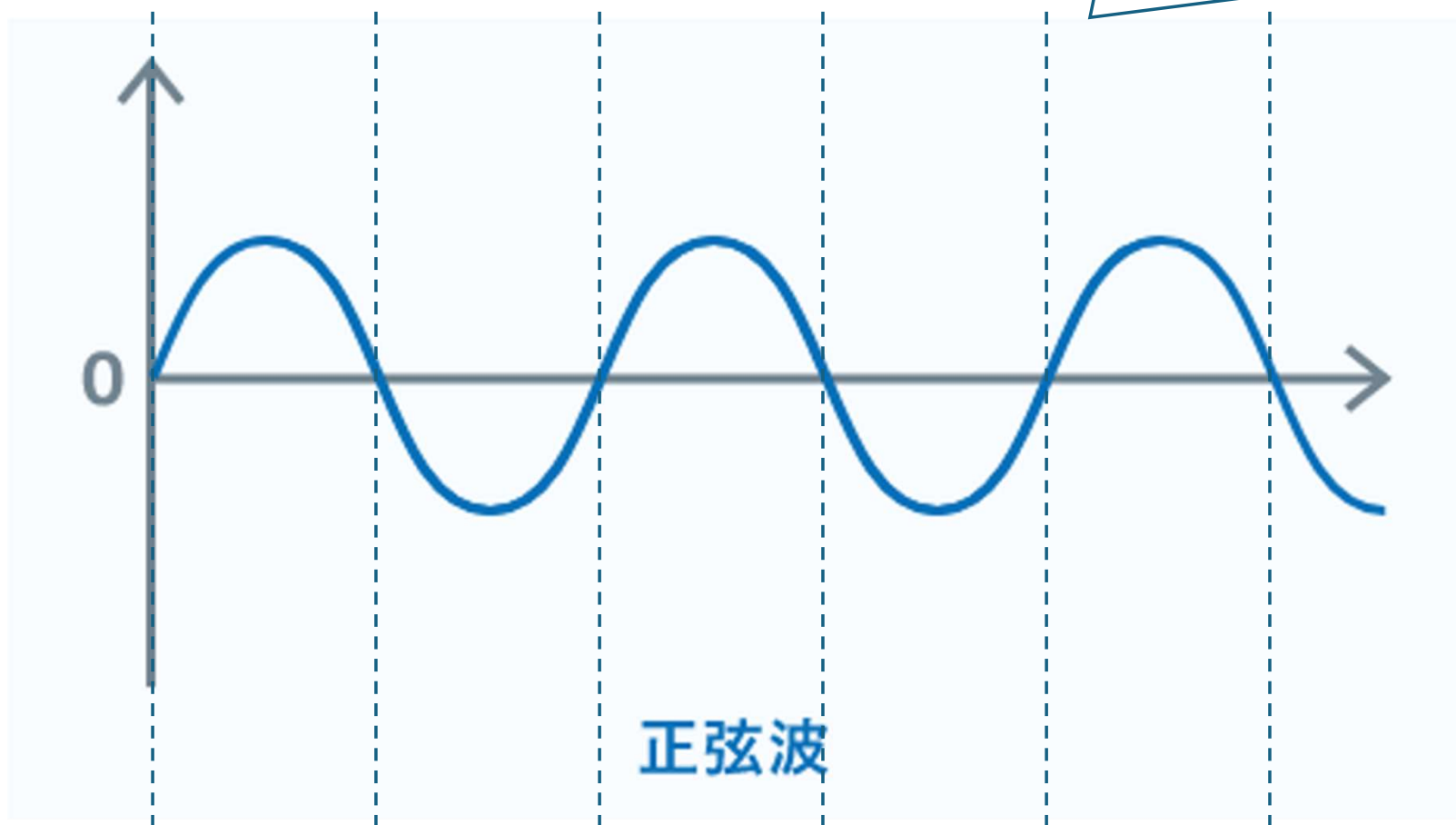


# 標本化定理と実際の運用



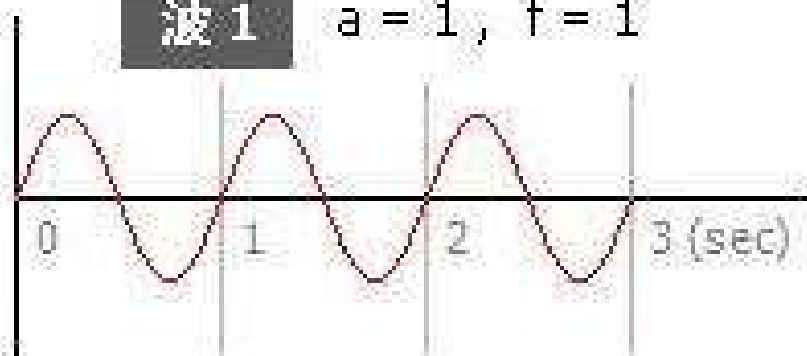
# 標本化定理と実際の運用

元信号の2倍の周波数で  
サンプリングしているのに  
元信号が再現できない

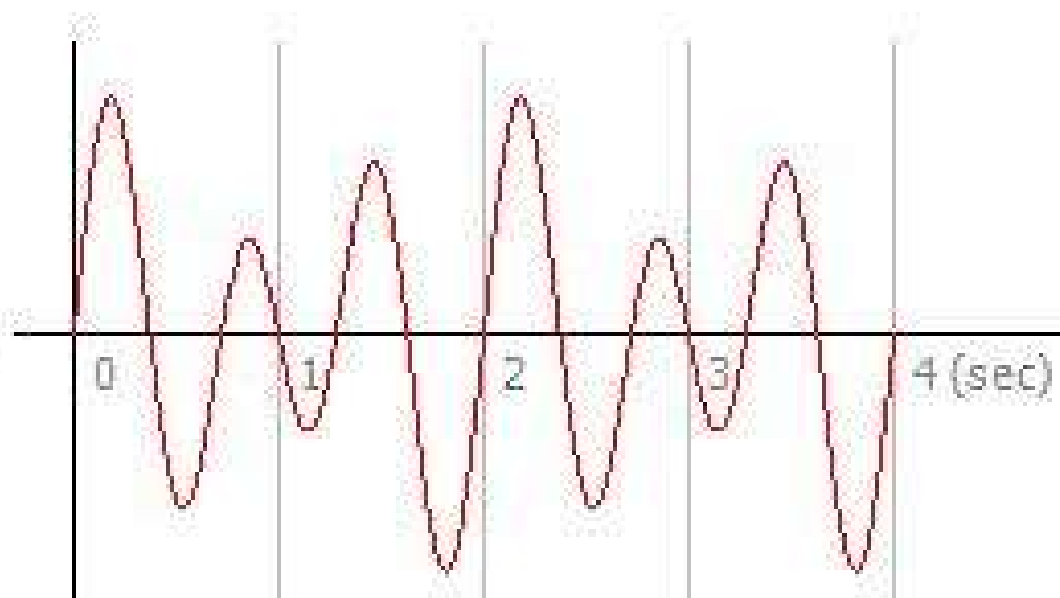
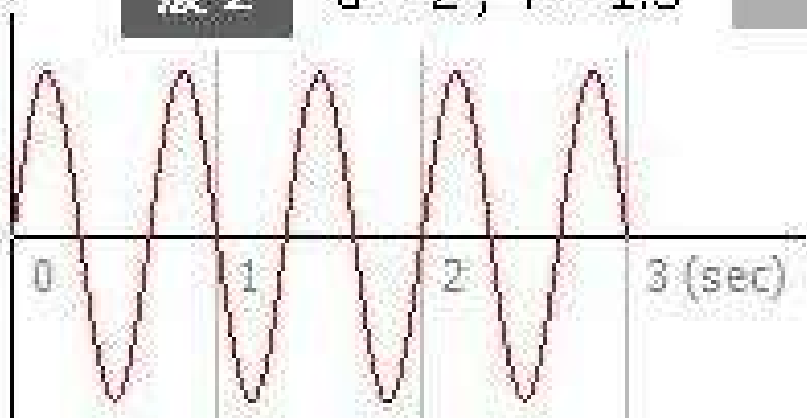


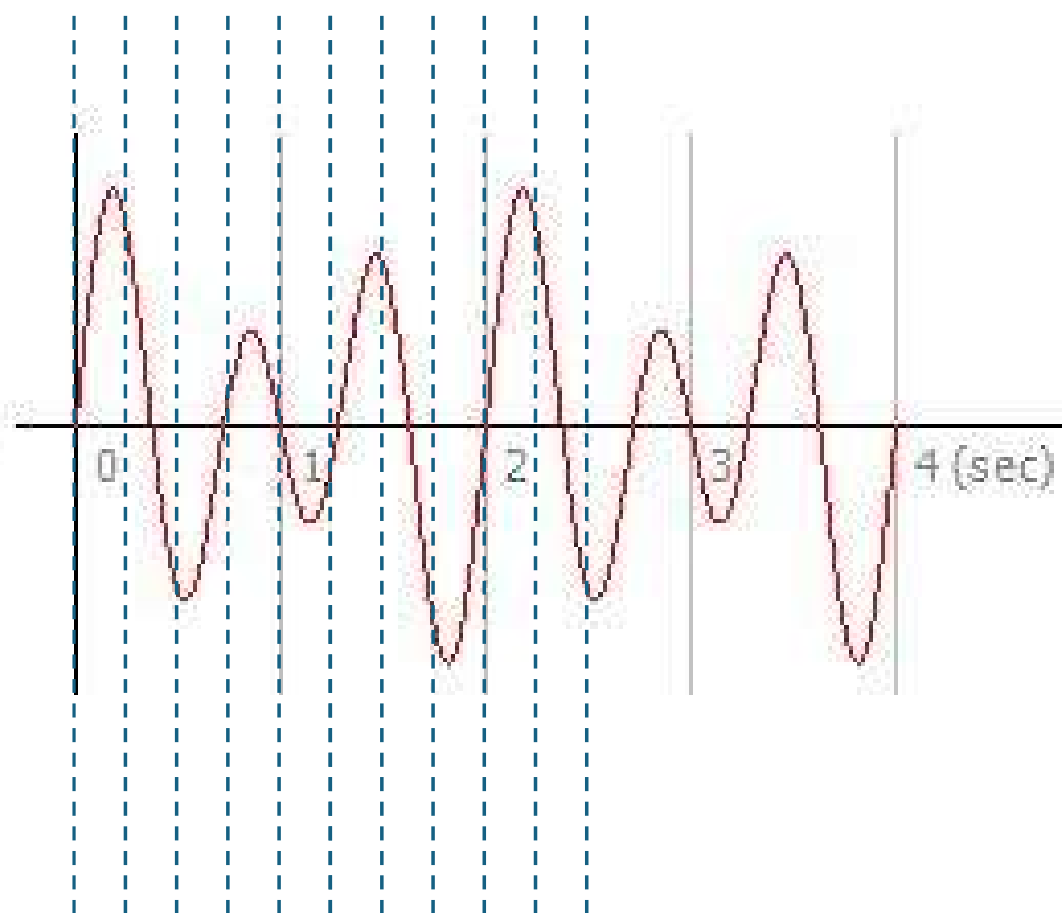
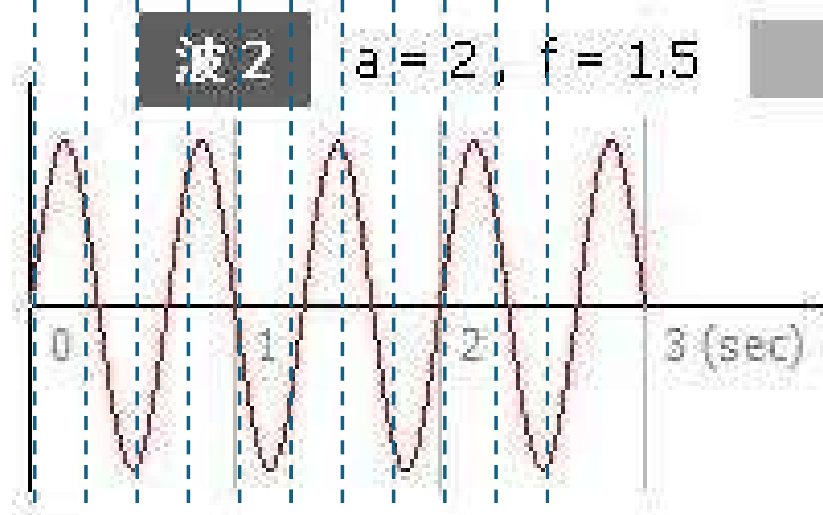
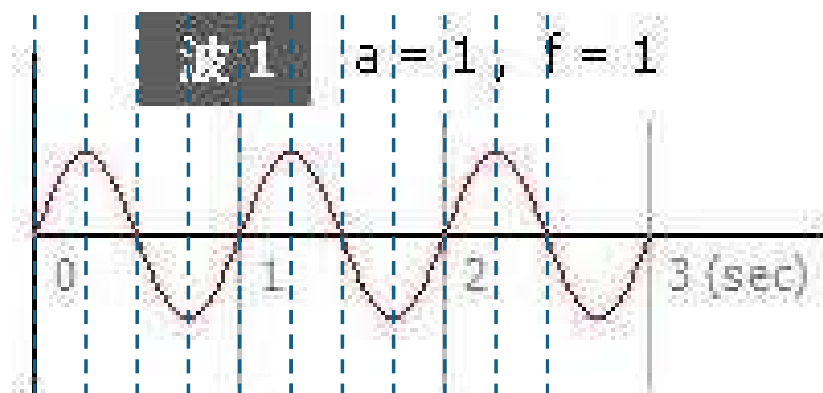


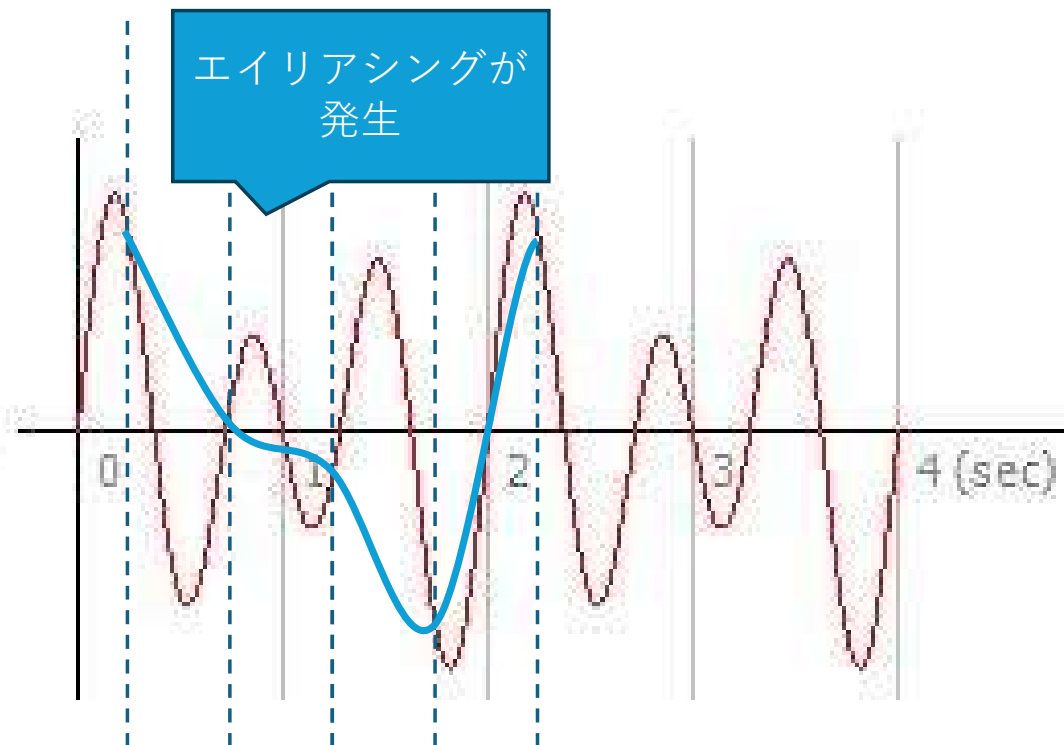
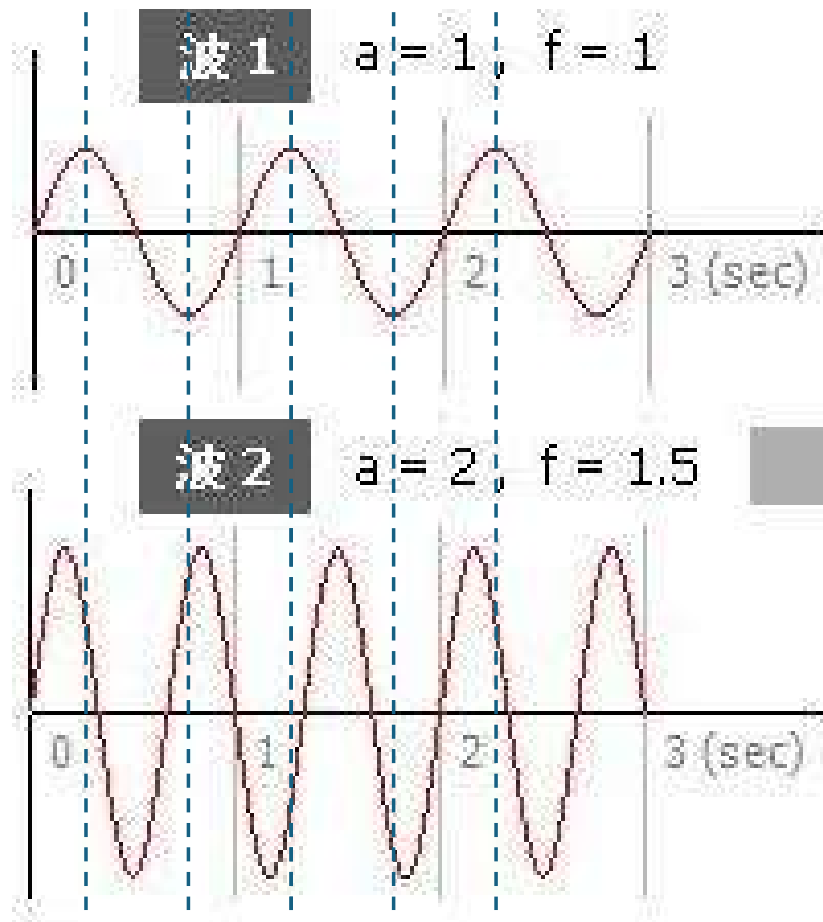
波 1  $a = 1, f = 1$



波 2  $a = 2, f = 1.5$







ある信号を標本化したとき、標本化周波数の $1/2$ 以上の周波数の成分は再構成できない  
(エイリアシングの原因となる)

# 結局

## 何倍でサンプリングすればいいの？

最低 4 倍 数理的な理由付けというより慣習

理想は10倍

可聴域

: 20~20kHz

CD

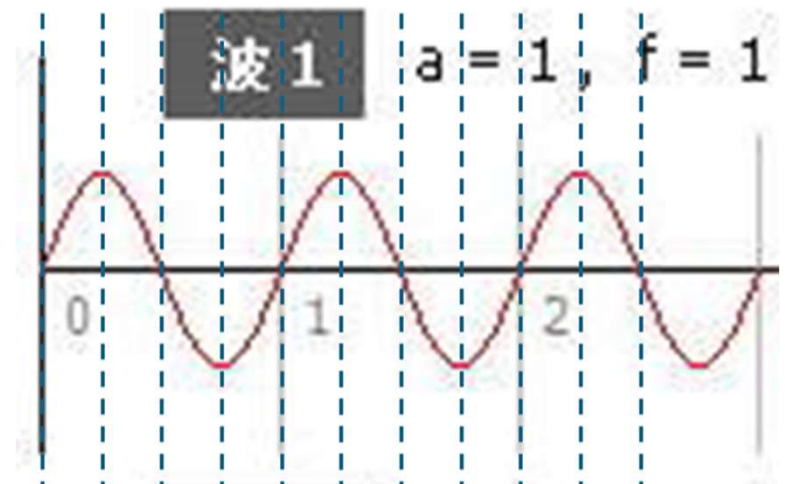
: 44.1kHz

ハイレゾ音源

: 96kHz~

2倍じゃ  
足りず、

4倍以上に



# フィルタリングによる 必要サンプリング周波数の低減

元信号の周波数成分：1~100Hzまで

→ 400Hz以上でサンプリング  
場合によっては不可能な場合も

例) そもそも計測機器が400Hzサンプリング不可能  
400Hzサンプリングだとデータが大きすぎて通信困難

- ・ 必要な周波数成分は30Hz以下
  - 本当に必要な情報は120Hzサンプリングで取得可能
  - ただし、120Hzサンプリングだと  
元信号の高周波成分によるエイリアシングが発生

# 生体電位の周波数（おさらい）

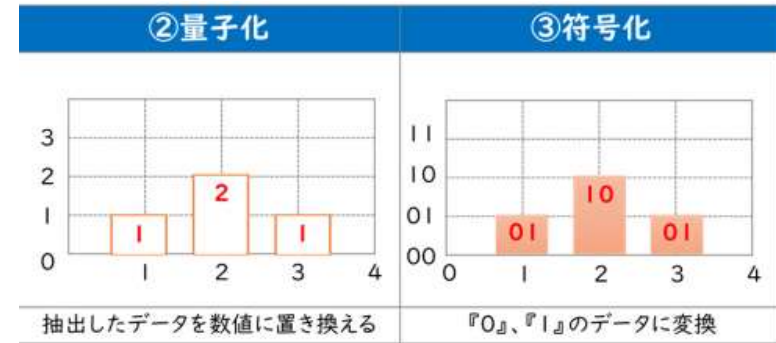
- EEG 0.5 ~ 30 Hz （40Hz以上はステージと無関係）
- EOG 0.1 ~ 10 Hz
- EMG 筋肉次第  
顎の下の方の筋肉だと100 Hz以下  
（顎舌骨筋、顎二腹筋、舌骨下筋群、茎突舌骨筋）

睡眠ステージ判定では、高周波成分をカットした上、  
200Hz程度でサンプリングするケースが多い  
# EMGは周波数成分よりも振幅が重要なので、  
100Hz以上をカットしてサンプリングするケースが多い

# 量子化 & 符号化

- 量子化分解能：  
符号化の際に何bit使用するか？  
多いほど量子化時に細かく分割する  
例) 16bit → 65536分割

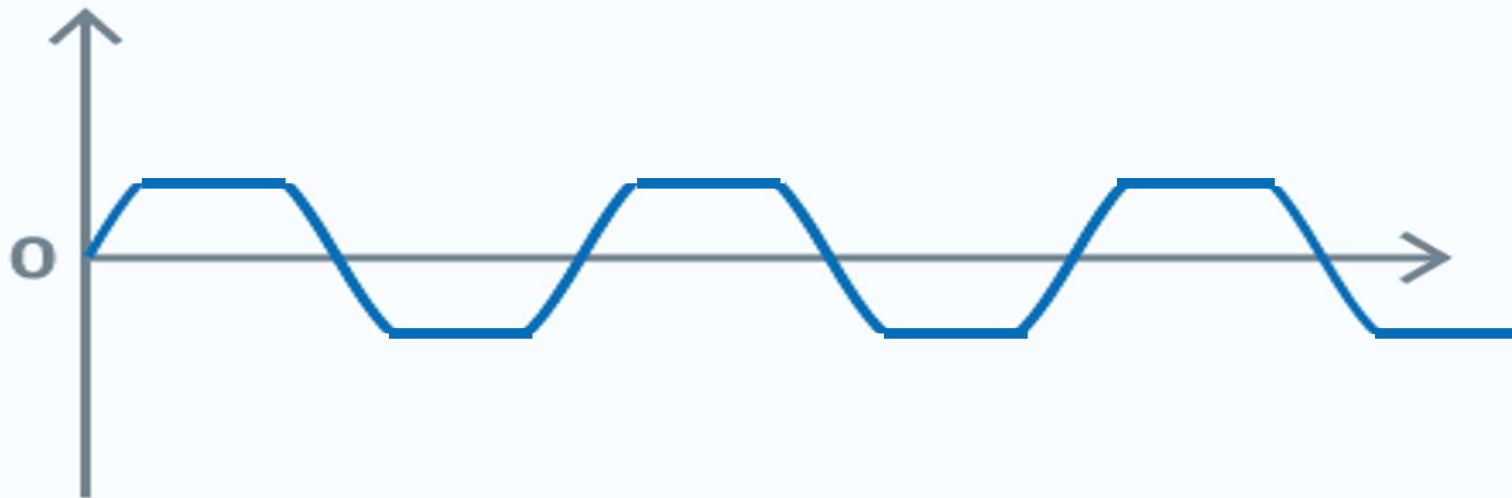
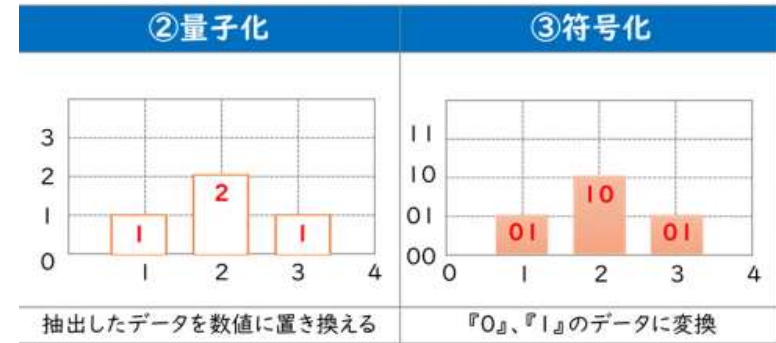
細かければ丸め誤差は小さくなるが、  
データサイズは大きくなるのでバランス注意





# 量子化 & 符号化

- 量子化範囲に注意  
アナログ値を分割して数値を割り当て  
→ 上限値, 下限値が存在



正弦波

# 具体的な計測機器のAD変換

PSG-1100 日本光電製  
日本で一番有名な睡眠時生体信号計測機器



アンプ＋  
AD変換の部分

増幅部

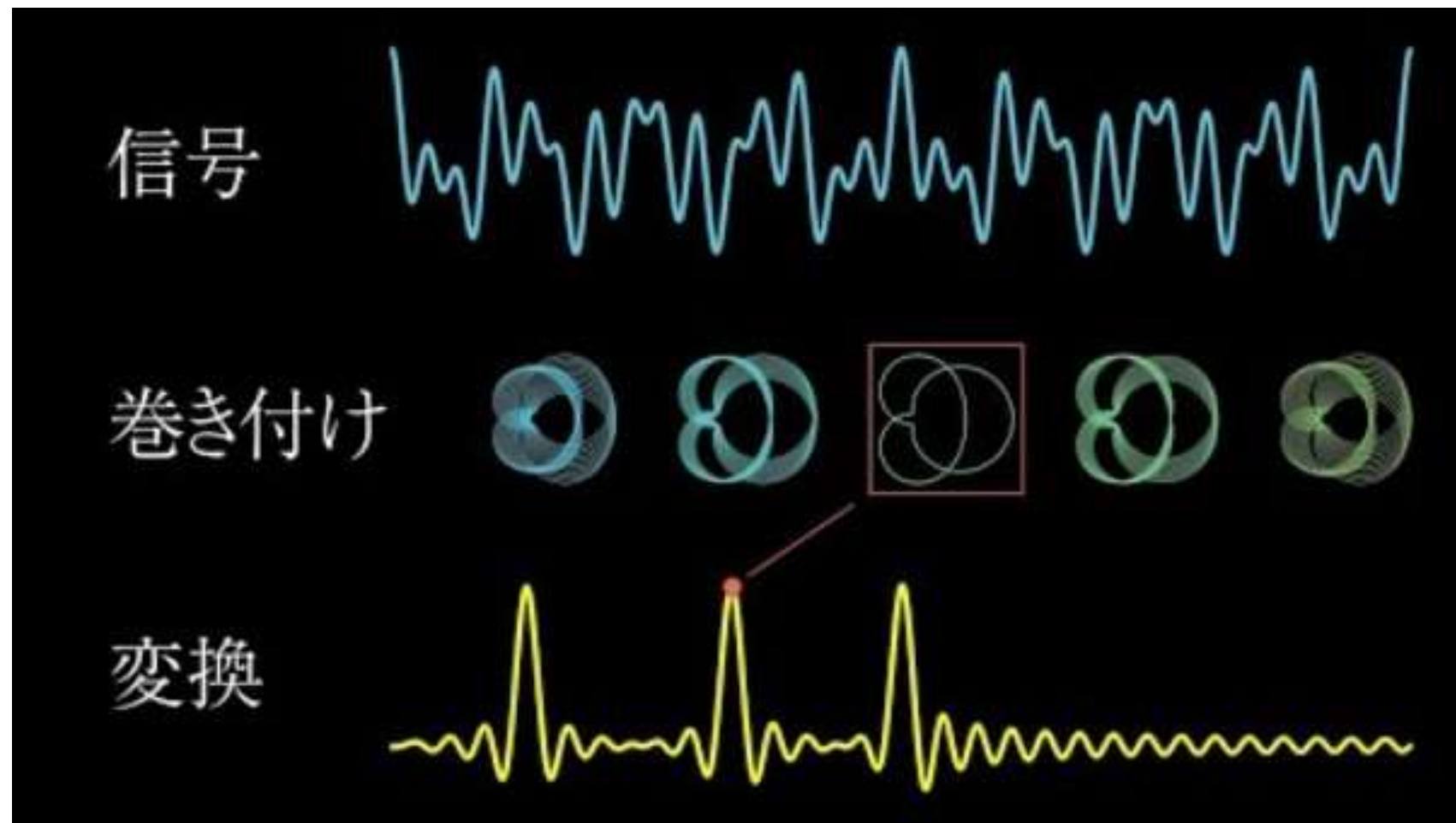
入力端子数	脳波(25)、多目的(8)、バイポーラ(8ペア)、体位(1)、呼吸圧(1)、SpO <sub>2</sub> (1)、ETCO <sub>2</sub> (1)
入力抵抗	100MΩ以上
雑音	1.5μVp-p (0.53～60Hz) 脳波・EOG・EMG・バイポーラ 3μVp-p (0.53～60Hz) ECG
弁別比	105dB以上 (60Hz)
A/D分解能	16bit
サンプリング	2000、1000、500、250、200Hz

# 信号計測における注意点まとめ

- サンプルング周波数に注意
  - 必要な周波数成分が計測可能なサンプルング周波数か？
  - エイリアシングが発生する状況か？
- 量子化に注意
  - 分解能は十分に小さいか？
  - 量子化のレンジは適切か？

# 信号の周波数変換

# フーリエ変換の基本的な考え方



# 短時間フーリエ変換と ウェーブレット変換

フーリエ変換：対象となる時間の間，波が定常的なこと前提  
→世の中そんな信号はほとんどない．．．

「短い時間」であれば定常的な波が仮定できる  
→短時間フーリエ変換

長周期の波と短周期の波を同じ窓幅で考えるのは変では？  
→ウェーブレット変換  
：マザーウェーブレット（特定周波数の波）との  
畳み込み計算

# 短時間フーリエ変換と ウェーブレット変換

フーリエ変換：対象となる時間の間 波が定常的なこと前提

→世  
「短  
→使う上では  
「時間軸ベースの信号を周波数成分に分解する方法」  
とっておけばOK.

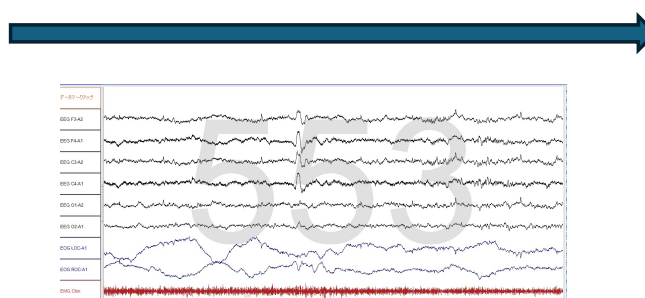
長周  
→対象となる周波数範囲や計算時間の許容を踏まえて  
ベストのものを選択しよう.

：マザーウェーブレット（特定周波数の波）との  
畳み込み計算



# 生体信号の前処理

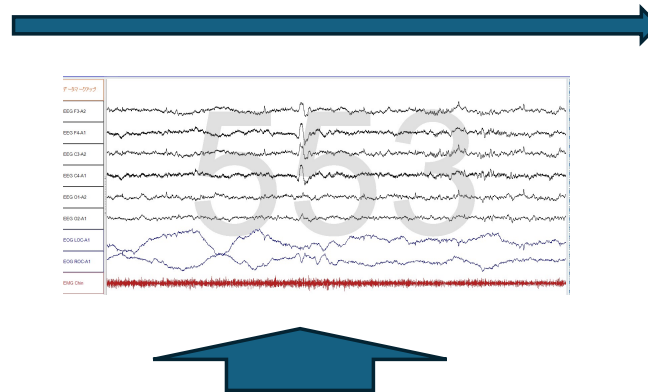
# 生体信号の認識, 大まかな流れ



Stage N2!

入力信号とステージラベルの  
間にある関連性を学習  
→前者から後者を推定

# 生体信号の認識, 大まかな流れ



Stage N2!

入力信号とステージラベルの  
間にある関連性を学習  
→前者から後者を推定

計測信号をモデルにとって  
使いやすい形に修正してあげることで  
判定精度の改善が可能

= 前処理

# 前処理の定義と適用先

数理モデルへの入力の前に  
モデルにとって都合の良い形状・特性に修正すること

- ノイズの除去
- リサンプリング
- 特徴の抽出

# ノイズの除去

数理モデルにとって，ノイズは精度低下の要因

特に，生体信号の場合

- 元々の電位変化が小さい ( $\mu V$ ,  $mV$ )
- 計測状況（センサの表面状態）が変化しやすい  
例) 体温変化, 汗の混入
  - 外乱によるノイズが混入しやすい
- 計測に手間がかかる
  - サンプル追加によるノイズの平滑化が困難.

→ ノイズを精査, その特徴を利用して除去

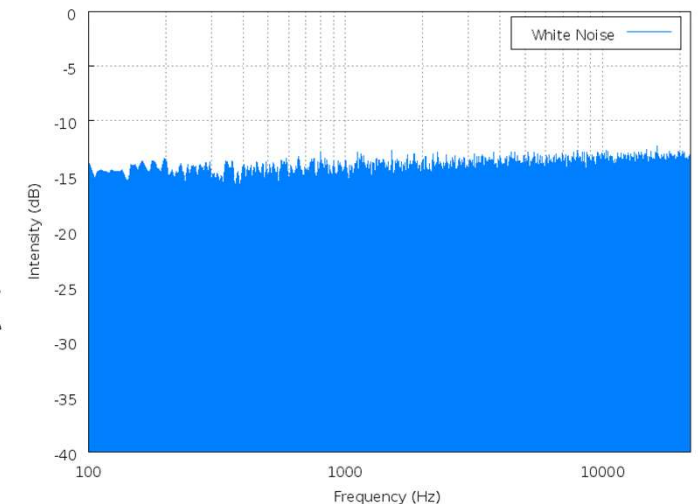
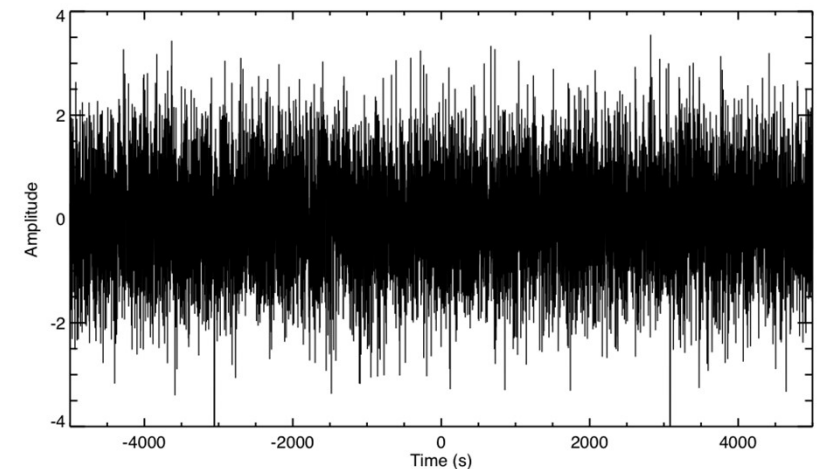
# 生体信号に入りやすいノイズの例

- 白色雑音

ランダム（に見える）な電位の変化  
均一な量の周波数成分が重畳して発生

汗，電極の接触不良，脱落時に  
観測されやすい

元信号がマスクされる，  
そもそも入っていないケース  
→ノイズ除去ではなく，信号排除で対応



# 生体信号に入りやすいノイズの例

- インパルスノイズ

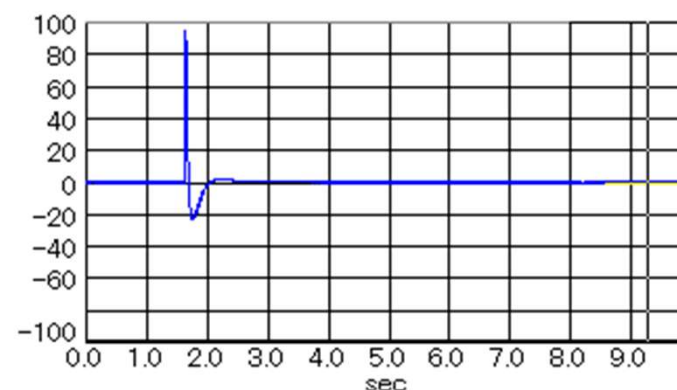
一瞬だけ大きな電位変化が発生

計測カイロに大きな電氣的な変化があった場合に発生.

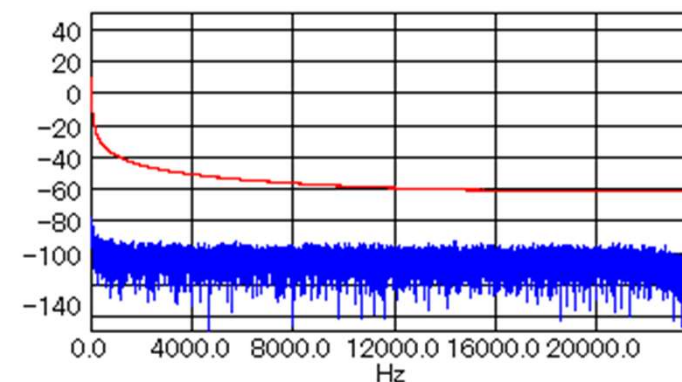
例) グラウンド電極が金属に接触  
信号線をアンプにつないだ瞬間  
部屋の蛍光灯をつけた瞬間

→周波数特性を犠牲にできるなら  
Medianフィルタリングで対応

赤線



線はホールドスペクトル 全体にレベルが上がる





# 生体信号に入りやすいノイズ

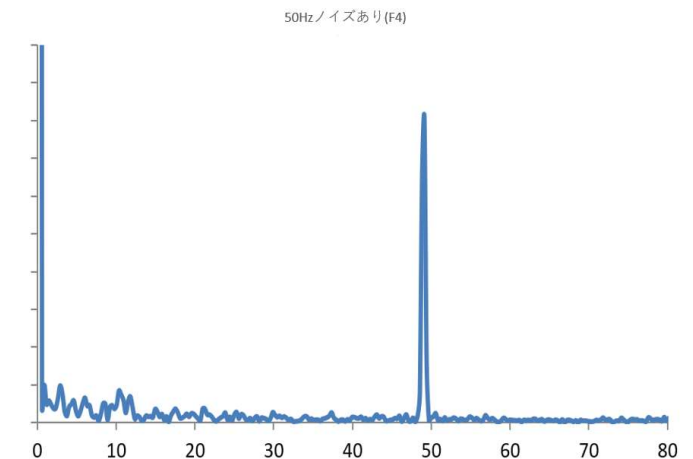
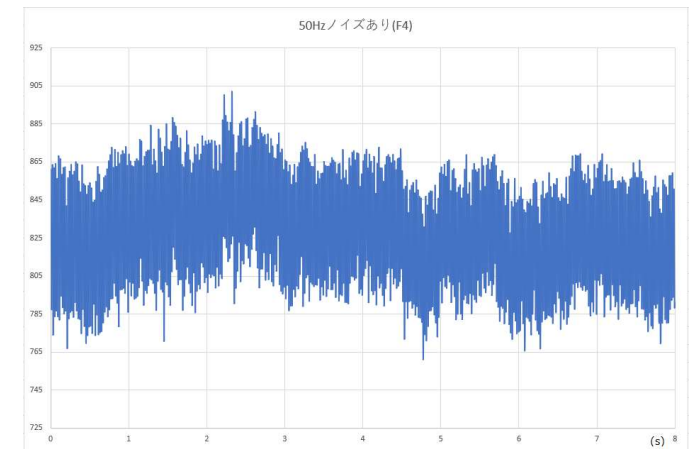
- 電源ノイズ

100V電源の電磁波を信号線などが拾ってしまった場合に発生

50/60Hzにピーク，ノイズ幅が一定  
→白色雑音と区別

直流蛍光灯を使っていると25Hzにも

電源周波数のノッチフィルタ  
信号線のシールド化



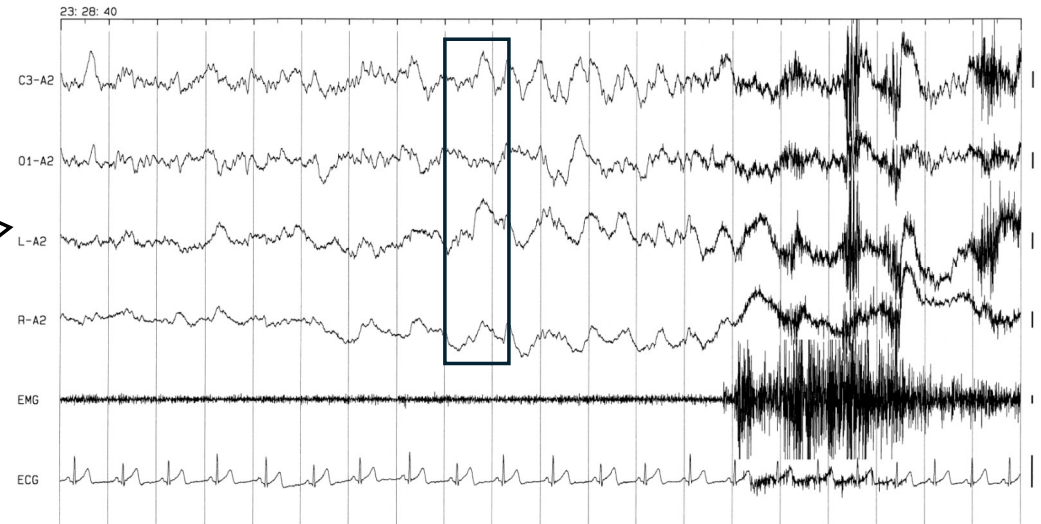
# 生体信号に入りやすいノイズ

- 別信号の混入

眼電位が同位相 = ノイズ  
ノイズ源はおそらく前頭部の脳波

センサ電極の距離が近いと  
見たい生体電位以外にも  
重畳して観測される

- 混入する信号によって対応が異なる
- 複数センサの同時計測して打ち消す
  - 周波数成分の違いを利用する（筋電位混入）



# 生体信号に入りやすいノイズ

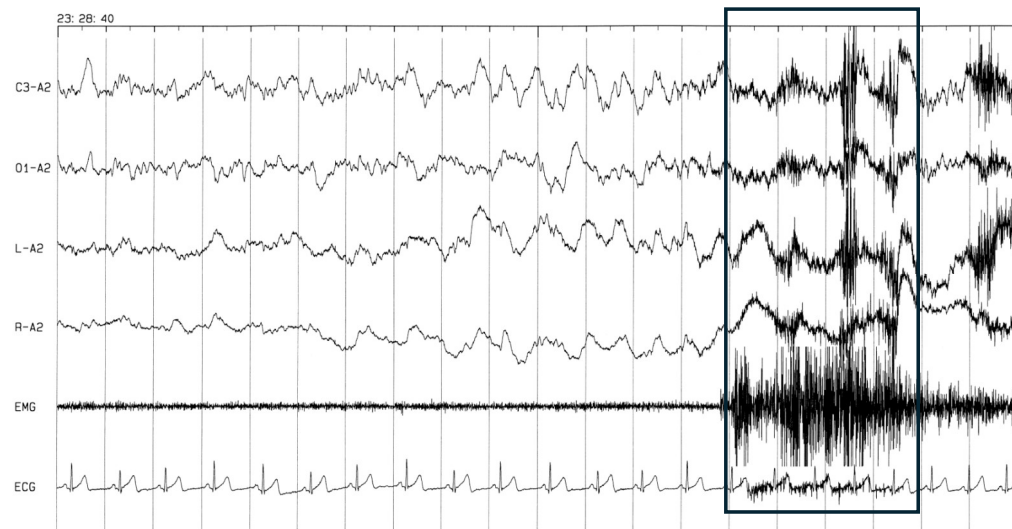
- 体動

体が動くことで

- 筋電位の混入
- 信号線の揺れに伴う電位変化
- 電極の設置状態の変化
  - 取り付け圧力
  - センサー皮膚間距離
  - センサが浮くとコンデンサ化&電氣的性質が変化

が発生，それが電位的な変化として観測される

睡眠ステージ判定では「体動はステージW情報」なので除去しない。



# ノイズと信号の捉え方

基本的に、

やりたい解析を行う上で有用な波＝信号

//

邪魔な波＝ノイズ

体動は普通「ノイズ」だが、  
ステージ判定では「信号」の一つ

K複合の形状の違いは、  
ステージ判定では「ノイズ」だが、  
ADHDの早期発見がしたいなら「信号」

# 逆に言うと、

解析に本当に必要な波形だけ計測するのが理想

- ノイズ源の除去
  - 蛍光灯を使わない，インバータ使わない
  - 同じ部屋で大電力を使わない（サーバーとかおかない）
- センサの脱落を防ぐ
  - シールの粘着性をあげる，導電性ジェルを増やす
  - 取り付け前にアルコールで皮膚表面を拭く
- 周波数フィルタの適用
  - 必要な周波数成分だけ計測する，抽出する  
→バンドパスフィルタの利用
- ノイズ入りの信号を解析から除外する

# リサンプリング

標準化，量子化後の信号のサンプリング周波数を変更すること

- 本質的には意味のある操作ではない  
(情報が増えるわけでも，不要な情報が除かれるわけでもない)
- が，現実的にはよく行う

例えば，

- 公開されている睡眠記録A,B,Cがそれぞれ違うサンプリング周波数を採用しているケース
- 公開されている睡眠ステージ判定モデルが，特定のサンプリング周波数前提でチューニングされている

# アップとダウン

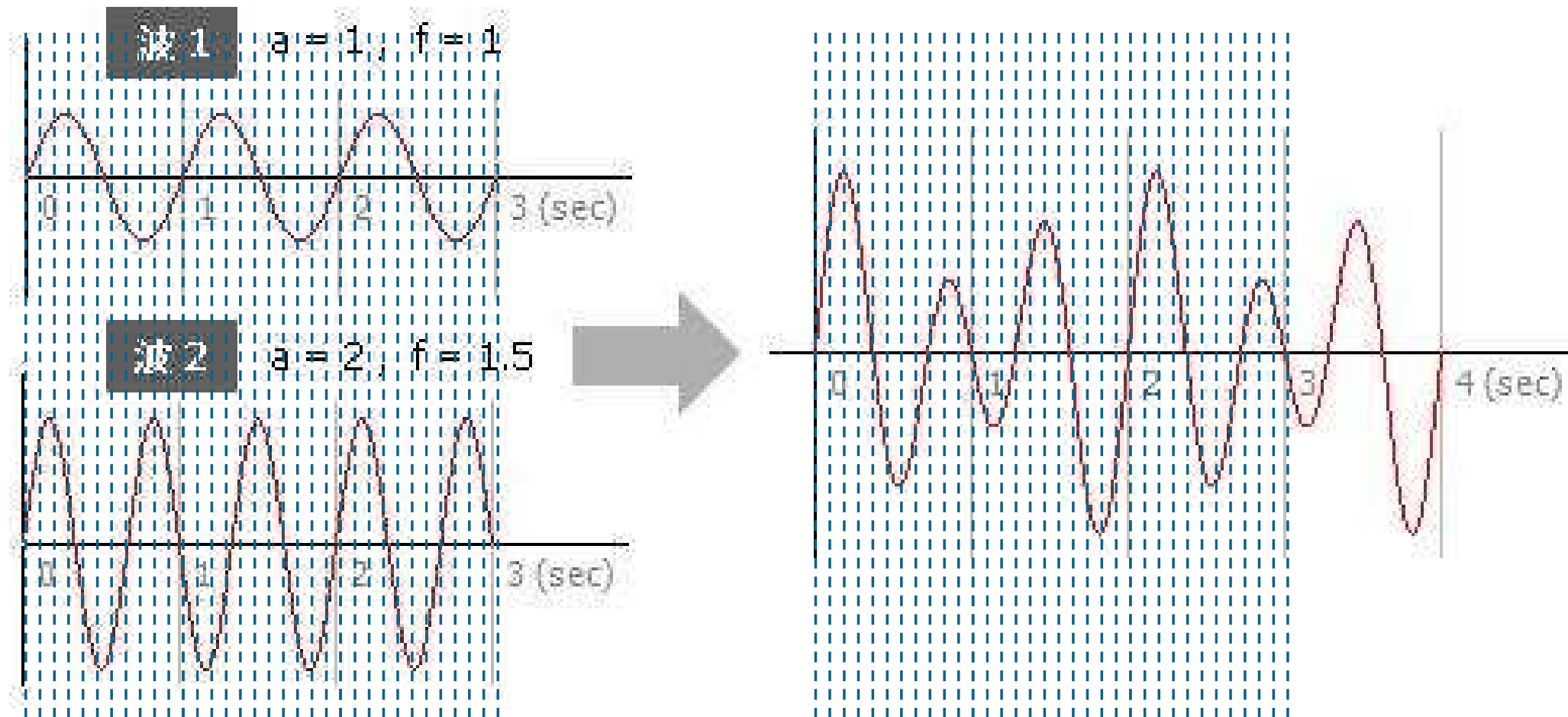
アップサンプリング：より高いサンプリング周波数に

ダウンサンプリング：より低いサンプリング周波数に

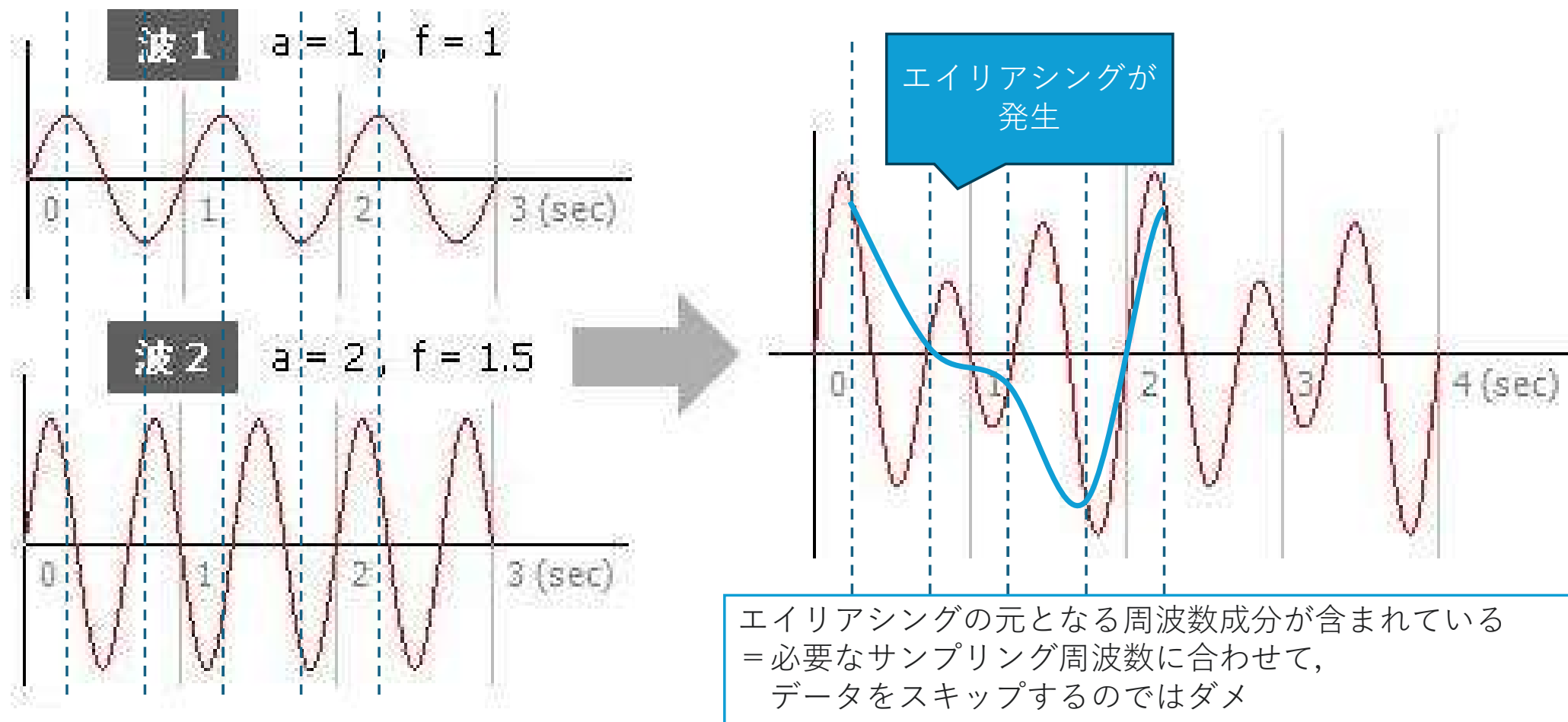
アップ・ダウン共に、  
へたなやり方でリサンプリングするとエイリアシングが発生



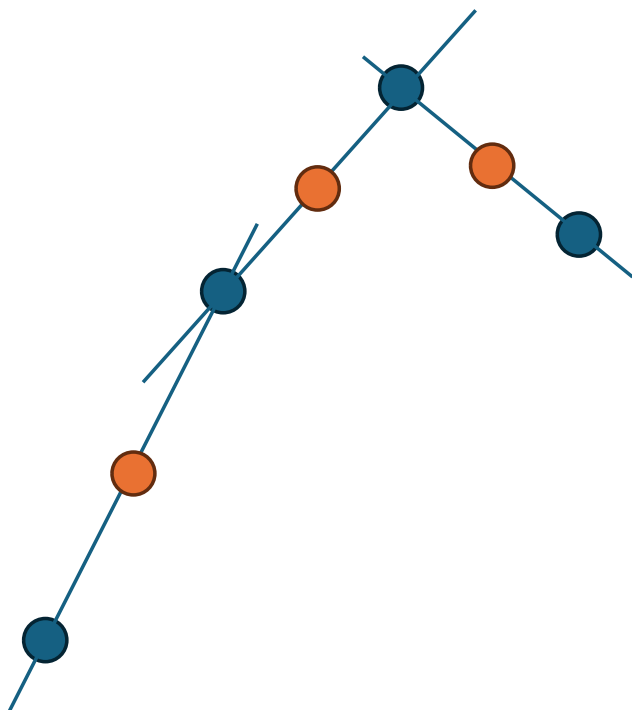
# ダメなダウンサンプリングの例



# ダメなダウンサンプリングの例



# ダメなアップサンプリングの例

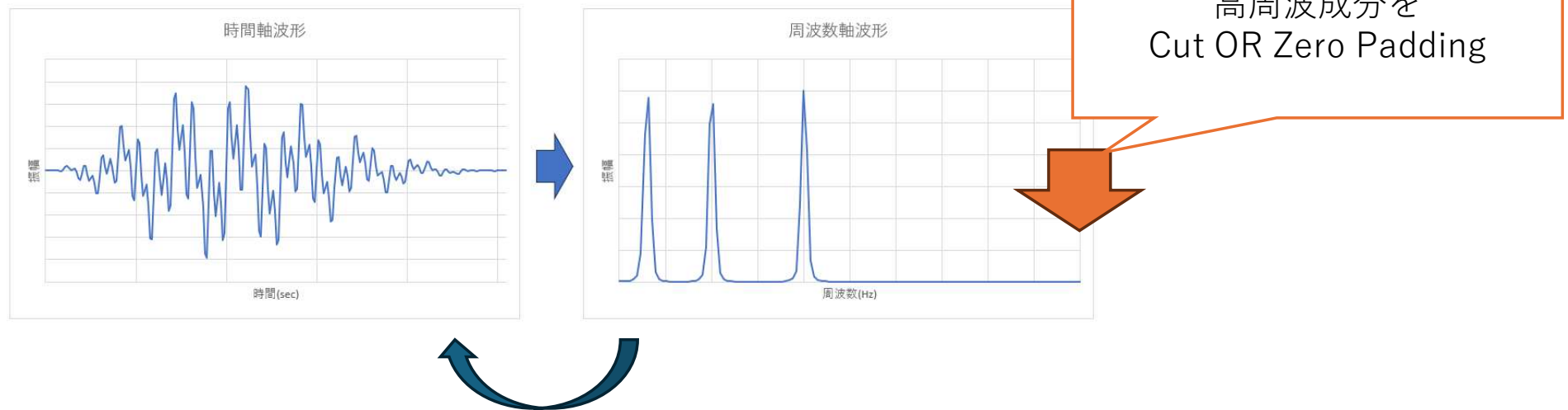


新しいサンプル点のY値を  
両側の計測点の平均値にする

アップサンプルは  
元々の信号形状を損ねずに  
できるはずなのに、  
周波数成分が変化してしまう

# 適切なリサンプリングの方法

フーリエ変換を経由する



これならスペクトラムを変えずにアップ/ダウンサンプリング可  
ただし計算量は多くて大変. . .

# 特徴抽出

判定に有用な情報・信号に変換する作業

- 非機械学習モデルや一般的な機械学習モデルでは特徴抽出の是非が全体の精度に大きく影響
- 深層学習モデルであっても、抽出が困難な特徴を事前に抽出しておくことで、学習サンプル，学習回数の節約が期待できる。

アライグマ



しっぽが縞模様  
目の上、鼻の周りが白い

タヌキ



しっぽが短い  
顔の毛が横に広がる

# 生体信号の特徴抽出例

- フーリエ・ウェーブレット  
→ 周波数特徴が重要な場合に,  
  # 1 次畳み込みニューラルネットで獲得（再現）可能  
  既に説明したので省略
- 全波整流平滑化
- 尖度・歪度計算

# 全波整流平滑化

信号の振幅を抽出する方法

$$s_{new} = f(|s_{old}|)$$

$f(x)$  : 平滑化 (ローパスフィルタ)

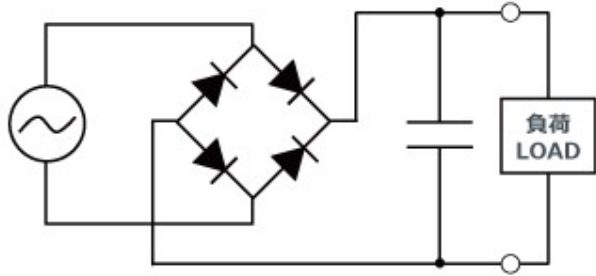
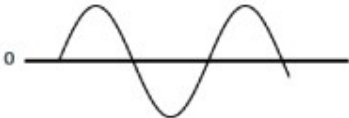

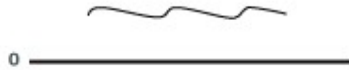
元々の信号に絶対値

→一定の窓幅で平均値をとる

筋電位から筋張力を見たいとき

一応、畳み込みと

ReLUの組み合わせで学習可能

	全波整流
回路構成	
入力 電圧波形	
整流後 電圧波形	
整流平滑後 電圧波形	

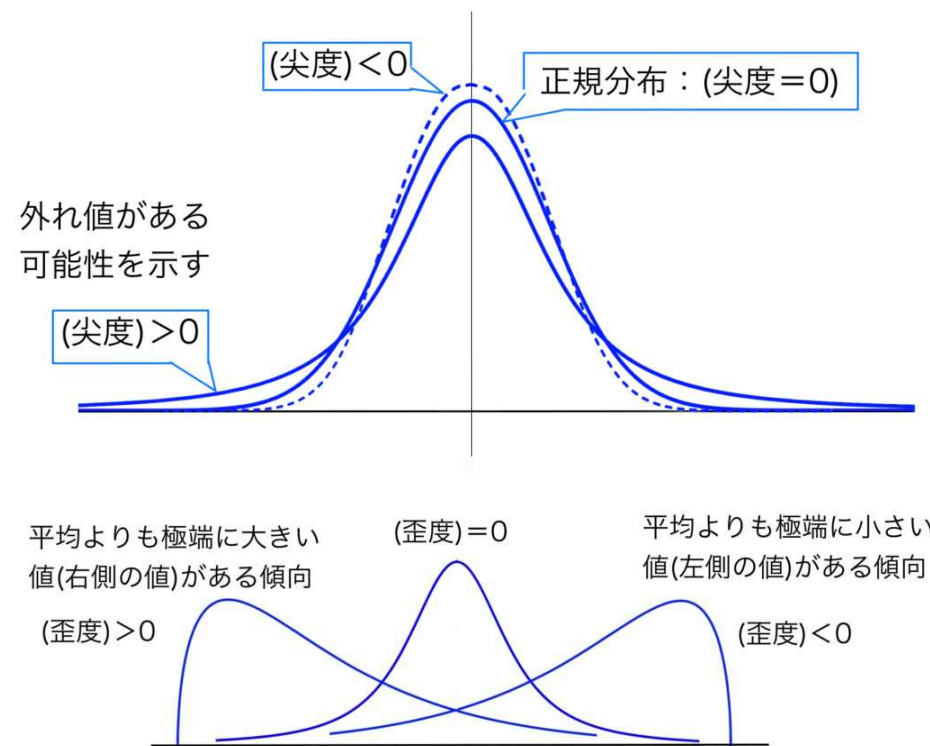
# 尖度・歪度

尖度：  
信号値の分布がどのくらい  
尖っているか

$$\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^n \frac{(x_i - \bar{x})^4}{s^4} - \frac{3(n-1)^2}{(n-2)(n-3)}$$

歪度：  
信号値の分布がどのくらい  
歪んでいるか

$$\frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{s} \right)^3$$



累乗計算があるので  
シンプルには獲得できない  
ノイズ混入量の見極め等に