# 東京大学 工学部 計数工学科/物理工学科

# 応用音響学：

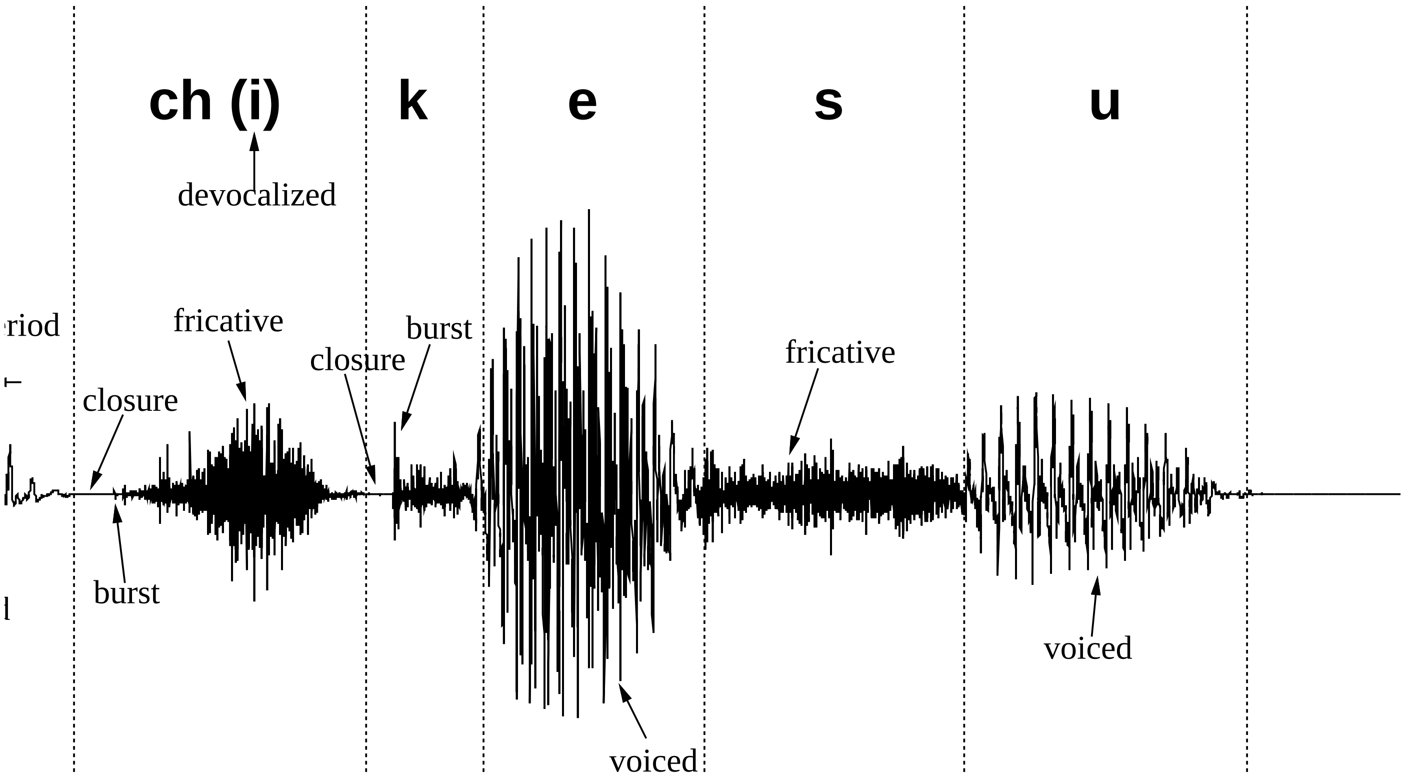
# 音声分析 (1) サンプリングと量子化

嵯峨山 茂樹 <sagayama@hil.t.u-tokyo.ac.jp>
東京大学 工学部 計数工学科 http://hil.t.u-tokyo.ac.jp/

- 音声のディジタル化
- サンプリング定理
- 量子化

# 音声波形

**「打ち消す」の音声波形を示す。「ち」は無声化して、母音が脱落している。有声音(ここでは u, e, u)には、周期性が見られる。これが、ピッチ(声の高さ)として感じられる。**




**Speech Waveform /uchikesu/**

**図1. 音声波形の例「打ち消す」**

# 音声のディジタル化

- 音・電気変換(マイクロフォン) (transducer)
- 増幅 (amplification)
- A/D 変換 (A-to-D conversion)
  - フィルタリング (filtering)
  - サンプリング (sampling) — 時間離散化
    - サンプリング定理：$2B$
  - 量子化 (quantization) — 値離散化
    - 量子化雑音：$\frac{q^2}{12}$
- ディジタル値の取り込み

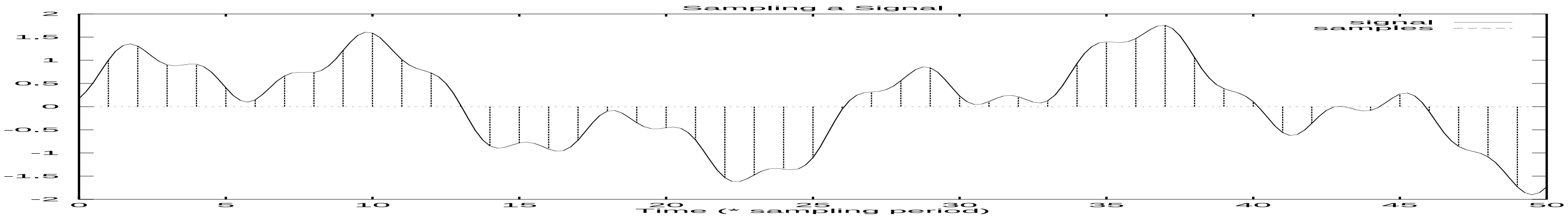
# 信号のディジタル処理の手順

## サンプリング周波数$\omega_0$による信号のサンプリングと復元の実際


1. 低域通過フィルタ (low-pass filter: LPF) により、周波数$\omega_0/2$以上をカット。
2. サンプリング周波数$\omega_0$により信号をサンプリング。
3. 目的のディジタル信号処理、ディジタル記録、ディジタル伝送などを行なう。
4. サンプル値を並べて、インパルス列信号を生成。
5. 低域通過フィルタ (low-pass filter: LPF) により、周波数$\omega_0/2$以上をカット。

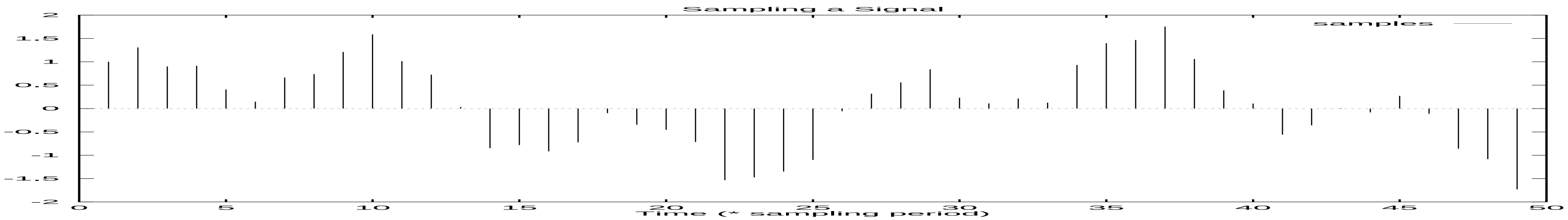
# 信号のサンプリング(標本化)と復元

- **目的: ディジタル信号処理、離散系への変換、etc.**
- **信号 $f(t)$ から等しい間隔の標本(サンプル)列を取り出す。**



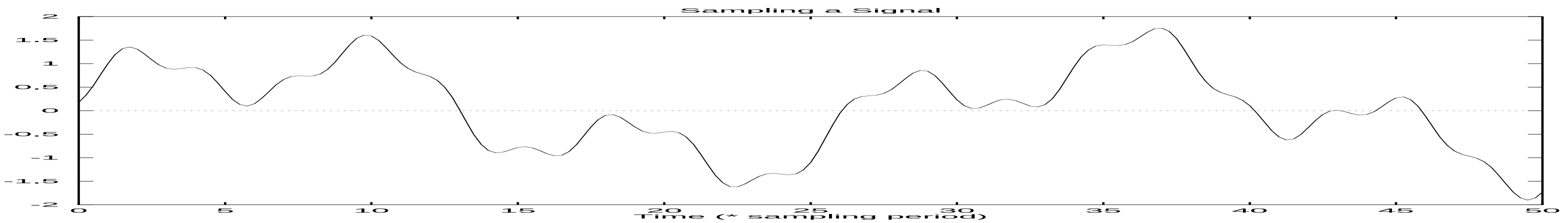

- **サンプリング(標本化)**

↓



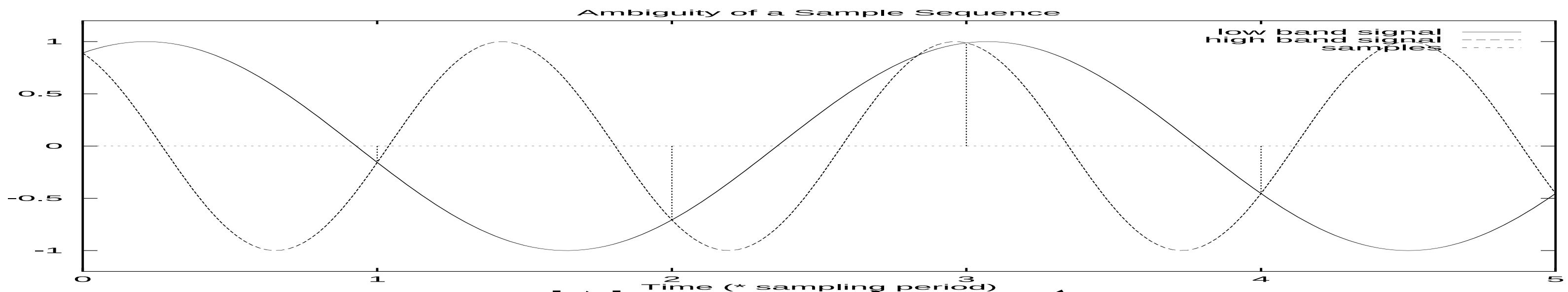

- **信号復元**

↓

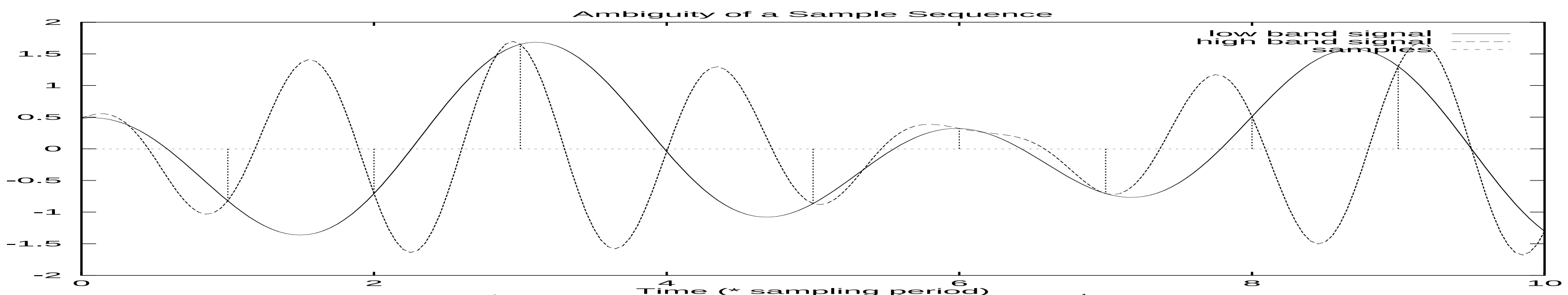


# 信号のサンプリング(標本化)と復元の問題

- 正確に復元できる(原情報の情報を保存している)サンプリングの条件は何か?
- サンプル値列の多義性




正弦波のサンプリング




一般信号のサンプリング

- 直感的サンプリング定理:
  - 「正弦波は1周期あたり、最低サンプル点2個必要。」

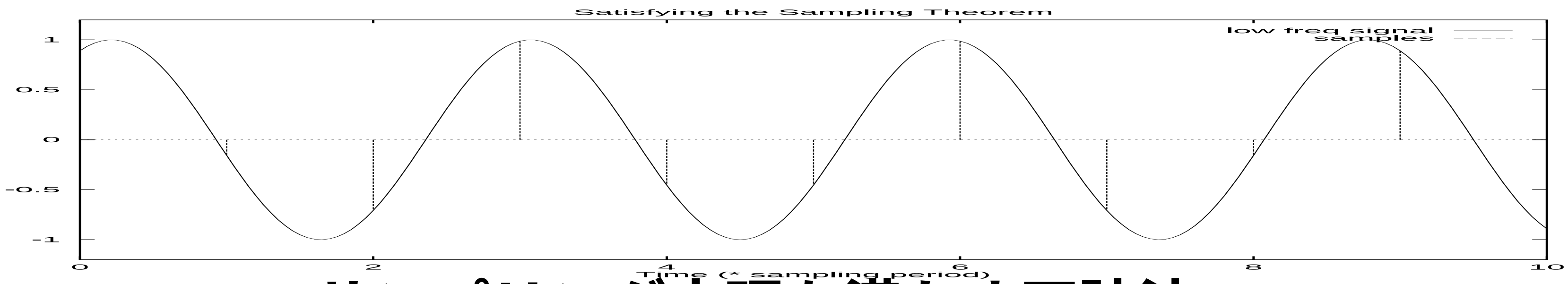
# サンプリング(標本化)定理

## ■ 標本化定理

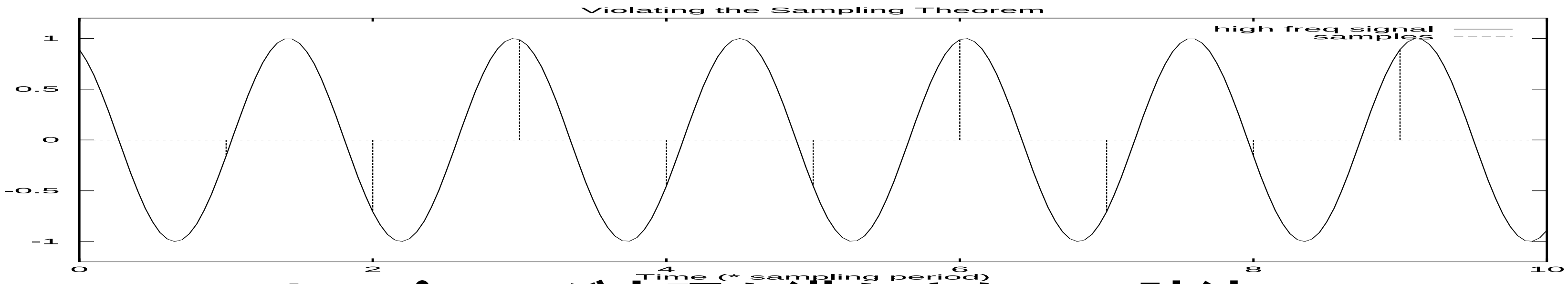
$f_m$[Hz]以上の周波数成分を持たない帯域制限信号 $f(t)$ は、$1/2f_m$ [秒]より短い等しい間隔の標本(サンプル)で完全に決定される。

$$\mathcal{F}[\ f(t)\ ] = F(\omega) = \begin{cases} 0, & |\omega| > \omega_m = 2\pi f_m \\ \textsf{any}, & |\omega| < \omega_m = 2\pi f_m \end{cases}$$

情報を失わずに連続データと離散データを相互に変換可能。



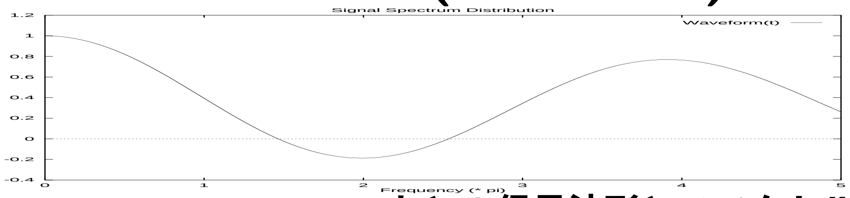

サンプリング定理を満たす正弦波

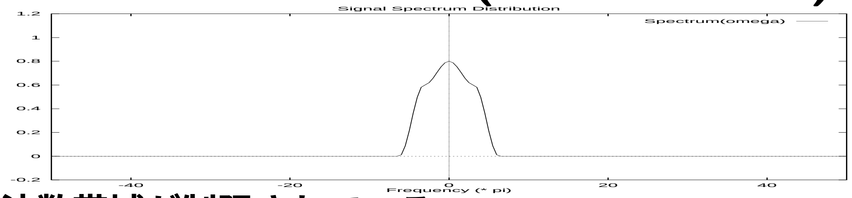



サンプリング定理を満たさない正弦波

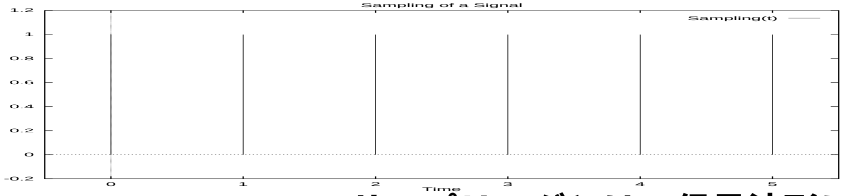
# サンプリング定理 (図解)

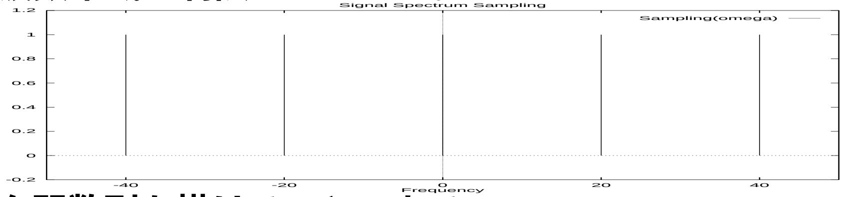
## 信号波形 (時間領域)　　スペクトル (周波数領域)



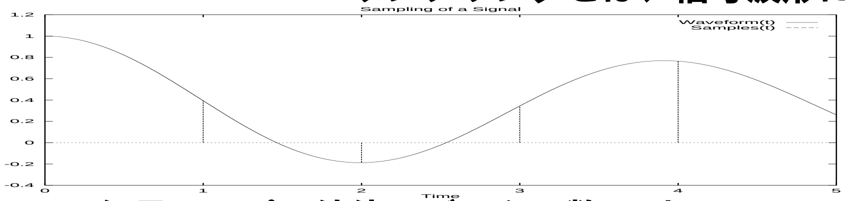



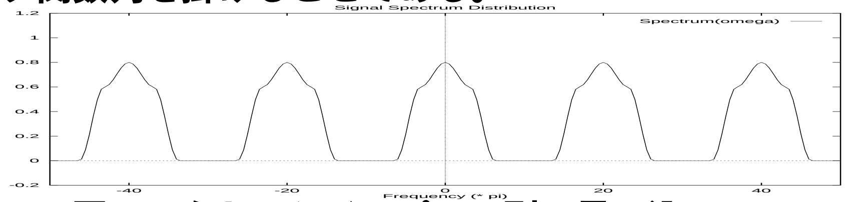

**もとの信号波形とスペクトル。周波数帯域が制限されている。**



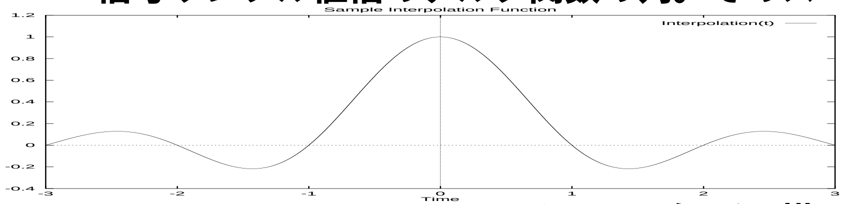



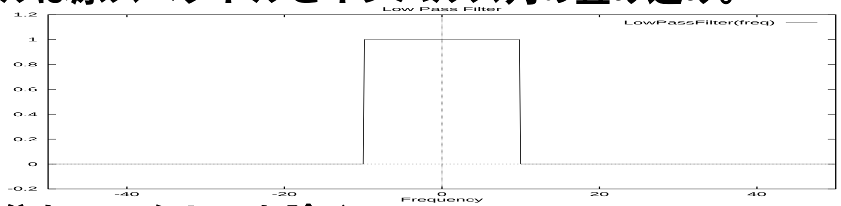

**サンプリングとは、信号波形にデルタ関数列を掛けることである。**



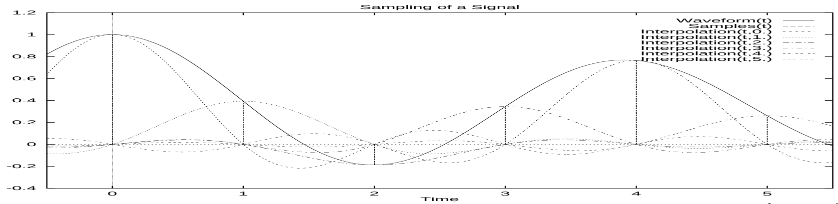

**信号サンプル値倍のデルタ関数の列。そのスペクトルは原スペクトルとインパルス列の畳み込み。**

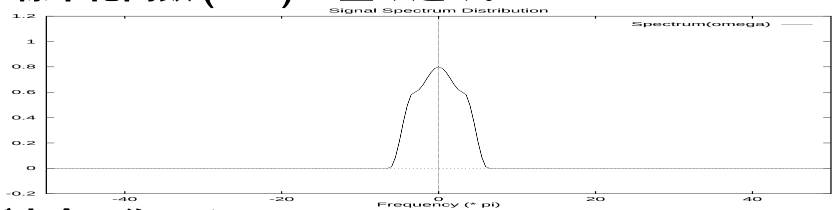



**スペクトルに矩形を掛けて余分なスペクトルを除く。**

**これは信号サンプル値インパルス列と標本化関数 (sinc) の畳み込み。**



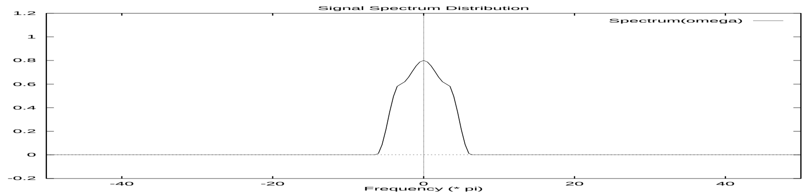

**こうしてもとの信号波形が完全に復元される。**

# サンプリング定理の導出

- **信号 $f(t)$ の帯域を $f_m$[Hz] 以下に制限**

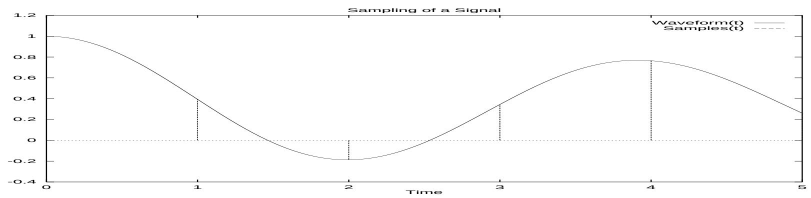
$$\mathcal{F}[\ f(t)\ ] = 0$$

$$\text{where} \qquad |\omega| > \omega_m = 2\pi f_m$$




- **$f(t)$ を間隔 $T$ でサンプリングした標本列信号 $f_s(t)$ は、**

$$f_s(t) = f(t)\delta_T(t)$$



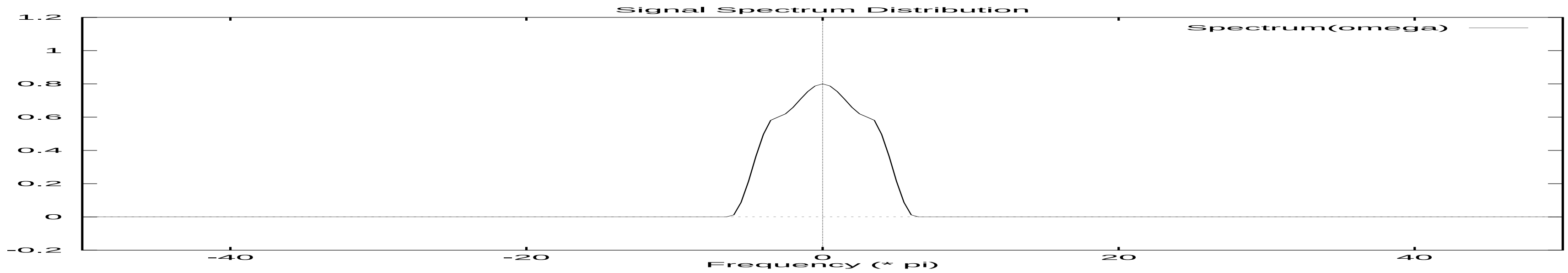

- $\mathcal{F}[\ f(t)\ ] = F(\omega),\ \mathcal{F}[\ f_s(t)\ ] = F_s(\omega),\ \omega_0 = 2\pi/T$ **とすると、**

$$\begin{aligned} F_s(\omega) &= \mathcal{F}[\ f(t)\delta_T(t)\ ] = \frac{1}{2\pi}F(\omega) * \mathcal{F}[\ \delta_T(t)\ ] \\ &= \frac{1}{2\pi}F(\omega) * \{\omega_0\delta_{\omega_0}(\omega)\} = \frac{1}{T}\int_{-\infty}^{\infty} F(\omega)\delta_{\omega_0}(\omega-\sigma)d\sigma \end{aligned}$$

$F_s(\omega)$ **は周波数領域で** $F(\omega)$ **を周期** $\omega_0$ **で並べたものとなる。**

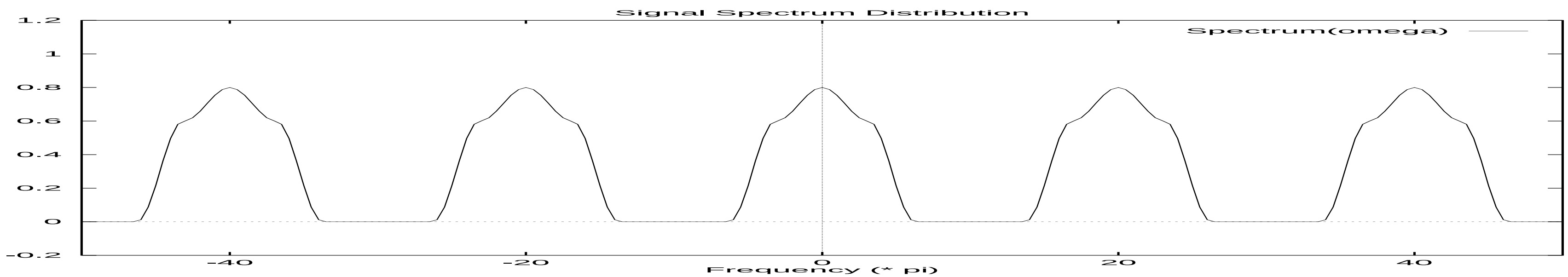
# エリアシング (aliasing)

- 信号の帯域 (0でないスペクトル成分の存在範囲): $B$、




原信号スペクトル。帯域は $B$

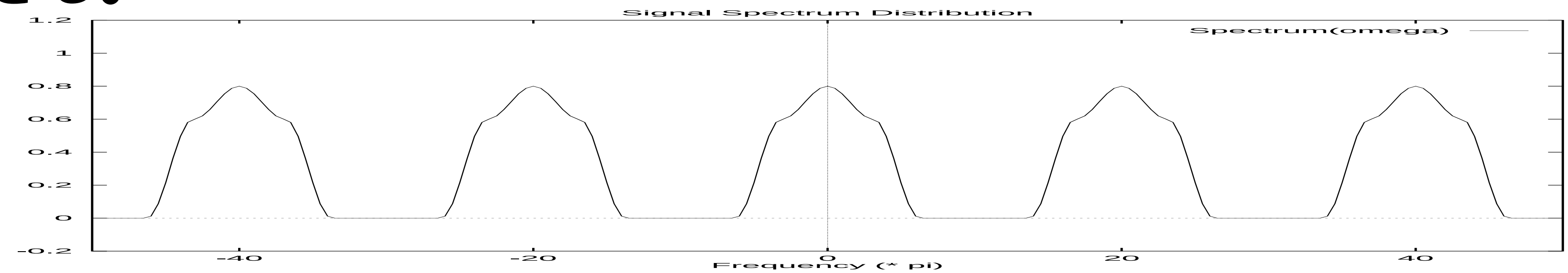
- サンプリング周波数 $f_m$ でサンプリングしたサンプル値のインパルス列信号のスペクトルは、$f_m$ 間隔で繰り返しパターン。→ エリアシング (aliasing)




サンプリングによって生じるエリアシング

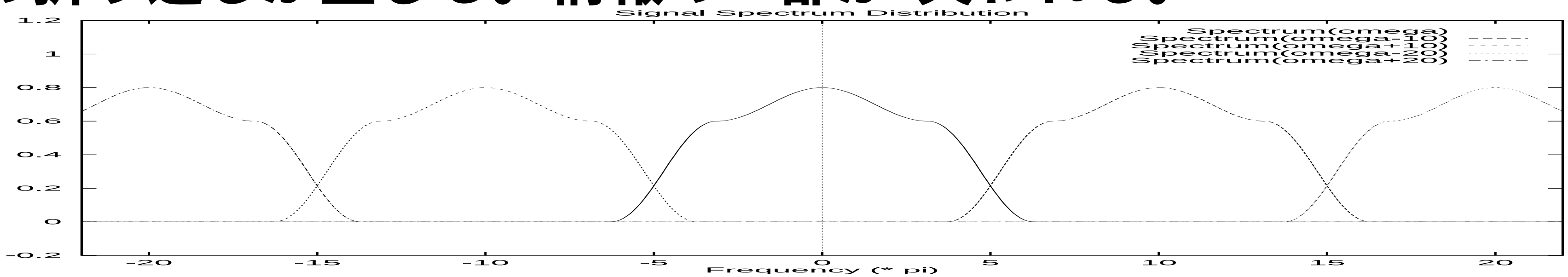
# 折り返し

- **Nyquist (ナイキスト) 周波数:** $f_m = 2B$
  **(Nyquist 間隔:** $T = \frac{1}{2f_m}$**、 最大サンプリング周期)**
- **Nyquist 周波数以上の周波数でサンプリングすれば情報が保存できる。**




**折り返しのないサンプル列スペクトル**

- **Nyquist 周波数未満の周波数でサンプリングすると、スペクトルの折り返しが生じる。情報の一部が 失われる。**




**折り返しの生じたサンプル列スペクトル**

# サンプリング定理による波形の復元

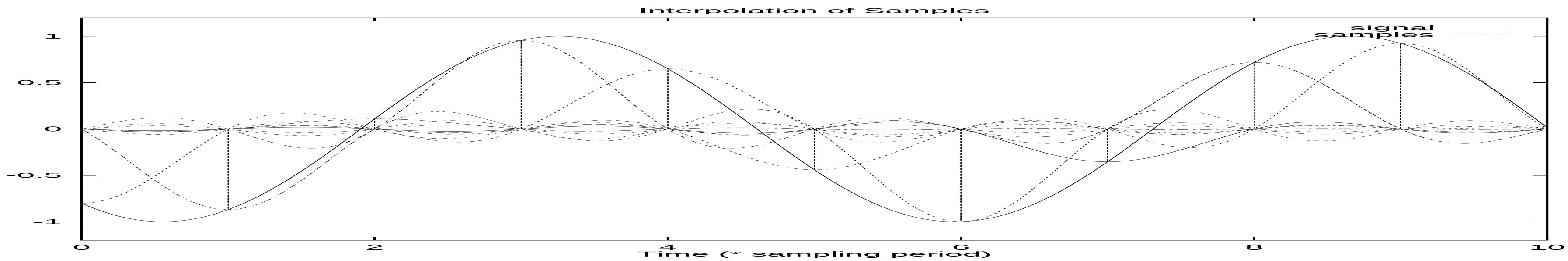
- **原信号** $f(t)$ **(**$f_m = \omega_m/2\pi$ **以下に帯域制限) をサンプル値系列** $f_s(t)$ **(サンプル間隔** $T = 1/2f_m$**) から復元するには?**
- **周波数領域で、(rect**$_W$ **は幅** $W$ **の矩形関数)**

$$
\begin{aligned}
f(t) &= \mathcal{F}^{-1}[TF_s(\omega)\ \mathrm{rect}_{2\omega_m}(\omega)] = Tf_s(t) * \left\{\frac{\omega_m}{\pi}\mathrm{sinc}(\omega_m t)\right\} \\
&= f_s(t) * \mathrm{sinc}(\omega_m t)
\end{aligned}
$$

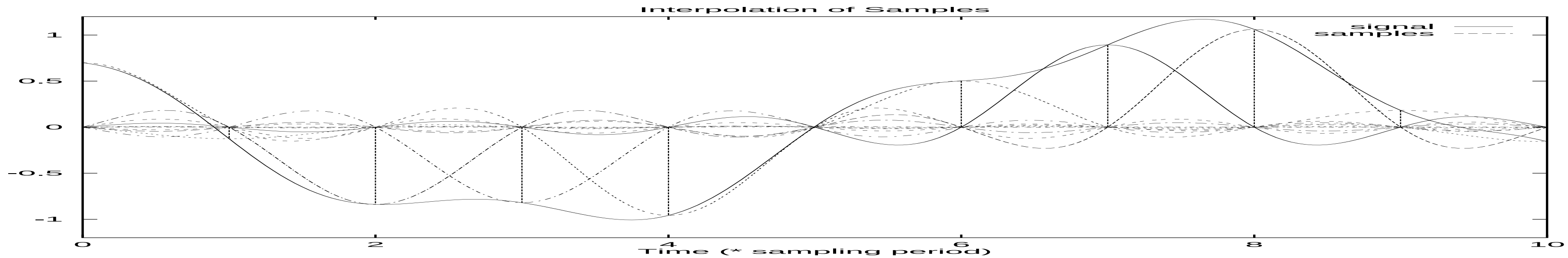
$k$ **番目の標本値を** $f_k$ **とすると** $f_s(t) = \sum_{-\infty}^{\infty} f_k \delta(t - kT)$

$$
\begin{aligned}
f(t) &= f_s(t) * \mathrm{sinc}(\omega_m t) = \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f_k \delta(s - kT)\mathrm{sinc}(\omega_m(t - s))ds \\
&= \sum_{k=-\infty}^{\infty} f_k \mathrm{sinc}(\omega_m(t - kT)) = \sum_{k=-\infty}^{\infty} f_k \frac{\sin \omega_m(t - kT)}{\omega_m(t - kT)}
\end{aligned}
$$

# サンプリング定理による波形の復元




**サンプル値列から原信号の復元(正弦波の例)**




**サンプル値列から原信号の復元(一般信号の例)**

# 周波数領域におけるサンプリング定理

- **サンプリング定理は時間領域と周波数領域で双対性を持つ。**
- **周波数領域のサンプリング定理:**

  **時刻** $|t| > T$ **[秒]で** $0$ **となる時間制限信号** $f(t)$ **は、** $1/2T$ **[Hz] (** $= \pi/T$ **[rad/s]) より小さな等間隔の周波数スペクトルのサンプルで完全に決定できる。**
- **(導出)** $|t| > T$ **で** $f(t) = 0$ **より、** $T = \pi/\omega_0$ **とすると、** $f(t) = \Sigma_{-\infty}^{\infty} c_k e^{jk\omega_0 t}$ **だから**

$$c_k = \frac{1}{2T}\int_{-T}^{T} f(t)e^{-jk\omega_0 t}dt = \frac{1}{2T}\int_{-\infty}^{\infty} f(t)e^{-jk\omega_0 t}dt = \frac{1}{2T}\mathcal{F}[\ f(t)\ ](k\omega_0)$$

$\mathcal{F}[\ f(t)\ ](k\omega_0) = F(k\omega_0)$ **は、周波数領域で** $\omega_0 = \pi/T$ **[rad/s] 間隔のサンプル列。**

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt = \int_{-T}^{T}\left\{\sum_{k=-\infty}^{\infty}\frac{1}{2T}F(k\omega_0)e^{jk\omega t}\right\}e^{-j\omega t}dt \\ &= \sum_{k=-\infty}^{\infty}\left\{F(k\omega_0)\frac{1}{2T}\int_{-T}^{T}e^{-jk(\omega-\omega_0)t}dt\right\} = \int_{-T}^{T}F(k\omega_0)\frac{\sin(\omega-k\omega_0)T}{(\omega-k\omega_0)T}dt \end{aligned}$$

# 信号のリサンプリング

- **目的/用途: ディジタル系の間の情報の適切な変換**
  **例: CD(44.1kHz) とDAT(48kHz) の変換、画像の画素数の変換、その他あらゆるデジタル系接続。**
- $x(t)$**: 原信号，** $x_k$**: サンプル列，**
  $\omega_m$**: サンプリング周波数 → サンプリング周波数**$\omega'_m$**への変換。**
- **任意の時刻**$t$**についての信号の補間計算**

$$x(t) = \sum_{k=-\infty}^{\infty} x_k \cdot \mathrm{sinc}\ \omega_m(t - kT) \qquad (T\text{はサンプリング周期})$$

- $\omega'_m > \omega_m$ **の場合 (アップサンプリング)**
  **新しいサンプリング点における信号の補間計算 (サンプリング周期**$T' = 2\pi/\omega'_m$**)**

$$x(iT') = \sum_{k=-\infty}^{\infty} x_k \cdot \mathrm{sinc}\ \omega_m(iT' - kT)$$

- $\omega'_m < \omega_m$ **の場合 (ダウンサンプリング)**
  **新しいサンプリング点における信号の補間計算 (サンプリング周期**$T' = 2\pi/\omega'_m$**)**

$$x(iT') = \sum_{k=-\infty}^{\infty} x_k \cdot \mathrm{sinc}\ \omega'_m(iT' - kT)$$

# 連続信号伝送とPCM伝送

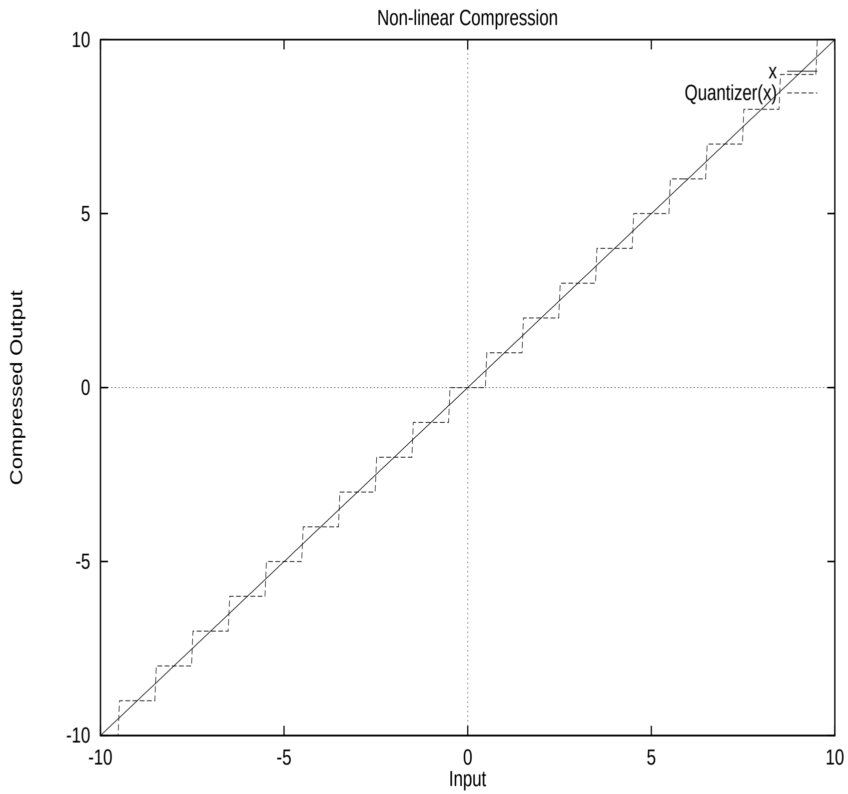
- 連続値信号(帯域 $f_m$[Hz])の伝送 → 毎秒 $2f_m$ 個のサンプルを伝送。(サンプリング定理)
- 連続値信号の1サンプルの情報量は $\infty$ (情報理論)
- 雑音のある通信路では、通信容量は有限 (Hartley-Shannon)
→ 連続値信号サンプルを送っても、受信側で完全に復元できない。
- サンプル値を離散化(量子化)
→ サンプルあたりの情報量を有限に制限
- PCM (Pulse-Coded Modulation, パルス符号化変調)
→ 量子化されたサンプル値を伝送・蓄積。

$$\text{信号の毎秒の情報量} < \text{伝送路の通信容量}$$

ならば、情報が完全に復元できる。

# 量子化

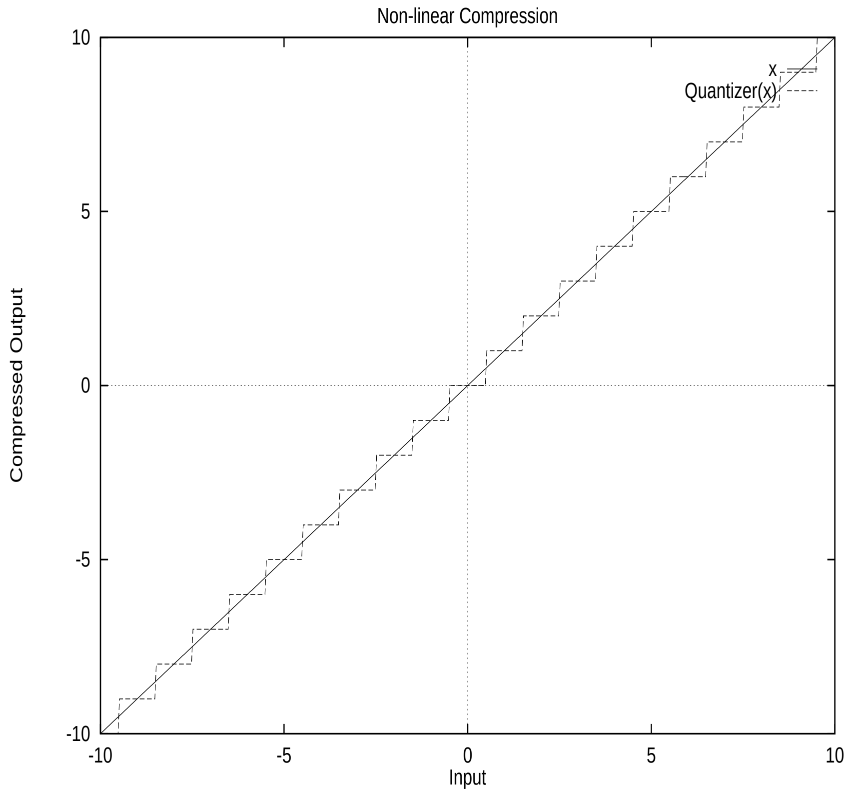
- 標本値の離散化： AD 変換 (analog-to-digital conversion)、量子化
- 標本値を $\Delta$ 間隔で離散化し、量子化区間番号 $n$[bit] の 2 進数で表現
- 線形量子化 (均一量子化)
    - 量子化器
      量子化幅を $\Delta$ とすると、
      区間 $-2^{n-1}\Delta \sim (2^{n-1}-1)\Delta$ にある連続値を量子化区間番号に変換
    - 量子化誤差
      → 入力値に依存した誤差値
      (右図の 45 度直線と階段形状との差)



# 量子化雑音

## 線形量子化 (均一量子化)

- 量子化誤差
  → 波形に依存した雑音の重畳と見なせる
  →「量子化雑音」




## 量子化雑音のパワー：入力を一様分布と仮定。

平均: $$\overline{e(t)} = \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} u du = 0$$

分散(パワー): $$\overline{e^2(t)} = \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} u^2 du = \frac{\Delta^2}{12}$$

# 不均一量子化

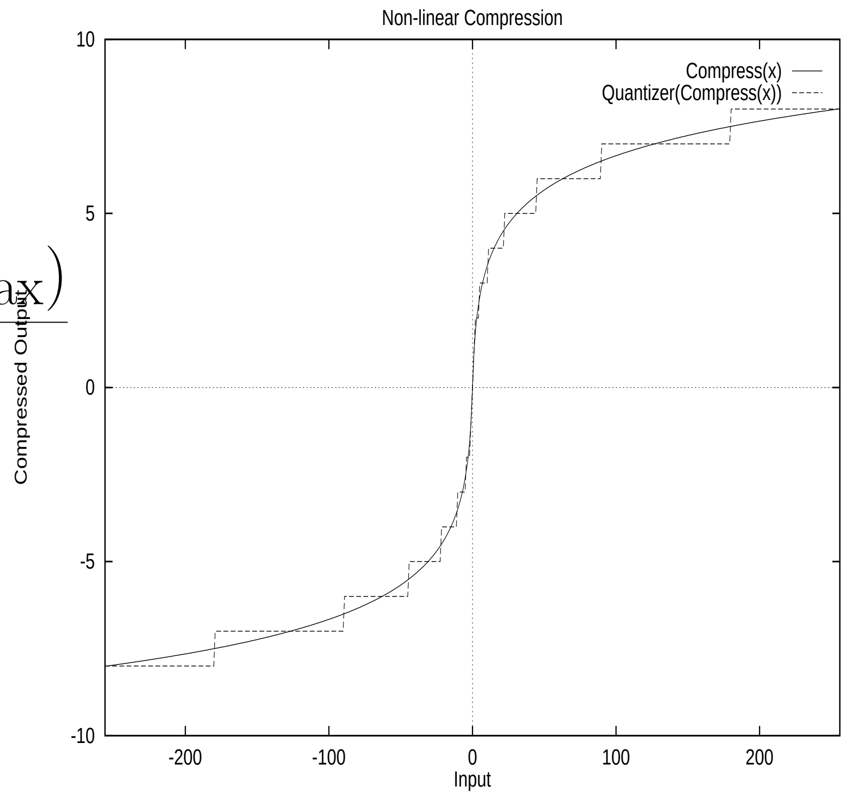
- **非線形量子化 (不均一量子化):**
  **非線形変換 + 均一量子化**
  **◆例: $\mu$圧伸 (ディジタル電話回線など)**

$$Y = 128\mathrm{sgn}(X)\frac{\log(1+\mu|X|/x_{\max})}{\log(1+\mu)}$$

$X$: **線形入力、**$Y$: **対数圧縮 PCM 符号.**

$x_{\max} = 8192, \mu = 255$

**→ 大振幅では量子化雑音も大、**
**小振幅では量子化雑音も小。**
**→ マスキング。S/N 比**




- **PCM 電話回線: 8kHz サンプリング × 8bit $\mu$圧伸量子化 = 64kb/s**
  **Cf. CD 録音: 2 チャネル × 44.1kHz サンプリング × 16 bit 直線量子化 = 1.4112Mb/s**