

Aodio Signal Processing

TANIGUCHI Taichi

October 16, 2018

Contents

I	音声データ	1
1	音の波形	2
2	音声データの保存	2
3	音声ファイル	3
II	フーリエ変換	3
4	離散フーリエ変換	3
4.1	離散フーリエ変換	3
4.2	逆離散フーリエ変換	6
4.3	離散フーリエ変換の性質	6
4.3.1	線形性	7
4.3.2	周期性	7
4.3.3	対称性	8
4.4	畳み込み定理	9
4.5	パーセバルの定理	9

Part I

音声データ

1 音の波形

音は波で表される。

周波数とは、音の波が一秒間に何回振動するかを表している。単位は、Hz である。一回の振動の長さを波長という。

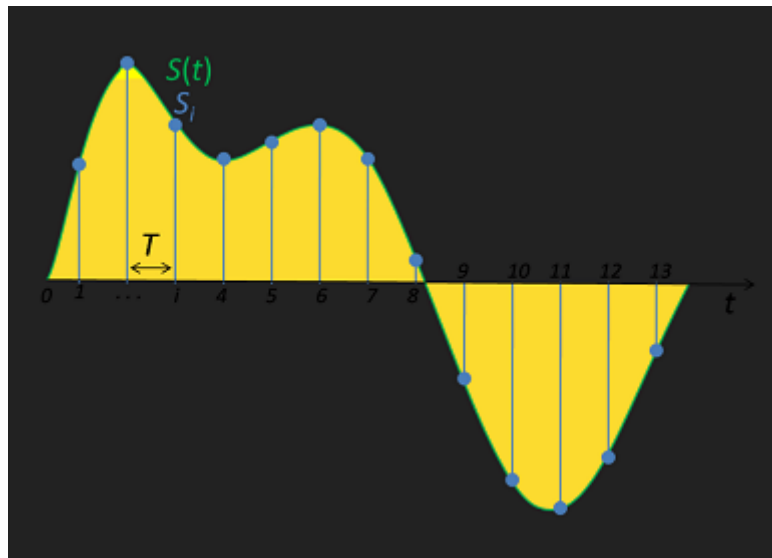
音の大きさは、振幅が大きいほど大きくなる。また、音の高さは振動数が多いほど高くなる。この二つの要素で音声は表されている。

2 音声データの保存

音声をデータについて保存する。まずは標本化定理をみる。[1] では次のように解説されている。

アナログ信号をデジタル信号へと変換する際に、どの程度の間隔で標本化（サンプリング）すればよいかを定量的に示す定理

標本化とは、一定の間隔で信号（音の波）を測定することで、信号を離散化して収集することである。間隔 T で標本化したとき、 $\frac{1}{T}$ をサンプリング周波数といい、 f_s で表す。標本化定理では、「周波数低域幅、つまり周波数がとりうる値が f_s 未満に制限された信号は、 f_s の 2 倍以上の標本化周波数で標本化すれば、その標本化された列から元の信号が一意に復元できることを指している。



単位時間あたりに標本を採る頻度をサンプリング周波数という。

3 音声ファイル

[2] では、音声をデータとして保存する際、以下のように行われていると説明されている。

音声をデジタルデータとして格納するには、一般に音声の波形を電圧で表したものを個々のチャンネル毎に特定の量子化ビット数（ビット深度）と一定の間隔（サンプリング周波数）で標本化する。

ここでは、一般的に wav ファイルを用いる。wav ファイルには、任意のサンプリング周波数とビットレートのデータが格納されている。

ビットレートとは、単位時間あたりに転送または処理されるビット数である。

Part II

フーリエ変換

4 離散フーリエ変換

4.1 離散フーリエ変換

離散化された信号 x の離散フーリエ変換とは、

$$\forall k = 0, 1, \dots, N-1 \quad X[k] = \sum_{n=0}^N x[n] e^{-i2\pi kn/N}$$

と表される。

ただし、

- n : 離散化された時間の index
- k : 周波数の index
- $w_k = \frac{2\pi k}{N}$: 単位時間あたりの角周波数
- $f_k = \frac{f_s k}{N}$: 周波数

また、

$$s_k = e^{-i2\pi kn/N} = \cos(2\pi kn/N) - i \sin(2\pi kn/N)$$

とすると、離散フーリエ変換は次のように内積を用いて表すことができる。

$$X[k] = \langle x, s_k \rangle$$

例えば、正弦波の離散フーリエ変換を考えてみる。正弦波とは、

$$x_1[n] = \exp \frac{i2\pi k_0 n}{N} \quad \text{for } n = 0, 1, 2, \dots, N-1$$

のように表される。信号を離散フーリエ変換をすると、

$$\begin{aligned} X_1[k] &= \sum_{n=0}^{N-1} x_1[n] \exp \frac{-i2\pi kn}{N} \\ &= \sum_{n=0}^{N-1} \exp \frac{i2\pi k_0 n}{N} \exp \frac{-i2\pi kn}{N} \\ &= \sum_{n=0}^{N-1} \left(\exp \frac{-i2\pi(k-k_0)n}{N} \right) \\ &= \frac{1 - \exp(-i2\pi(k-k_0))}{1 - \exp(-i2\pi(k-k_0)/N)} \\ &= \frac{1 - \exp(\frac{-i2\pi}{N} N(k-k_0))}{1 - \exp(-i2\pi(k-k_0)/N)} \\ &= \frac{1 - (\cos(\frac{2N(k-k_0)\pi}{N}) - i \sin(\frac{2N(k-k_0)\pi}{N}))}{1 - (\cos(\frac{2(k-k_0)\pi}{N}) - i \sin(\frac{2(k-k_0)\pi}{N}))} \\ &= N\delta_{k=k_0} \end{aligned}$$

ただし、

$$\delta_{k=k_0} = \begin{cases} 1 & \text{if } k = k_0 \\ 0 & \text{if } k \neq k_0 \end{cases}$$

より一般的に任意に位相 ϕ だけずれたときは正弦波は

$$x_2[n] = \exp(i(2\pi f_0 n + \phi))$$

と表され、その離散フーリエ変換は

$$\begin{aligned} X_2[k] &= \sum_{n=0}^{N-1} x_2[n] \exp(-i2\pi kn/N) \\ &= \sum_{n=0}^{N-1} \exp(i(2\pi f_0 + \phi)) \exp \frac{-i2\pi kn}{N} \\ &= \exp i\phi \sum_{n=0}^{N-1} \exp(-i2\pi(\frac{k}{N} - f_0)n) \\ &= \exp i\phi \frac{1 - \exp(-i2\pi(\frac{k}{N} - f_0))}{1 - \exp(-i2\pi(\frac{k}{N} - f_0))} \end{aligned}$$

と表される。

この正弦波の式は極刑式によって、複素数で表された正弦波である。一方、高校物理
までで見る正弦波の式は

$$x_3[n] = A_0 \cos\left(\frac{2\pi k_0 n}{N}\right)$$

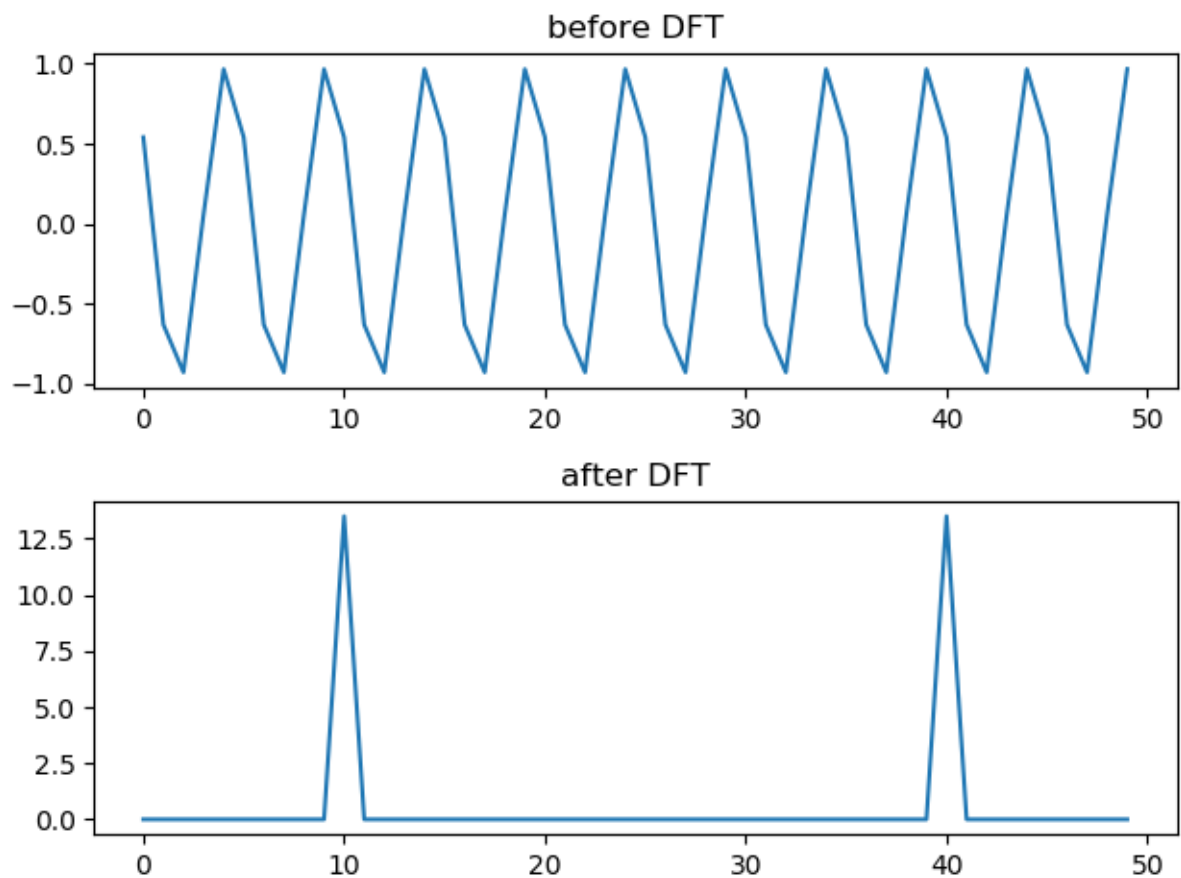
であり、オイラーの方程式を用いると

$$A_0 \cos\left(\frac{2\pi k_0 n}{N}\right) = \frac{A_0}{2} \exp\left(\frac{i2\pi k_0 n}{N}\right) + \exp\left(-\frac{i2\pi k_0 n}{N}\right)$$

と表される。この場合の離散フーリエ変換は

$$\begin{aligned} \sum_{n=-N/2}^{N/2} x_3[n] \exp\left(\frac{-i2\pi k n}{N}\right) &= \sum_{n=-N/2}^{N/2} \left(\frac{A_0}{2} \exp\left(\frac{i2\pi k_0 n}{N}\right) + \exp\left(-\frac{i2\pi k_0 n}{N}\right)\right) \exp\left(\frac{-i2\pi k n}{N}\right) \\ &= \sum_{n=-N/2}^{N/2} \frac{A_0}{2} \exp\left(-\frac{i2\pi(k-k_0)n}{N}\right) \sum_{n=-N/2}^{N/2} \frac{A_0}{2} \exp\left(-\frac{i2\pi(k+k_0)n}{N}\right) \\ &= N \frac{A_0}{2} \delta_{k=k_0} \end{aligned}$$

ここに離散フーリエ変換を行ったときの波形を載せておく。



4.2 逆離散フーリエ変換

$X[k]$ の逆離散フーリエ変換とは、

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \exp\left(\frac{i2\pi kn}{N}\right)$$

と表される。これは、内積を用いると

$$X[k] = \langle X[k], s_k[n] \rangle$$

と表される。

4.3 離散フーリエ変換の性質

離散フーリエ変換の性質をいくつか挙げる。

4.3.1 線形性

信号 $x_1[n], x_2[n]$ をある信号とする。この時、

$$ax_1[n] + bx_2[n]$$

を離散フーリエ変換すると

$$\begin{aligned} DFT(ax_1[n] + bx_2[n]) &= \sum_{n=0}^{N-1} (ax_1[n] + bx_2[n]) \exp(-i2\pi kn/N) \\ &= \sum_{n=0}^{N-1} ax_1[n] \exp(-i2\pi kn/N) + \sum_{n=0}^{N-1} bx_2[n] \exp(-i2\pi kn/N) \\ &= a \sum_{n=0}^{N-1} x_1[n] \exp(-i2\pi kn/N) + b \sum_{n=0}^{N-1} x_2[n] \exp(-i2\pi kn/N) \end{aligned}$$

このように各信号の離散フーリエ変換の線形結合で表される。これを線形性という。

4.3.2 周期性

サンプル数を n_0 とする。この性質は波形が時刻 n における s 信号とそこから n_0 だけ移動した、つまり時刻 $n + n_0$ における信号をそれぞれフーリエ変換したときの関係を見ている。

$x[n - n_0]$ を離散フーリエ変換すると

$$\begin{aligned} DFT(x[n - n_0]) &= \sum_{n=0}^{N-1} x[n - n_0] \exp(-i2\pi kn/N) \\ &= \sum_{m=-n_0}^{N-1-n_0} x[m] \exp(-i2\pi k(m + n_0)/N) \\ &= \sum_{m=0}^{N-1} x[m] \exp(-i2\pi km/N) \exp(-i2\pi kn_0/N) \\ &= \exp(-i2\pi kn_0/N) \sum_{m=0}^{N-1} x[m] \exp(-i2\pi km/N) \\ &= \exp(-i2\pi kn_0/N) X[k] \end{aligned}$$

4.3.3 対称性

離散フーリエ変換は

$$\overline{X[-k]} = DFT(X[k])$$

という対称性を満たす。

$$\begin{aligned}\overline{X[-k]} &= \sum_{n=0}^{N-1} x[n] \exp(i2\pi(-k)/N) \\ &= \sum_{n=0}^{N-1} x[n] \exp(-i2\pi k/N) \\ &= X[k]\end{aligned}$$

よって、

$$\overline{X[k]} = X[-k]$$

も成り立つことがわかる。

また、フーリエ変換後の値を実部と虚部分けると
実部は共役をとっても変わらないので

$$\begin{aligned}Re(\overline{X[k]}) &= Re(X[-k]) \\ Re(\overline{X[k]}) &= Re(X[k])\end{aligned}$$

よって

$$Re(X[-k]) = Re(X[k])$$

これはフーリエ変換の実部は偶関数であることを意味している。

また、虚部は共役をとると-1 倍されるので、

$$\begin{aligned}Im(\overline{X[k]}) &= Im(X[-k]) \\ Im(\overline{X[k]}) &= Im(-X[k])\end{aligned}$$

よって

$$Im(X[-k]) = Im(-X[k]) = -Im(X[k])$$

これはフーリエ変換の虚部は奇関数であることを意味している。

4.4 畳み込み定理

離散フーリエ変換には代数的構造が見られることがある。

信号 $x_1[n]$ と $x_2[n]$ の畳み込み積分を記号 $*$ として表す。つまり、

$$x_1[n] * x_2[n] = \sum_{m=0}^{N-1} x_1[m]x_2[n-m]$$

とする。

畳み込み積分 $x_1[n] * x_2[n]$ の離散フーリエ変換は次のようになります。

$$\begin{aligned} DFT(x_1[n] * x_2[n]) &= \sum_{n=0}^{N-1} (x_1[n] * x_2[n]) \exp(-i2\pi kn/N) \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} x_1[m]x_2[n-m] \exp(-i2\pi kn/N) \\ &= \sum_{m=0}^{N-1} x_1[m] \sum_{n=0}^{N-1} x_2[n-m] \exp(-i2\pi kn/N) \\ &= \left(\sum_{m=0}^{N-1} x_1[m] \exp(-i2\pi km/N) \right) X_2[k] \\ &= X_1[k]X_2[k] \end{aligned}$$

このように代数的な構造が確認できる。

4.5 パーセバルの定理

$$\sum_{n=-N/2}^{N/2-1} |x[n]|^2 = \frac{1}{N} \sum_{n=-N/2}^{N/2-1} |X[k]|^2$$

この不等式はパーセバルの不等式と呼ばれている。この式が意味しているところは、ヒルベルト空間 H から自分自身との間のユニタリ作用素ということ表している。詳しくは関数解析学のフーリエ解析で紹介されている。

References

[1] <https://ja.wikipedia.org/wiki/>

[2] <https://ja.wikipedia.org/wiki/>

- [3] <http://www.eri.u-tokyo.ac.jp/people/knishida/Seismology/Timeseries.pdf>
- [4] <http://universe-of-sound.jp/page8.html>
- [5] <https://www.math.nagoya-u.ac.jp/~yamagami/teaching/fourier/fourier2013.pdf>
- [6] <https://www.overleaf.com/project/5bb05a9e489bb32c475defe4>