

# 信号処理

～ 第3部 非定常信号解析，ケプストラム解析～

横田 康成

平成 15 年 5 月 22 日

# 目次

第 1 章	非定常信号の時間 - 周波数解析	2
1.1	非定常信号のスペクトル	2
1.2	短時間スペクトルとガボール変換	4
1.3	ウィグナー分布	5
1.4	コーエンのクラス	8
1.5	演習の解答 (Matlab プログラム)	11
第 2 章	ウェーブレット変換	15
2.1	連続ウェーブレット変換	15
2.2	離散ウェーブレット変換と直交ウェーブレット変換	16
2.3	ウェーブレットの種類	17
第 3 章	多重解像度近似とウェーブレット変換	19
3.1	多重解像度近似	19
3.2	各解像度空間の直交展開	20
3.3	各解像度における離散表現の関係	21
3.4	各解像度空間の直交補空間とその直交展開	22
3.5	各直交補空間における離散表現の関係	23
3.6	直交ウェーブレット表現からの合成	24
3.7	サブバンドフィルタとの関係	25
3.8	Daubechies の直交ウェーブレット	26
第 4 章	ケプストラム解析とヒルベルト変換	29
4.1	パワーケプストラム	29
4.2	複素ケプストラム	34
4.3	ケプストラム解析と最小 (最大) 位相性, 線形位相性	35
4.4	ヒルベルト変換	39
4.5	解析信号	41

# 第1章 非定常信号の時間 - 周波数解析

## 1.1 非定常信号のスペクトル

時刻と共に信号に含まれる周波数成分，つまりスペクトルが変化してゆく信号  $x(t)$  を考えよう．一般に，信号の統計的性質が時刻により変化する信号を非定常信号 (non-stationary process)，もしくは時変信号 (time variant process) と呼ぶ．こうした非定常信号  $x(t)$  のある瞬間の時刻  $t_0$  でのスペクトルをどのように推定したらよいのだろうか．フーリエ変換は無数の時刻にわたる三角関数を基底関数とする展開であるため，こうした非定常信号をフーリエ変換しても，その信号の時刻  $t_0$  での瞬間のスペクトル（以下，瞬時スペクトル）を求めることはできない．そこで，時刻  $t_0$  の前後  $[t_0 - T/2, t_0 + T/2]$  の幅  $T$  の信号を切り出し，その区間でフーリエ変換することを考えよう．これにより，その区間におけるスペクトルが求められるから，時間幅  $T$  を無限に小さくしてゆけば，時刻  $t_0$  における瞬間スペクトルを厳密に求められるように思える．このことの真偽を以下で検証してみよう．

非定常信号を考える前に，角周波数  $\pm\omega_0$  のみに成分を持つ信号，すなわち正弦波（あるいは余弦波） $x(t) = \sin(\omega_0 t)$  の  $t = 0$  での瞬時スペクトルを推定することを考えよう．この信号  $x(t)$  を区間  $[-T/2, T/2]$  で切り出した信号を  $x_{T/2}(t)$  とする．これは，信号  $x(t)$  に，

$$P_{T/2}(t) = \begin{cases} 1, & -\frac{T}{2} \leq t \leq \frac{T}{2} \\ 0, & \text{others} \end{cases} \quad (1.1)$$

なる窓関数を乗じることにより， $x_{T/2}(t) = x(t)P_{T/2}(t)$  として表される．ここで，時間軸での積は周波数軸では畳み込み積分で表されることを思い出そう．フーリエ変換を  $\mathcal{F}$ ，畳み込み積分を  $*$  で表すとすると， $x_{T/2}(t) = x(t)P_{T/2}(t)$  のフーリエ変換は，

$$X_{T/2}(\omega) = \mathcal{F}[x_{T/2}(t)] = \mathcal{F}[x(t)] * \mathcal{F}[P_{T/2}(t)]$$

と表されることになる．窓関数  $P_{T/2}(t)$  のフーリエ変換は， $\mathcal{F}[P_{T/2}(t)] = T \operatorname{sinc}(T\omega/2) = (2/\omega) \sin(T\omega/2)$ ， $x(t)$  のフーリエ変換  $\mathcal{F}[x(t)]$  が  $\pm\omega_0$  にピークを持つデルタ関数  $(\delta(\omega - \omega_0) + \delta(\omega + \omega_0))$  となることから， $X_{T/2}(\omega)$  は，

$$\begin{aligned} X_{T/2}(\omega) &= \mathcal{F}[x(t)] * \mathcal{F}[P_{T/2}(t)] \\ &= (\delta(\omega - \omega_0) + \delta(\omega + \omega_0)) * T \operatorname{sinc}(T\omega/2) \\ &= T \operatorname{sinc}(T(\omega - \omega_0)/2) + T \operatorname{sinc}(T(\omega + \omega_0)/2) \end{aligned} \quad (1.2)$$

と表される．つまり，本来，デルタ関数であったスペクトルが sinc 関数として観測されることになる．sinc 関数の幅は， $T$  が小さいと大きくなるから，切り出される時間幅  $T$  が小さい程，広い周波数にわたりスペクトルが分散することになる．

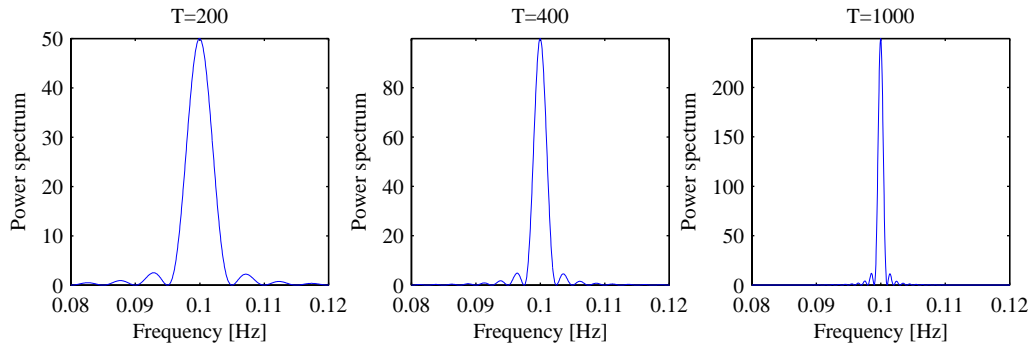


図 1.1: 周波数 0.1[Hz] の正弦波を時間幅  $T$  の窓関数 (矩形窓) を通して観測した際のパワースペクトル

演習 1: 周波数 0.1[Hz] の正弦波を時間幅  $T = 200, 400, 1000[\text{sec}]$  の窓関数を通して観測し, それをフーリエ変換することにより, パワースペクトル  $|X_{T/2}(\omega)|^2$  を求め, 窓関数の幅とスペクトルの形を比較してみよう. ただし, 計算機を用いて離散フーリエ変換により行う場合には, 標本化周波数を 1[Hz] としよう. また, 縦軸のスケールをあわせるため, 推定されたパワースペクトルを  $T$  で割り, 規格化しておこう.

解答: 結果を図 1.1 に示す. 見やすくするため, 0.1[Hz] 付近を拡大して描画してある. また, 負の周波数成分を無視してあるが,  $-0.1[\text{Hz}]$  付近にも, この図に示したものと同一成分がある. この図より, 時間幅  $T$  が短いほど, より広い周波数に渡りスペクトルが分散することがわかる. (Matlab プログラム ex1.m)

一般的な信号  $x(t)$  に対しても, 時間幅  $T$  で切り出された信号のスペクトルは, sinc 関数により真のスペクトル  $X(\omega) = \mathcal{F}[x(t)]$  が移動平均されるため, なまって観測されることになる. 時間幅  $T$  が小さいほど, そのなまり方は大きくなる. 例えば,  $\omega_0$  と  $\omega_1$  にピークを持つスペクトルは,  $\omega_0$  と  $\omega_1$  が接近している場合には, 時間幅  $T$  が小さいと分離して観測することが困難になることが起こりうる. すなわち, 時間幅  $T$  を小さくすると周波数分解能が低下することになる. 一方, 非定常信号のスペクトルの時間変化を高い時間分解能で捉えるためには, 時間幅  $T$  が小さいことが要求される. つまり, ある信号成分の存在する時刻と周波数を同時に知ることは決してできないことになる.

演習 1a: 周波数 0.1[Hz] の正弦波と周波数 0.102[Hz] の正弦波の和として表現される信号を時間幅  $T = 200, 400, 1000[\text{sec}]$  の窓関数を通して観測し, それをフーリエ変換することにより, パワースペクトル  $|X_{T/2}(\omega)|^2$  を求め, 窓関数の幅と 2 つの正弦波のスペクトルの分離の程度を比較してみよう. ただし, 計算機を用いて離散フーリエ変換により行う場合には, 標本化周波数を 1[Hz] としよう. また, 縦軸のスケールをあわせるため, 推定されたパワースペクトルを  $T$  で割り, 規格化しておこう.

解答: 結果を図 1.2 に示す. 窓関数の幅  $T$  が 200 の場合には, 2 つのスペクトルを分離することができないが,  $T = 400, 1000$  と大きくするに従い, 分離して観測できるようになる. (Matlab プログラム ex1a.m)

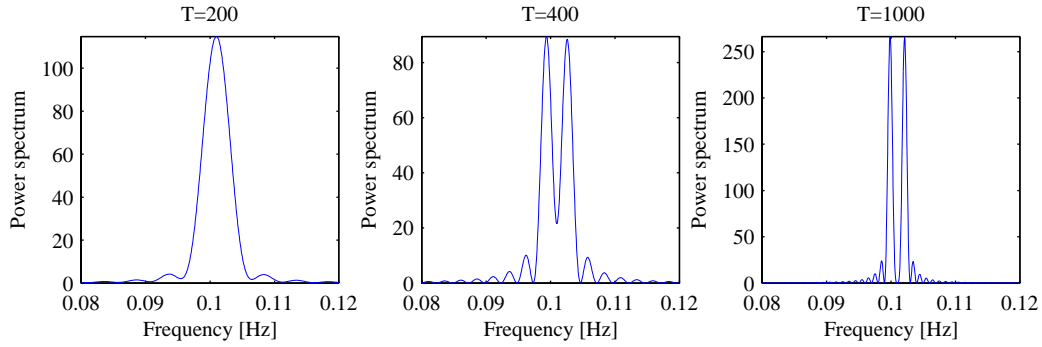


図 1.2: 周波数 0.1[Hz] の正弦波と周波数 0.102[Hz] の正弦波の和として表現される信号を時間幅  $T$  の窓関数 (矩形窓) を通して観測した際のパワースペクトル

## 1.2 短時間スペクトルとガボール変換

前節で述べた  $P_{T/2}(t)$  などのように、 $t = 0$  を中心とする時間軸で局在した窓関数  $h(t)$  を移動しながら、その窓関数を通してみた信号  $x(t)$  のフーリエ変換

$$X(t, \omega) = \int_{-\infty}^{\infty} x(t') h(t' - t) e^{-i\omega t'} dt' \quad (1.3)$$

は、短時間フーリエ変換 (Short term Fourier transform) と呼ばれる。短時間フーリエ変換の像  $X(t, \omega)$  は、短時間スペクトル (Short term Fourier spectrum)、あるいはスペクトログラム (Spectrogram) といわれ、時刻  $t$  と周波数  $\omega$  の関数となる。定常信号の場合、スペクトログラム  $X(t, \omega)$  は、時刻  $t$  に無関係になるが、非定常信号に対しては、時刻  $t$  に依存して瞬時スペクトルが変化することになる。非定常信号のスペクトル推定は、一般に、時間 - 周波数解析 (Time-Frequency Analysis) と呼ばれている。また、短時間スペクトルの絶対値の 2 乗は、短時間パワースペクトル、位相は短時間位相スペクトルと言われる。 $P_{T/2}(t)$  は矩形窓と呼ばれ、これを窓関数  $h(t)$  に用いた場合には、もっとも基本的な短時間フーリエ変換となるが、他に、カイザー窓、ブラックマン窓、ハミング窓などを用いた短時間フーリエ変換もある。

ところで、瞬時スペクトルは、その時刻  $t$  と周波数  $\omega$  を同時に精度よく求められないことを前節で述べた。ここで、その精度の上限はどの程度か、また、その上限を達成するにはどのような窓関数  $h(t)$  を用いればよいのだろうかという疑問が湧いてきたことであろう。次に、窓関数  $h(t)$ 、及びそのフーリエ変換  $H(\omega) = \mathcal{F}[h(t)]$  のそれぞれの軸での広がり方の関係を定量的に調べてみよう。広がり方を分散により表現するものとし、窓関数  $h(t)$  の時間軸での広がりを  $\Delta t^2$ 、 $H(\omega)$  の周波数軸での広がりを  $\Delta \omega^2$  と書くことにし、これらをそれぞれ次式で定義する。

$$\Delta t^2 = \frac{\int_{-\infty}^{\infty} t^2 |h(t)|^2 dt}{\int_{-\infty}^{\infty} |h(t)|^2 dt} \quad (1.4)$$

$$\Delta \omega^2 = \frac{\int_{-\infty}^{\infty} \omega^2 |H(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |H(\omega)|^2 d\omega} \quad (1.5)$$

これらの式を変形することにより，次式で与えられる関係式が得られる．

$$\Delta t \Delta \omega \geq \frac{1}{2} \quad (1.6)$$

この不等式は，周波数軸での広がり時刻軸での広がり，それらの積が一定値  $1/2$  を下回ることができないことを意味しており，時間 - 周波数解析における不確定性原理と呼ばれている．さらに，上式において等号を満たすのは，窓関数  $h(t)$  が正規分布

$$h(t) = ae^{-bt^2} \quad (1.7)$$

となる場合である．正規分布のフーリエ変換もまたフーリエ変換であることから， $H(\omega)$  もまた正規分布

$$H(\omega) = a\sqrt{\frac{\pi}{b}}e^{-\frac{\omega^2}{4b}} \quad (1.8)$$

になる．こうした正規分布を窓関数  $h(t)$  として用いた短時間フーリエ変換

$$G(t, \omega) = \int_{-\infty}^{\infty} x(t')ae^{-b(t'-t)^2} e^{-i\omega t'} dt' \quad (1.9)$$

は，ガボール変換 (Gabor transform) と呼ばれている．ガボール変換では，核関数  $g_{\omega, t_0}(t)$  の時間窓  $e^{-b(t-t_0)^2}$  が周波数  $\omega$  に依存しないため，各周波数で一定の周波数分解能を持つ．

演習 2：1000 秒間に，周波数が  $0.01[\text{Hz}]$  から  $0.4[\text{Hz}]$  まで直線的に変化する正弦波信号（チャープ信号という）と周波数が  $0.3[\text{Hz}]$  から  $0.1[\text{Hz}]$  まで直線的に変化するチャープ信号の和として与えられる信号の短時間パワースペクトルを求めてみよう．ただし，標準化周波数を  $1[\text{Hz}]$  とする．窓関数  $h(t)$  としては，矩形窓とブラックマン窓，時間幅  $T$  としては， $T = 32, 64, 128[\text{sec}]$  をそれぞれ選択した際の結果を比較せよ．  
 解答：スペクトログラムの算出結果を濃淡図として図 1.3 に示す．図中左列から  $T = 32, 64, 128[\text{sec}]$  の時間幅で，上段は矩形窓，下段はブラックマン窓を使用した際の短時間パワースペクトルを表す．ただし，表示を見やすくするため，パワースペクトルの平方根，つまり振幅スペクトルを描いた．窓関数の形状，時間幅により，得られる短時間パワースペクトルの局在性が異なることがわかるであろう．窓関数，時間幅は，対象に応じて，適切に選択しなければならないことになる．(Matlab プログラム ex2.m)

### 1.3 ウィグナー分布

ひとたび，短時間フーリエ変換（スペクトログラム）から離れて，ウィナーとヒンチンの定理を利用した定常過程のパワースペクトル密度の推定を思い出そう．ウィナーとヒンチンの定理は，定常確率過程  $x(t)$  の自己相関関数とパワースペクトル密度がフーリエ変換対の関係にあることを示したものであった．このことから，非定常確率過程  $x(t)$  の自己相関関数

$$R(t, \tau) = E[x(t + \tau/2)x^*(t - \tau/2)] \quad (1.10)$$

を  $\tau$  に関してフーリエ変換することにより，時刻  $t$  をパラメータに持つ非定常信号のパワースペクトル密度が求められるように思える．但し， $\cdot^*$  は， $\cdot$  の複素共役を表すが，確率過程  $x(t)$  が実数値をとる場合には意味を持たないため，以降，無視して考えてもよい．また， $E[\cdot]$  は期待値を表すが， $x(t)$  の複数の見本過程を観測できる場合，この期待値は集合平均で代用可能である．一方，一つの見本過程しか観測できない場

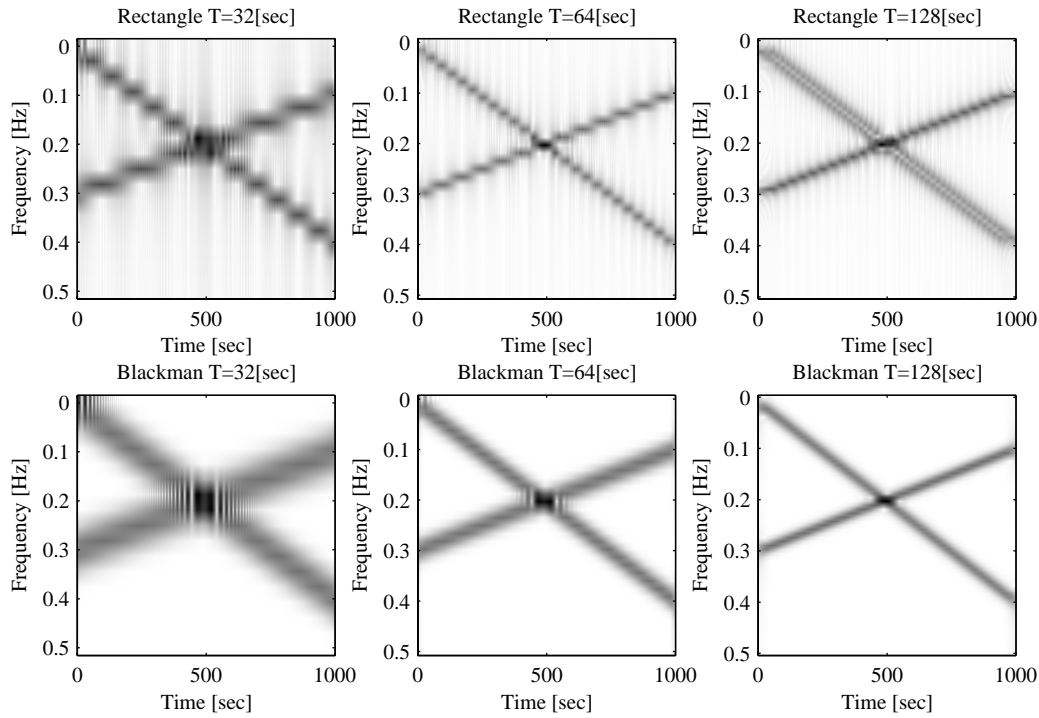


図 1.3: 0.01[Hz] から 0.4[Hz] まで変化するチャープ信号と 0.3[Hz] から 0.1[Hz] まで変化するチャープ信号の和の短時間パワースペクトル．上段は矩形窓，下段はブラックマン窓を使用．左列から， $T = 32, 64, 128$ [sec] の時間幅の窓関数を使用．

合でも， $x(t)$  が定常エルゴード過程であるならば，自己相関関数などの統計的性質が時刻  $t$  に依存しないから，期待値を時間平均に置き換え，

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t + \tau/2)x^*(t - \tau/2) dt \quad (1.11)$$

として自己相関関数を推定することができる．

一方，非定常過程では，その統計的性質が時刻  $t$  に依存するため，期待値を時間平均に置き換えることができないため，一つの見本過程からは，自己相関関数，等価的にパワースペクトル密度を求められないことになる<sup>1</sup>．そこで，式 (1.10) において期待値をとることをあきらめ，単に，

$$x(t + \tau/2)x^*(t - \tau/2) \quad (1.12)$$

を  $\tau$  に関してフーリエ変換して得られる時刻  $t$  と周波数  $\omega$  を変数に持つ分布

$$W(t, \omega) = \int_{-\infty}^{\infty} x(t + \tau/2)x^*(t - \tau/2)e^{-i\omega\tau} d\tau \quad (1.13)$$

を非定常過程のパワースペクトルの代用分布として用いることを考えてみよう．この分布  $W(t, \omega)$  は，ウィグナー分布 (Wigner distribution) と呼ばれている．

<sup>1</sup>ただし，複数の標本過程があったとしても，時間 - 周波数の不確定性原理があるために，時刻と周波数を同時に精度よく求められるわけではない

ウィグナー分布の時間平均は，

$$\begin{aligned}
 \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} W(t, \omega) dt &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-\infty}^{\infty} x(t + \tau/2) x^*(t - \tau/2) e^{-i\omega\tau} d\tau dt \\
 &= \int_{-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t + \tau/2) x^*(t - \tau/2) dt \right\} e^{-i\omega\tau} d\tau \\
 &= \int_{-\infty}^{\infty} R(\tau) e^{-i\omega\tau} d\tau \\
 &= |X(\omega)|^2
 \end{aligned} \tag{1.14}$$

となり， $x(t)$  を定常エルゴード過程とみなしたときの  $x(t)$  のパワースペクトルになる．一方，周波数  $\omega$  に関しての積分は，

$$\begin{aligned}
 \int_{-\infty}^{\infty} W(t, \omega) d\omega &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t + \tau/2) x^*(t - \tau/2) e^{-i\omega\tau} d\tau d\omega \\
 &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} e^{-i\omega\tau} d\omega \right\} x(t + \tau/2) x^*(t - \tau/2) d\tau \\
 &= \int_{-\infty}^{\infty} \delta(\tau) x(t + \tau/2) x^*(t - \tau/2) d\tau = |x(t)|^2
 \end{aligned} \tag{1.15}$$

となり，各時刻における瞬時パワーに等しくなる．時刻，および周波数に関する積分，すなわち周辺分布がこのような性質を持つ条件は周辺分布条件と呼ばれる．

本来，パワースペクトルは負値をとらないものであるが，ウィグナー分布  $W(t, \omega)$  は，パワースペクトルに似た性質を持つものの，負の値をとることもありうる．また，ウィグナー分布には，例えば，対象とする信号  $x(t)$  が周波数  $\omega_1, \omega_2$  の2つの成分を持つ場合，この信号のウィグナー分布は，周波数  $\omega_1, \omega_2$  のみならず，これらの周波数の中間の周波数にも非ゼロの値を持つという性質がある．同様に，時刻  $t_1, t_2$  で非ゼロの成分を持つ信号のウィグナー分布は，時刻  $t_1, t_2$  の他に，これらの時刻の中間の時刻においても非ゼロの値を持つ．こうした成分はクロス項と呼ばれている．離散時間系列を扱う場合には，ウィグナー分布の計算に離散フーリエ変換を利用することになるが，その場合，スペクトルが周期性を持つので，周期的に繰り返されるそれらのスペクトルの間にもクロス項が現れ，さらに複雑なことになる．

演習 3： 周波数が 0.01[Hz] から 0.4[Hz] まで直線的に変化する 1000 秒間のチャープ信号のウィグナー分布を求めてみよう．ただし，標本化周波数を 1[Hz] とする．また，また，このチャープ信号に，0.01[Hz] から 0.4[Hz] まで変化するチャープ信号を加えた信号のウィグナー分布を求め，クロス項の発生を確認してみよう．

解答： 周波数が 0.01[Hz] から 0.4[Hz] まで直線的に変化する 1000 秒間のチャープ信号のウィグナー分布  $W(t, \omega)$  を図 1.4(a) に示す．また，このチャープ信号に，周波数が 0.3[Hz] から 0.1[Hz] まで直線的に変化するチャープ信号を加えた信号のウィグナー分布  $W(t, \omega)$  を同図 (b) に示す．真の成分に加え，様々なクロス項が現れていることが分かるであろう．多くの周波数成分が時間と共に変化する複雑な非定常過程では，多数のクロス項を生じることになり，真の成分がどれであるか判別が難しくなる．したがって，ウィグナー分布を直接的に用いて，非定常過程のスペクトルを推定することは，一般的にはほとんど行われぬ．(Matlab プログラム ex3.m)



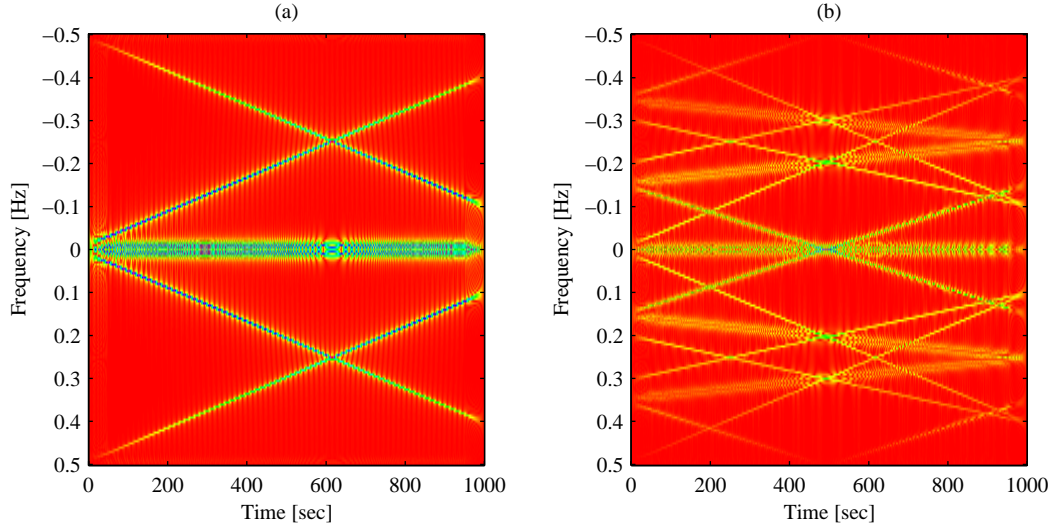


図 1.4: チャープ信号のウィグナー分布 . (a)0.01[Hz] から 0.4[Hz] まで変化するチャープ信号のウィグナー分布 , (b)0.01[Hz] から 0.4[Hz] まで変化するチャープ信号と 0.3[Hz] から 0.1[Hz] まで変化するチャープ信号の和のウィグナー分布

## 1.4 コーエンのクラス

クロス項の発生により, ウィグナー分布を, 時間 - 周波数解析に, 直接, 用いることは, 実際にはほとんどない. しかし, ウィグナー分布は, 短時間フーリエ変換をはじめとする様々な時間 - 周波数解析を関連づけ, それらを統合的に理解するためには, 非常に有用である. 以下, このことを学んでゆこう.

信号  $x(t)$  のウィグナー分布  $W(t, \omega)$  と任意の 2 変数関数  $\Phi(t, \omega)$  の 2 重畳み込み積分で表現可能な分布

$$C(t, \omega) = W(t, \omega) * \Phi(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t', \omega') \Phi(t - t', \omega - \omega') dt' d\omega' \quad (1.16)$$

の集合をコーエンのクラス (Cohen's class) という. 但し,  $*$  は畳み込み積分を表す. 2 変数関数  $\Phi(t, \omega)$  は核関数と呼ばれ, ウィグナー分布  $W(t, \omega)$  そのものは, 核関数がデルタ関数  $\Phi(t, \omega) = \delta(t)\delta(\omega)$  として表されるコーエンのクラスの特種なメンバである.

さて, 式 (1.16) における  $C(t, \omega)$ ,  $W(t, \omega)$ ,  $\Phi(t, \omega)$  をそれぞれ 2 重逆フーリエ変換したものを

$$M(\theta, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C(t, \omega) e^{i\theta t} e^{i\omega\tau} dt d\omega \quad (1.17)$$

$$A(\theta, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, \omega) e^{i\theta t} e^{i\omega\tau} dt d\omega \quad (1.18)$$

$$\phi(\theta, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi(t, \omega) e^{i\theta t} e^{i\omega\tau} dt d\omega \quad (1.19)$$

とおけば, 式 (1.16) は, 畳み込み積分を用いずに

$$M(\theta, \tau) = A(\theta, \tau)\phi(\theta, \tau) \quad (1.20)$$

と表現できる. また,  $A(\theta, \tau)$  は, 信号  $x(t)$  のウィグナー分布  $W(t, \omega)$  の 2 重逆フーリエ変換であるから, 式 (1.13) を用いれば,

$$A(\theta, \tau) = \int_{-\infty}^{\infty} x(t + \tau/2) x^*(t - \tau/2) e^{i\theta t} dt \quad (1.21)$$

と書ける．この  $A(\theta, \tau)$  は曖昧度関数とも呼ばれている．すなわち，コーエンのクラスとは，信号  $x(t)$  の曖昧度関数  $A(\theta, \tau)$  を任意の 2 変数関数  $\phi(\theta, \tau)$  で重みづけした後，2 重フーリエ変換して得られるすべての関数

$$C(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\theta, \tau) \phi(\theta, \tau) e^{-i\theta t} e^{-i\omega \tau} d\theta d\tau \quad (1.22)$$

からなる集合であると言い換えることができる．

ここで，窓関数  $h(t)$  を通してみた信号  $x(t')h(t' - t)$  の自己相関関数を

$$R(t, \tau) = \int_{-\infty}^{\infty} x(t' + \tau/2)h(t' - t + \tau/2)x^*(t' - \tau/2)h^*(t' - t - \tau/2)dt' \quad (1.23)$$

と定義しよう．ここで， $x(t)$  が定常過程であっても，窓関数を通して見た  $x(t')h(t' - t)$  は，時間軸で局在していることから，定常過程ではなくなることに注意しよう．したがって，定常エルゴード過程に対する自己相関関数とは異なり，式 (1.23) で定義される自己相関関数は，積分範囲の長さで規格化，すなわち平均を求めてはいない．しかしながら，ウィナー - ヒンチンの定理と同様に，信号  $x(t')h(t' - t)$  のフーリエ変換  $X(t, \omega)$  の絶対値の 2 乗  $|X(t, \omega)|^2$  は，式 (1.23) で定義される自己相関関数  $R(t, \tau)$  のフーリエ変換

$$|X(t, \omega)|^2 = \int_{-\infty}^{\infty} R(t, \tau) e^{-i\omega \tau} d\tau \quad (1.24)$$

に一致する．さらに，こうした短時間パワースペクトルは，窓関数  $h(t)$  のウィグナー分布を核関数  $\Phi(-t, \omega)$  に選んだ時のコーエンのクラス  $C(t, \omega)$  であることも示される．つまり，すべての短時間パワースペクトルは，コーエンのクラスに属する分布ということになる．また，すべての短時間パワースペクトルは，正值条件  $C(t, \omega) \geq 0$  を満たすことも示される．

非定常過程の短時間パワースペクトル表現として，コーエンのクラスに属するすべての分布が有効であるわけではない．パワースペクトルとして特に望まれる性質としては，正值条件を満たすこと，周辺分布を満たすことが挙げられる．しかしながら，これら 2 つの条件を同時に満足する分布は存在しないことがウィグナーにより証明されている．ウィグナー分布は前者を，短時間パワースペクトルは後者の条件を満たさない分布である．

前に述べたようにウィグナー分布にはクロス項が大きいという問題があるが，適当な核関数  $\Phi(t, \omega)$ ，等価的に  $\phi(\theta, \tau)$  を選びクロス項の低減を狙った分布が提案されている．チュウイ - ウィリアムス分布 (Chui - Williams distribution) は，こうした分布の代表的な分布であり，次式で表現される核関数  $\phi(\theta, \tau)$  を持つ．

$$\phi(\theta, \tau) = \exp\left(\frac{-(2\pi)^2 \theta^2 \tau^2}{\sigma}\right), \quad \sigma > 0 \quad (1.25)$$

時刻  $t_1$  と  $t_2$  の付近にそれぞれ周波数  $\omega_1$  と  $\omega_2$  の成分が存在している信号  $x(t)$  の曖昧度関数  $A(\theta, \tau)$  には， $(\theta, \tau)$  平面において， $(\omega_1 - \omega_2, t_1 - t_2)$  と  $(\omega_2 - \omega_1, t_2 - t_1)$  付近にクロス項の要因となる非ゼロ成分が出現することが分かっている．一方，信号の本来の成分は原点付近に現れる． $C(t, \omega)$  は，曖昧度関数  $A(\theta, \tau)$  と核関数  $\phi(\theta, \tau)$  の積の 2 重フーリエ変換で表されるから，核関数  $\phi(\theta, \tau) = 1$  を持つウィグナー分布では，クロス項の要因をすべて捨ててしまうことになる．これに対し，チュウイ-ウィリアムス分布では，式 (1.25) から分かるように  $\theta$  軸，あるいは  $\tau$  軸から離れるに従い急速に減衰する核関数  $\phi(\theta, \tau)$  を持つから，クロス項の要因を拾いにくくなり，その出現を抑えることが可能になる．

演習 4： 身近にある非定常信号に対し，短時間フーリエ変換を用いて時間 - 周波数解析を行ってみよう．窓関数としては，矩形窓，ブラックマン窓をそれぞれ使ってみよう．また，いずれも，窓関数の幅は  $T = 1024/8000[\text{msec}]$  とする．

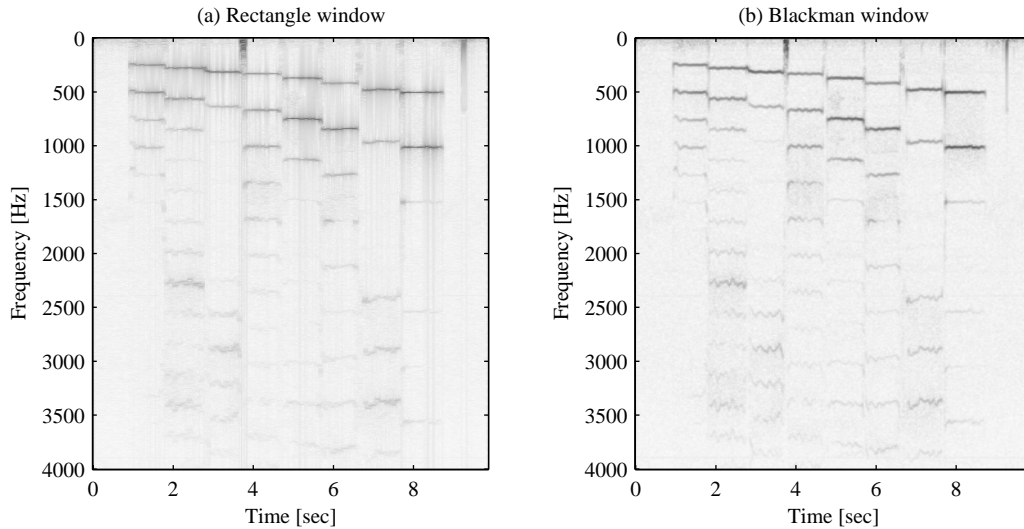


図 1.5: 音声信号の時間 - 周波数スペクトル推定例

解答: ある女性が「ドレミファソラシド」と発声した際の約 10[sec] の音声信号 (サンプリング周波数 8[kHz]) に対し、時間 - 周波数解析を行う。この信号の短時間パワースペクトルを図 1.4 に示す。(a) は矩形窓、(b) はブラックマン窓を窓関数に用いた際の短時間パワースペクトルである。これより、音声は、基本周波数とその高調波で構成され、これらは線状のスペクトルとして観測されることがわかる。基本周波数をピッチと呼んでいるおり、声帯の振動数である。ピッチは、男性は 100[Hz] ~ 200[Hz]、女性は 150[Hz] ~ 300[Hz] の範囲にあることが普通である。また、「ドレミファソラシド」と音階があがると、ピッチが高周波数側に移動してゆくことがわかる。また、「ミ」、「シ」は、ピッチは異なるものの、高調波の出方、つまり、スペクトルの包絡線は似ており、他の音とは異なることがわかる。これは、声帯で発生されたピッチとその高調波からなる信号が口腔の形状により、フィルタリングされるためである。母音に口腔形状が異なるため、スペクトルの包絡線もまた異なる。したがって、スペクトルの包絡線をケプストラム解析することにより、母音の認識が可能である。こうした声の短時間パワースペクトルは、一般に声紋と呼ばれ、人の声帯や口腔の形状などの影響を受け個々人により異なることから、個人を識別するために、犯罪捜査などに利用されている。

## 1.5 演習の解答 (Matlab プログラム)

1章で出題した演習を Matlab で行う際のスクリプト (プログラム) の例を以下に示す。以下のプログラムを実行すれば、図 1.1~1.5 を描くことができる。処理の流れの分かりやすさを重視するため、実行速度を犠牲にするプログラムとなっている。実行速度を重視するためには、ベクトル化したプログラムを書く必要がある。

```
function ex1
% 演習 1

set(0,'DefaultAxesFontName','TimesNewRoman');
set(0,'DefaultAxesFontSize',8);
set(0,'DefaultTextFontSize',8);

LL=[200 400 1000];
for k=1:3
    subplot(1,3,k)
    sub1(LL(k))
    title(['T=' num2str(LL(k))])
end

set(gcf,'PaperUnits','centimeters')
set(gcf,'PaperPosition',[0 6 16 5])
print -depsc ex1.eps

function sub1(L)
N=2^14;
x=zeros(1,N);
y=sin((1:L)*(2*pi*0.1));
x(1:length(y))=y;
xx=fftshift(fft(x));
px=abs(xx).^2/L;
f=[-N/2:N/2-1]/N;
plot(f,px,'Linewidth',0.2)
axesresize([1 0.8])
xlim([0.08 0.12]);
ylim([0 max(px)]);
xlabel('Frequency [Hz]');
ylabel('Power spectrum');

-----

function ex1a
% 演習 1a

set(0,'DefaultAxesFontName','TimesNewRoman');
set(0,'DefaultAxesFontSize',8);
set(0,'DefaultTextFontSize',8);

LL=[200 400 1000];
for k=1:3
    subplot(1,3,k)
    sub1(LL(k))
    title(['T=' num2str(LL(k))])
end

set(gcf,'PaperUnits','centimeters')
set(gcf,'PaperPosition',[0 6 16 5])
print -depsc ex1a.eps

function sub1(L)
N=2^14;
x=zeros(1,N);
y=sin((1:L)*(2*pi*0.1))+sin((1:L)*(2*pi*0.102));
x(1:length(y))=y;
xx=fftshift(fft(x));
```

```

px=abs(xx).^2/L;
f=[-N/2:N/2-1]/N;
plot(f,px,'Linewidth',0.2)
axesresize([1 0.8])
xlim([0.08 0.12]);
ylim([0 max(px)]);
xlabel('Frequency [Hz]');
ylabel('Power spectrum');

-----

function ex2
% 演習 2

set(0,'DefaultAxesFontName','TimesNewRoman');
set(0,'DefaultAxesFontSize',8);
set(0,'DefaultTextFontSize',8);

t=0:999;
x=chirp(t,0.01,t(end),0.4)+chirp(t,0.3,t(end),0.1);

subplot(2,3,1)
nfft=32;
window=rectwin(nfft);
stft(x,nfft>window);
title(['Rectangle T=' int2str(nfft) '[sec]']);

subplot(2,3,2)
nfft=64;
window=rectwin(nfft);
stft(x,nfft>window);
title(['Rectangle T=' int2str(nfft) '[sec]']);

subplot(2,3,3)
nfft=128;
window=rectwin(nfft);
stft(x,nfft>window);
title(['Rectangle T=' int2str(nfft) '[sec]']);

subplot(2,3,4)
nfft=32;
window=blackman(nfft);
stft(x,nfft>window);
title(['Blackman T=' int2str(nfft) '[sec]']);

subplot(2,3,5)
nfft=64;
window=blackman(nfft);
stft(x,nfft>window);
title(['Blackman T=' int2str(nfft) '[sec]']);

subplot(2,3,6)
nfft=128;
window=blackman(nfft);
stft(x,nfft>window);
title(['Blackman T=' int2str(nfft) '[sec]']);

set(gcf,'PaperUnits','centimeters')
set(gcf,'PaperPosition',[0 6 16 10])
print -depsc ex2.eps

function stft(x,nfft>window)
x=[zeros(1,nfft/2) x zeros(1,nfft/2)];
[y,f,t]=specgram(x,nfft,1>window,nfft-1);
zz=abs(y);
zzmax=max(zz(:));
zzmin=min(zz(:));
pz=fix((zz-zzmin)/(zzmax-zzmin)*256);
pz=uint8(255-pz);
image(t,f,pz)

```

```

colormap(gray(256));
ylabel('Frequency [Hz]');
xlabel('Time [sec]');

-----

function ex3
% 演習 3

set(0,'DefaultAxesFontName','TimesNewRoman');
set(0,'DefaultAxesFontSize',8);
set(0,'DefaultTextFontSize',8);

t=0:999;
subplot(1,2,1)
x=chirp(t,0.01,t(end),0.4);
wigner(x);
title('a')

subplot(1,2,2)
x=chirp(t,0.01,t(end),0.4) + chirp(t,0.3,t(end),0.1);
wigner(x);
title('b');

set(gcf,'PaperUnits','centimeters')
set(gcf,'PaperPosition',[0 6 16 7])
print -depsc2 ex3.eps

function wigner(x)
n=length(x);
maxTau=n/10; % less than n
x=[zeros(n,1) ; x(:) ; zeros(n,1)];
z=zeros(n,2*maxTau+1);

for t=1:n
    for tau=-maxTau:maxTau;
        z(t,tau+maxTau+1) = x(t+n+ceil(tau/2))*x(t+n-ceil(tau/2));
    end
end

zz=abs(fft(z,[],2));
zz=fftshift(zz,2)';
zzmax=max(zz(:));
zzmin=min(zz(:));
pz=fix((zz-zzmin)/(zzmax-zzmin)*256);
pz=uint8(pz);
f=0:size(pz,2)-1;
image([0 1000],[-0.5 0.5],pz)
colormap(hsv(256));
ylabel('Frequency [Hz]');
xlabel('Time [sec]');

function y=shift(x,n)
x=x(:)';
y=[x(n:end) x(1:n-1)];

-----

function ex4
% 演習 4

set(0,'DefaultAxesFontName','TimesNewRoman');
set(0,'DefaultAxesFontSize',8);
set(0,'DefaultTextFontSize',8);

[x,fs,nbits]=wavread('onkai.ume.wav');

subplot(1,2,1)
stft(x,1024,boxcar(1024),fs);
title('a) Rectangle window')

```

```

subplot(1,2,2)
stft(x,1024,blackman(1024),fs);
title('(b) Blackman window');

set(gcf,'PaperUnits','centimeters')
set(gcf,'PaperPosition',[0 6 16 7])
print -depsc2 ex4.eps

function pz=stft(x,nfft>window,fs)
x=x(:);
x=[zeros(nfft/2,1) ; x ; zeros(nfft/2,1)];
[y,f,t]=specgram(x,nfft,fs>window,nfft/4*3);
zz=abs(y);
zzmax=max(zz(:));
zzmin=min(zz(:));
pz=fix(((zz-zzmin)/(zzmax-zzmin)).^0.35*255);
pz=uint8(255-pz);
image(t,f,pz)
colormap(gray(256));
xlabel('Time [sec]');
ylabel('Frequency [Hz]');

```

## 第2章 ウェーブレット変換

### 2.1 連続ウェーブレット変換

さざなみのような振動した小さな波  $\psi(t)$  を考えよう．さざなみであるから，ある幅で局在している必要がある．そこで，このさざなみは2乗可積分であると仮定しよう．この波を時間軸で  $a$  倍に伸縮する操作と時間軸に沿って  $b$  だけ平行移動する操作により，様々なスケール  $a$  と位置  $b$  を持ったさざなみ状の関数の集合

$$\{ \psi_{a,b}(t) \mid a(\neq 0), b \in \mathbf{R} \}, \quad \psi_{a,b}(t) \equiv \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \quad (2.1)$$

を作成する． $\mathbf{R}$  は，実数からなる集合を表す．上式において，さざなみ  $\psi(t)$  を時間軸方向に  $a$  倍すると同時に振幅方向に  $1/\sqrt{|a|}$  倍しているが，これは， $\psi_{a,b}(t)$  の2乗積分値，すなわちエネルギーを一定にするための規格化である．

さて，これらのさざなみ  $\psi_{a,b}(t)$  と信号  $x(t)$  の内積

$$W(a,b) = \int_{-\infty}^{\infty} \psi_{a,b}(t)x(t) dt \quad (2.2)$$

を求めたとする．パラメータ  $a, b$  は，それぞれスケール，位置を表しており，おおよそ周波数，時刻にそれぞれ対応させることができる．したがって，さざなみ状の関数と信号  $x(t)$  の内積は，各時刻での周波数特性を求めることに似ている．すなわち，非定常信号の時間・周波数解析のようなことを行えることになる．こうしたさざなみ状の関数  $\psi(t)$  はウェーブレット (wavelet) と呼ばれている．また，式 (2.2) のような  $x(t)$  から  $W(a,b)$  への変換を広義にウェーブレット変換 (wavelet transform) と呼んでいる．

短時間フーリエ変換やガボール変換においては，窓関数  $h(t)$  の幅が一定で三角関数の周波数のみを変化させたのに対し，ウェーブレット変換では，スケールを縮めることにより，周波数を高めるだけでなく，窓関数の幅を短くすることにもなるため，高周波領域において時間分解能が改善する．これが，短時間フーリエ変換やガボール変換とウェーブレット変換の大きな違いの一つである．

次に，このウェーブレット変換に逆変換が存在するかどうかを考えてみよう．逆変換が存在するということは，係数  $W(a,b)$  から元の信号  $x(t)$  を再構成できる，すなわち，

$$x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(a,b) \tilde{\psi}_{a,b}(t) da db \quad (2.3)$$

なる関数系  $\tilde{\psi}_{a,b}(t)$  が存在することである．この関数系  $\tilde{\psi}_{a,b}(t)$  が，1つの関数  $\tilde{\psi}(t)$  を定め， $\psi(t)$  から  $\psi_{a,b}(t)$  を作成した時と同じ方法で，

$$\tilde{\psi}_{a,b}(t) \equiv \frac{1}{\sqrt{|a|}} \tilde{\psi}\left(\frac{t-b}{a}\right), \quad \forall a(\neq 0), b \in \mathbf{R} \quad (2.4)$$

として生成できるとき，この逆変換を特に逆ウェーブレット変換 (inverse wavelet transform) と呼ぶ．また， $\tilde{\psi}(t)$  を  $\psi(t)$  の双対ウェーブレット (dual wavelet) という．双対ウェーブレットが存在する場合でも，それが  $\psi(t)$  から一意に求められるかどうかはわからない．



一般に、2乗可積分関数  $\psi(t)$  に対し、そのフーリエ変換  $\Psi(f)$  が

$$C_\psi \equiv \int_{-\infty}^{\infty} \frac{|\Psi(f)|^2}{|f|} df < \infty \quad (2.5)$$

を満たす条件をアドミッシブル条件 (admissible condition) という。  $\psi(t)$  が絶対可積分ならばアドミッシブル条件は、  $\Psi(0) = 0$ 、すなわち、  $\int_{-\infty}^{\infty} \psi(t) dt = 0$  に等価である。ウェーブレット  $\psi(t)$  は、アドミッシブル条件を満たす、つまり、平均が0ならば、基本ウェーブレットと呼ばれる。基本ウェーブレット  $\psi(t)$  では、ウェーブレット  $\psi(t)$  の複素共役  $\psi^*(t)$  が双対ウェーブレット  $\tilde{\psi}(t)$  の一つになることが知られている。したがって、式 (2.3) で与えられる逆ウェーブレット変換は、

$$x(t) = \frac{1}{C_\psi} \frac{1}{a^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(a, b) \psi_{a,b}^*(t) da db \quad (2.6)$$

と書ける。連続ウェーブレットと基本ウェーブレットの関係を図 2.2(a) に示す。

## 2.2 離散ウェーブレット変換と直交ウェーブレット変換

さて、基本ウェーブレットを用いたウェーブレット変換に対し、逆ウェーブレット変換が存在することが示された後は、ウェーブレット変換が直交変換となりうるかどうかに興味を湧くであろう。しかしながら、パラメータ  $a, b$  が連続である限り、いかなるウェーブレットに対しても直交性を満足しないことが分かっている。直交変換を実現するためには、パラメータ  $a, b$  を離散化することを考えなければならない。その準備として、関数展開におけるフレームについて学ぼう。

2乗可積分なすべての実関数を元とする線形空間を  $L^2(\mathbf{R})$ 、整数からなる集合を  $\mathbf{Z}$  として表すことにする。関数系  $\{\psi_k(t) \in L^2(\mathbf{R}) \mid k \in \mathbf{Z}\}$  は、任意の関数  $x(t) \in L^2(\mathbf{R})$  に対し、次式を満たす2つの実数  $0 < A, B < \infty$  が存在するならば、 $L^2(\mathbf{R})$  に対するフレーム (frame) と呼ばれる。

$$A \leq \frac{\sum_{k \in \mathbf{Z}} \left( \int_{-\infty}^{\infty} \psi_k(t) x(t) dt \right)^2}{\int_{-\infty}^{\infty} |x(t)|^2 dt} \leq B \quad (2.7)$$

上式において、 $A, B$  は、フレーム限界 (frame bound) と呼ばれる。 $A = B$  の時、フレームはタイト (tight) であるといい、直交関数展開におけるエネルギー保存に相当する。一つでも  $\psi_k(t)$  を取り除くとフレームにならなくなる場合、フレームは完全 (complete) であるといい、有限次元ベクトル空間における基底に相当する。また、 $A > 0$  であることから、フレームによる  $x(t)$  の展開係数から元の信号  $x(t)$  を再構成できることになる。すなわち、任意の  $x(t) \in L^2(\mathbf{R})$  に対し、

$$x(t) = \sum_{k \in \mathbf{Z}} \left( \int_{-\infty}^{\infty} \psi_k(t) x(t) dt \right) \tilde{\psi}_k(t) \quad (2.8)$$

を満足するフレーム  $\{\tilde{\psi}_k(t) \mid k \in \mathbf{Z}\}$  が存在することになる。フレーム  $\{\tilde{\psi}_k(t) \mid k \in \mathbf{Z}\}$  をフレーム  $\{\psi_k(t) \mid k \in \mathbf{Z}\}$  の双対フレーム (dual frame) という。一般に、双対フレームは無数個存在するが、フレーム  $\{\psi_k(t) \mid k \in \mathbf{Z}\}$  が線形独立ならば、唯一の双対フレームが定められる。この際、このフレームをリース基底 (Riesz basis) という。フレームとリース基底の関係を図 2.2(b) に示す。

話をウェーブレットに戻そう。ウェーブレット  $\psi(t)$  に対し、パラメータ  $a, b$  を離散化して生成されるウェーブレット関数系

$$\{\psi_{i,j}(t) \mid i, j \in \mathbf{Z}\}, \quad \psi_{i,j}(t) \equiv 2^{i/2} \psi(2^i t - j) \quad (2.9)$$

による  $x(t)$  のウェーブレット変換を

$$W_{i,j} = \int_{-\infty}^{\infty} \psi_{i,j}(t)x(t) dt, \quad i, j \in \mathbf{Z} \quad (2.10)$$

とする．これを離散ウェーブレット変換 (discrete wavelet transform) という．ここで， $\psi_{i,j}(t)$  が  $L^2(\mathbf{R})$  に対するフレームになるならば，式 (2.8) より

$$x(t) = \sum_{i,j \in \mathbf{Z}} W_{i,j} \tilde{\psi}_{i,j}(t) \quad (2.11)$$

として再構成できる，すなわち逆変換が存在することになる．しかし， $\tilde{\psi}_{i,j}(t)$  がある一つの関数  $\tilde{\psi}(t)$  をウェーブレットとして生成できる，すなわち，逆変換が逆離散ウェーブレット変換であることは保証されない．しかし，ウェーブレット関数系  $\{\psi_{i,j}(t) \mid i, j \in \mathbf{Z}\}$  が  $L^2(\mathbf{R})$  に対するリース基底であれば，唯一の双対なリース基底  $\{\tilde{\psi}_{i,j}(t) \mid i, j \in \mathbf{Z}\}$  が定まり，更に，これらが

$$\int_{-\infty}^{\infty} \psi_{i,j}(t)\tilde{\psi}_{k,l}(t) dt = \begin{cases} 1, & i = k \text{ and } j = l \\ 0, & \text{others} \end{cases} \quad (2.12)$$

を満たすならば，双対リース基底  $\tilde{\psi}_{i,j}(t)$  は，一つの関数  $\tilde{\psi}(t)$  から

$$\tilde{\psi}_{i,j}(t) = 2^{i/2}\tilde{\psi}(2^i t - j), \quad i, j \in \mathbf{Z} \quad (2.13)$$

として生成できることが示されている．

リース基底が自分自身に双対，すなわち  $\psi_{i,j}^*(t) = \tilde{\psi}_{i,j}(t)$  ならば，式 (2.12) より

$$\int_{-\infty}^{\infty} \psi_{i,j}(t)\psi_{k,l}^*(t) dt = \begin{cases} 1, & i = k \text{ and } j = l \\ 0, & \text{others} \end{cases} \quad (2.14)$$

となり，リース基底が正規直交関数系を形成することになる．この時，そのリース基底を生成するウェーブレット  $\psi(t)$  を直交ウェーブレット (orthogonal wavelet) という．直交ウェーブレット関数系による関数  $x(t)$  の変換

$$W_{i,j} = \int_{-\infty}^{\infty} \psi_{i,j}(t)x(t) dt = 2^{i/2} \int_{-\infty}^{\infty} \psi(2^i t - j)x(t) dt, \quad i, j \in \mathbf{Z} \quad (2.15)$$

および，逆変換

$$x(t) = \sum_{i,j \in \mathbf{Z}} W_{i,j} \psi_{i,j}^*(t) = 2^{i/2} \sum_{i,j \in \mathbf{Z}} W_{i,j} \psi^*(2^i t - j) \quad (2.16)$$

を，それぞれ直交ウェーブレット変換 (orthogonal wavelet transform)，逆直交ウェーブレット変換 (inverse orthogonal wavelet transform) という．離散ウェーブレットと直交ウェーブレットの関係を図 2.2(c) に示す．

## 2.3 ウェーブレットの種類

最も簡単な直交ウェーブレットは，ウェーブレット変換という言葉が使われる以前から，そのウェーブレットとしての性質が知られていたハール関数 (Haar function)

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{others} \end{cases} \quad (2.17)$$

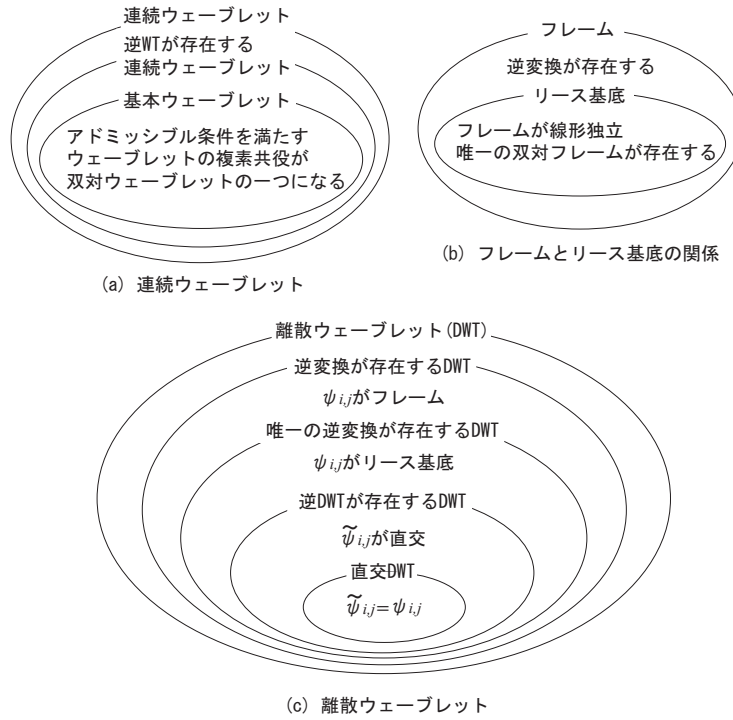


図 2.1: フレームとウェーブレットの関係

である．ハール関数により生成されたウェーブレットが正規直交性を満たすことは明らかであり，かつ局在性も満足している．しかし，ウェーブレット関数系が不連続であるため，展開の収束が遅く，限定された場合にしか実用にはならない．そこで，滑らかさをもちかつ局在性を持つウェーブレットを探す試みがなされた．

局在性を持たせるために急速に減衰させた関数では，滑らかさが失われることから，滑らかさと局在性は相反する性質である．さらに，直交変換としての性質も満たさなければならず，直交ウェーブレットは，相当に厳しい制約を受けている．具体的には，ウェーブレット関数系が正規直交性を有し，しかも  $\psi(t) \in C^m$ ， $|t| \rightarrow \infty$  で十分速く 0 になるならば，

$$\int_{-\infty}^{\infty} t^k \psi(t) dt = 0, \quad 0 \leq k \leq m \tag{2.18}$$

とならなければならないことが証明されている．Meyer は，すべてのモーメントが 0 となり，無限回微分可能，かつ急速に減衰する直交ウェーブレットを設計した．Mayer の直交ウェーブレットと並び，広く利用されている直交ウェーブレットとしては，Daubechies のウェーブレットがある．Daubechies のウェーブレットについては後述する．

# 第3章 多重解像度近似とウェーブレット変換

## 3.1 多重解像度近似

前に紹介した直交ウェーブレット変換を計算機で実現するためには、信号を連続時間信号としてではなく、離散時間信号として扱った方がはるかに便利である。しかしながら、直交ウェーブレット変換は、連続時間信号  $x(t)$  に対する直交変換であるから、離散時間信号に対し直接的に適用することはできない。ウェーブレット  $\psi(t)$  を信号  $x(t)$  に合わせて時刻  $t$  を離散化すれば良いように思うかもしれないが、連続な変数を持つ直交ウェーブレット関数系が、その変数を離散化した場合にも直交している保証はない。ちなみに、離散化された変数に対する関数系の直交性を選点直交性という。

離散時間信号に対する直交ウェーブレット変換を導くためには、多重解像度近似の考え方を新たに導入し、信号  $x(t)$  の標準化の問題に立ち返る必要がある。本章では、2乗可積分な関数空間  $L^2(\mathbf{R})$  に対する多重解像度近似について学ぶことにしよう。

まず、2乗可積分な関数をさまざまな解像度、すなわち精度で近似することを考える。便宜上、関数  $x(t) \in L^2(\mathbf{R})$  を解像度  $2^j$ 、 $j \in \mathbf{Z}$  で近似する作用素を  $A_{2^j}$  とし、近似された関数を  $A_{2^j}[x(t)]$  と表すことにする。解像度  $2^j$  は、値が小さいほど粗くなることを意味するものとする。

さて、解像度  $2^j$  に変換する作用素  $A_{2^j}$  が次の性質を持つものと仮定しよう。

1.  $A_{2^j}$  は、線形作用素であり、かつ、あるベクトル空間  $V_{2^j} \subset L^2(\mathbf{R})$  への射影子である。すなわち、 $A_{2^j}[x(t)]$  を解像度  $2^j$  で近似した  $A_{2^j}[A_{2^j}[x(t)]]$  は  $A_{2^j}[x(t)]$  に一致する。これを  $A_{2^j} \circ A_{2^j} = A_{2^j}$  と書く。したがって、ベクトル空間  $V_{2^j}$  は、 $L^2(\mathbf{R})$  に含まれるすべての関数を解像度  $2^j$  で近似したものの集合となる。
2. 作用素  $A_{2^j}$  は、ベクトル空間  $V_{2^j}$  への正射影子である。言い替えれば、 $A_{2^j}[x(t)]$  は、解像度  $2^j$  のすべての関数の中で、2乗ノルムの意味で  $x(t)$  にもっとも近い関数である。すなわち、

$$\int_{-\infty}^{\infty} |y(t) - x(t)|^2 dt \geq \int_{-\infty}^{\infty} |A_{2^j}[x(t)] - x(t)|^2 dt, \quad \forall y(t) \in V_{2^j} \quad (3.1)$$

となる。

3. 任意の関数  $x(t)$  の解像度  $2^{j+1}$  における表現は、その関数  $x(t)$  を解像度  $2^j$  で表現するために必要な情報を完全に有している。すなわち、

$$V_{2^j} \subset V_{2^{j+1}}, \quad \forall j \in \mathbf{Z} \quad (3.2)$$

となる。

4. 解像度  $2^j$  の関数空間  $V_{2^j}$  に属する関数を  $x(t)$  とした場合、関数  $x(2^k t)$  は解像度  $2^{j+k}$  の関数空間  $V_{2^{j+k}}$  に属する。すなわち、

$$x(t) \in V_{2^j} \Leftrightarrow x(2^k t) \in V_{2^{j+k}}, \quad \forall j, k \in \mathbf{Z} \quad (3.3)$$

となる .

5.  $A_{2^j}[x(t)]$  は , ある一定の長さあたり  $2^j$  点で完全に特徴づけられる . すなわち ,  $V_{2^j}$  から  $L^2(\mathbf{Z})$  への同型写像  $I$  が存在する .
6. 関数  $x(t) \in L^2(\mathbf{R})$  を  $2^{-j}$  の整数  $k$  倍の値  $2^{-j}k$  だけ平行移動した後 , 解像度  $2^j$  で近似した関数  $A_{2^j}[x(t-2^{-j}k)]$  は , 同じ関数  $x(t)$  を解像度  $2^j$  で近似した後 ,  $2^{-j}k$  だけ平行移動した関数  $A_{2^j}[x(t)]|_{t=t-2^{-j}k}$  に一致する . すなわち ,

$$A_{2^j}[x(t-2^{-j}k)] = A_{2^j}[x(t)]|_{t=t-2^{-j}k}, \quad \forall j, k \in \mathbf{Z} \quad (3.4)$$

となる . 性質 5. と組み合わせれば ,

$$I[A_{2^j}[x(t)]] = \alpha_i \Leftrightarrow I[A_{2^j}[x(t-2^{-j}k)]] = \alpha_{i-k}, \quad \forall i, j, k \in \mathbf{Z} \quad (3.5)$$

となる .

7. 関数  $x(t)$  の解像度  $2^j$  における表現では ,  $x(t)$  に関するなんらかの情報が失われている . 解像度を  $+\infty$  に増加すれば , 元の関数  $x(t)$  に収束する . 逆に , 解像度を 0 にすれば , 近似された関数の情報は限りなく少なくなり , 0 に収束する .

$$\left. \begin{aligned} \lim_{j \rightarrow +\infty} V_{2^j} &= \bigcup_{j=-\infty}^{+\infty} V_{2^j} \text{ is dense in } L^2(\mathbf{R}) \\ \lim_{j \rightarrow -\infty} V_{2^j} &= \bigcap_{j=-\infty}^{+\infty} V_{2^j} = \{0\} \end{aligned} \right\} \quad (3.6)$$

性質 1.-7. を満たす作用素  $A_{2^j}$  の集合は ,  $L^2(\mathbf{R})$  に属する任意の関数の解像度  $2^j$  における妥当な近似を与えることになる . また一般に , 性質 3.-7. を満たす任意のベクトル空間の集合  $\{V_{2^j} \mid j \in \mathbf{Z}\}$  を  $L^2(\mathbf{R})$  の多重解像度近似 (multi-resolution approximation) という .

## 3.2 各解像度空間の直交展開

$\{V_{2^j} \mid j \in \mathbf{Z}\}$  を  $L^2(\mathbf{R})$  における多重解像度近似とする . この時 , 関数系

$$\{2^{j/2}\phi(2^j t - n) \mid n \in \mathbf{Z}\} \quad (3.7)$$

が  $V_{2^j}$  の正規直交基底となるような関数  $\phi(t) \in V_1 \subset L^2(\mathbf{R})$  が一意に存在することが知られている . 関数  $\phi(t)$  は , スケーリング関数 (scaling function) と呼ばれる . 関数  $x(t) \in L^2(\mathbf{R})$  を解像度  $2^j$  で近似した関数  $A_{2^j}[x(t)]$  は , こうした正規直交基底を用いて ,

$$A_{2^j}[x(t)] = 2^{j/2} \sum_{n \in \mathbf{Z}} A_{2^j} x_n \phi(2^j t - n) \quad (3.8)$$

と表現可能である . ここで ,  $A_{2^j} x_n$  は , 展開係数であり , 上式の両辺に  $2^{j/2}\phi^*(2^j t - m)$  を乗じて  $t$  で積分し ,  $\{2^{j/2}\phi(2^j t - n) \mid n \in \mathbf{Z}\}$  の正規直交性を利用すれば ,

$$A_{2^j} x_n = 2^{j/2} \int_{-\infty}^{\infty} \phi^*(2^j t - n) A_{2^j}[x(t)] dt, \quad n \in \mathbf{Z} \quad (3.9)$$

として与えられる．ここで， $\{2^{j/2}\phi(2^j t - n) \mid n \in \mathcal{Z}\}$  が  $V_{2^j}$  の基底であることと， $A_{2^j}$  が  $V_{2^j}$  への正射影子であることから，式 (3.9) において， $A_{2^j}[x(t)]$  は，単に  $x(t)$  としても結果は同じになる．したがって，式 (3.9) は，

$$A_{2^j}x_n = 2^{j/2} \int_{-\infty}^{\infty} \phi^*(2^j t - n)x(t) dt, \quad n \in \mathcal{Z} \quad (3.10)$$

と書ける．これは， $x(t)$  の解像度  $2^j$  における離散表現とも呼ばれる．

式 (3.10) の意味を考えてみよう．式 (3.10) は，時間軸の正負の反転と伸縮を無視すれば， $2^{j/2}\phi^*(2^j t - n)$  を線形インパルス応答に持つフィルタに  $x(t)$  を通し，間隔  $2^{-j}$  毎の値を取り出すことを意味している．一般に，信号の標本化は，信号をナイキスト周波数以下の帯域に制限するため，低域通過フィルタ (LPF) に通した後，信号の一定間隔毎の値を取り出すことにより行われる． $2^{j/2}\phi(2^j t - n)$  は，解像度  $2^j$  の関数空間の基底関数，つまりそれ自身その解像度空間に属する関数であるから，ある程度滑らかであるはずなので，LPF の線形インパルス応答とみなすことができる．

こうしたことから，式 (3.10) は， $x(t)$  の標本化に似たことを行っているといえる．例えば， $2^{j/2}\phi(2^j t)$  が sinc 関数，すなわち  $2^{j/2}\phi(2^j t) = \sin(2^j \pi t)/(2^j \pi t)$  である場合，ナイキスト周波数で完全に遮断する理想 LPF を用いた理想的な標本化になる． $x(t)$  の解像度  $2^j$  における離散表現  $A_{2^j}x_n$  は，標本化の概念を一般化した  $x(t)$  の標本化系列となる．本章で，連続時間信号に立ち戻った理由は，多重解像度近似の立場から信号の標本化を考え直す必要があったからである．

現実の信号解析では，無限の解像度で信号  $x(t)$  を観測・記録できることは，まずありえない．観測される段階で信号  $x(t)$  が解像度  $2^j$  の信号  $A_{2^j}[x(t)]$  に劣化し，もしくは標本化するための前処理として LPF がかけられており，その後，標本化されることにより，その離散表現  $A_{2^j}x_n$  が得られると考えることが自然である．観測した時点での解像度  $2^j$  をどのように設定しても一般性を失うことはないので，本章では，離散時間信号を解像度  $2^0 = 1$  での離散表現  $A_1x_n$  とみなすことにする．

### 3.3 各解像度における離散表現の間の関係

観測される離散時間信号，つまり解像度  $2^0 = 1$  での  $x(t)$  の離散表現  $A_1x_n$  から，元々の信号  $x(t)$  を知ることなしに，それ以下の解像度での離散表現  $A_{2^j}x_n$ ， $j < 0$  が求められると便利である．関数  $x(t)$  の各解像度での離散表現  $A_{2^j}x_n$ ， $j < 0$  は，多重解像度近似の性質 3. 及び 5. より， $A_1x_n$  から完全に求められることは保証されている．次に，こうした各解像度における離散表現の間の関係を導くことにしよう．

線形空間  $V_{2^0}$  と線形空間  $V_{2^1}$  のそれぞれの離散表現の間の関係を例に考えよう．線形空間  $V_{2^0}$  は線形空間  $V_{2^1}$  の部分線形空間であるから， $V_{2^0}$  に属する関数  $\phi(t)$  は， $V_{2^1}$  の正規直交基底  $\{2^{1/2}\phi(2t - n) \mid n \in \mathcal{Z}\}$  の線形結合

$$\phi(t) = \sum_{n \in \mathcal{Z}} l_0(-n)2^{1/2}\phi(2t - n) \quad (3.11)$$

で表現できるはずである．ここで， $l_0(-n)$  は，展開係数であり， $n$  の符号は便宜上のものである．展開係数  $l_0(-n)$  は， $\{2^{1/2}\phi(2t - n) \mid n \in \mathcal{Z}\}$  の正規直交性から，

$$l_0(-n) = \int_{-\infty}^{\infty} 2^{1/2}\phi^*(2t - n)\phi(t) dt \quad (3.12)$$

として与えられる．式 (3.11) で規程される関数  $\phi(t)$  の 2 つのスケール間の関係をトゥー・スケール (two scale) 関係という．任意の多重解像度近似に対し，スケーリング関数  $\phi(t)$  が一意に存在することは前に述べたが，トゥー・スケール関係を満たす関数  $\phi(t)$  と多重解像度近似は 1 対 1 に対応することも分かっている

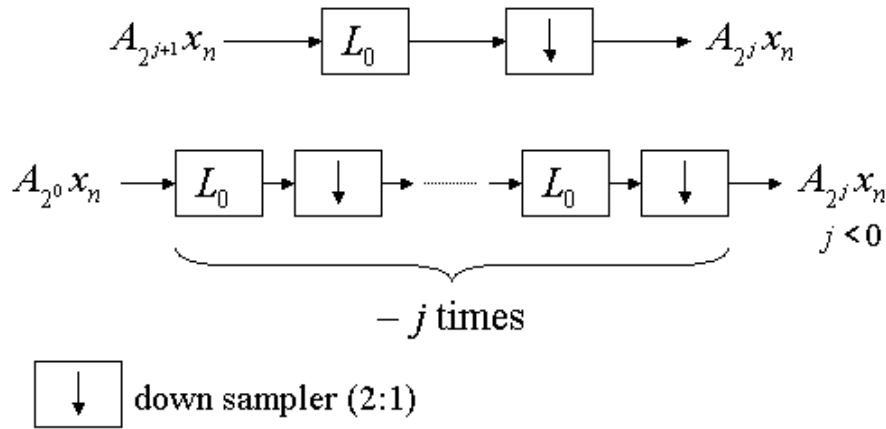


図 3.1: 離散表現  $A_{2^j}x_n$  と  $A_{2^{j+1}}x_n$  の間の関係 (上段), 離散表現  $A_{2^0}x_n$  から  $A_{2^j}x_n, j < 0$  を得る方法 (下段)

る. 式 (3.11) の  $t$  を  $2^j t - k$  とおき, 両辺で  $x(t)$  との内積をとり, 式 (3.10) を適用することにより, 各解像度における離散表現の間の関係式

$$A_{2^j}x_n = \sum_{k \in \mathbb{Z}} l_0(-k) A_{2^{j+1}}x_{2n+k}$$

が得られる. さらに,  $2n + k = m$  と変数変換を行えば,

$$A_{2^j}x_n = \sum_{m \in \mathbb{Z}} l_0(2n - m) A_{2^{j+1}}x_m \quad (3.13)$$

と表現することができる.

式 (3.13) の意味を考えてみよう.  $l_0(n)$  を離散時間 FIR フィルタ  $L_0$  のインパルス応答,  $A_{2^j}x_n, A_{2^{j+1}}x_n$  を共に標本化された離散信号とみなせば,  $A_{2^j}x_n$  は,  $A_{2^{j+1}}x_n$  を FIR フィルタ  $L_0$  に通した後, 標本点を 1 点おきに間引いたものとなっている (図 3.1 の上段の図参考). すなわち, 離散信号  $A_{2^{j+1}}x_n$  に,  $L_0$  によるフィルタリングとダウンサンプリングを反復的に適用することにより, 離散信号  $A_{2^j}x_n, j < 0$  が求められることになる (図 3.1 の下段の図参考).

### 3.4 各解像度空間の直交補空間とその直交展開

解像度  $2^{j+1}$  のベクトル空間  $V_{2^{j+1}}$  から, 解像度  $2^j$  のベクトル空間  $V_{2^j}$  に変換した際に失われる成分を定式化しよう.  $V_{2^j} \subset V_{2^{j+1}}$  であるから,  $V_{2^{j+1}}$  における  $V_{2^j}$  の直交補空間を  $O_{2^j}$  と記述する. すなわち,

$$V_{2^{j+1}} = V_{2^j} \oplus O_{2^j} \quad (3.14)$$

である. また,  $L^2(\mathbb{R})$  から, ベクトル空間  $O_{2^j}$  への正射影子を  $D_{2^j}$ , 関数  $x(t) \in L^2(\mathbb{R})$  を  $O_{2^j}$  へ正射影して得られる関数を  $D_{2^j}[x(t)]$  と表すことにする. この時, 関数系

$$\{ 2^{j/2} \psi(2^j t - n) \mid n \in \mathbb{Z} \} \quad (3.15)$$

が  $O_{2^j}$  の正規直交基底となるような関数  $\psi(t) \in O_1 \subset L^2(\mathbb{R})$  が一意に存在することが知られている. 式 (3.15) は式 (2.9) に等しく, 関数  $\psi(t)$  は直交ウェーブレットそのものである.

関数  $D_{2^j}x(t)$  は、ベクトル空間  $O_{2^j}$  に属するから、そのベクトル空間の正規直交基底  $\{2^{j/2}\psi(2^j t - n) \mid n \in \mathbf{Z}\}$  を用いて、

$$D_{2^j}x(t) = 2^{j/2} \sum_{n \in \mathbf{Z}} D_{2^j}x_n \psi(2^j t - n) \quad (3.16)$$

と表現可能である。ここで、 $D_{2^j}x_n$  は展開係数であり、上式の両辺で  $2^{j/2}\psi(2^j t - m)$  との内積をとり、 $\{2^{j/2}\psi(2^j t - n) \mid n \in \mathbf{Z}\}$  の正規直交性を利用すれば、

$$D_{2^j}x_n = 2^{j/2} \int_{-\infty}^{\infty} \psi^*(2^j t - n)x(t) dt \quad (3.17)$$

として求められる。これは、離散表現  $A_{2^{j+1}}x_n$  と  $A_{2^j}x_n$  の間の差分成分を表す。

関数  $x(t)$  の解像度 1 における離散表現  $A_1x_n$  は、任意の  $J$  に対し最も粗い解像度空間における離散表現  $A_{2^{-J}}x_n$  と各解像度空間の直交補空間の離散表現  $\{D_{2^j}x_n \mid -J \leq j \leq -1\}$  により完全に表現される。ゆえに、ベクトル空間  $V_1$  は、各解像度のベクトル空間の直和

$$V_1 = O_{2^{-1}} \oplus O_{2^{-2}} \oplus \cdots \oplus O_{2^{-J}} \oplus V_{2^{-J}} \quad (3.18)$$

で表現される。こうした  $V_1$  の分解を直交ウェーブレット分解 (orthogonal wavelet decomposition) という。

### 3.5 各直交補空間における離散表現の間の関係

各解像度における離散表現がより高い解像度における離散表現から逐次的に計算されるのと同様に、それらの直交補空間の離散表現もまた逐次的に計算される。 $\psi(t)$  は、 $O_1$  に属するから、 $V_2$  の正規直交基底  $\{2^{1/2}\phi(2t - n) \mid n \in \mathbf{Z}\}$  の線形結合

$$\psi(t) = \sum_{n \in \mathbf{Z}} h_0(-n)2^{1/2}\phi(2t - n) \quad (3.19)$$

として表現できる。ここで、 $h_0(-n)$  は、展開係数であり、 $n$  の符号は便宜上のものである。展開係数  $h_0(-n)$  は、 $\{2^{1/2}\phi(2t - n) \mid n \in \mathbf{Z}\}$  の正規直交性から、

$$h_0(-n) = \int_{-\infty}^{\infty} 2^{1/2}\phi^*(2t - n)\psi(t) dt \quad (3.20)$$

として与えられる。式 (3.19) の  $t$  を  $2^j t - k$  とおき、両辺で  $x(t)$  との内積をとり、式 (3.10),(3.17) を適用することにより、各解像度における離散表現の間の関係式

$$D_{2^j}x_n = \sum_{k \in \mathbf{Z}} h_0(2n - k)A_{2^{j+1}}x_k \quad (3.21)$$

が得られる。

式 (3.21) の意味を考えてみよう。 $h_0(n)$  を離散時間 FIR フィルタ  $H_0$  のインパルス応答、 $D_{2^j}x_n$ 、 $A_{2^{j+1}}x_n$  を離散時間信号とみなせば、 $D_{2^j}x_n$  は、 $A_{2^{j+1}}x_n$  を FIR フィルタ  $H_0$  に通した後、標本点を 1 点おきに間引いたものとなっている (図 3.2 の上段の図参考)。すなわち、観測信号とみなされる離散時間信号  $A_1x_n$  に、 $L_0$  によるフィルタリングと 2:1 のダウンサンプリングを反復的に  $(-j-1)$  回適用した後、 $H_0$  によるフィルタリングと 2:1 のダウンサンプリングを行うことにより、離散信号  $D_{2^j}x_n$ 、 $j < 0$  が求められることになる (図 3.2 の下段の図参考)。



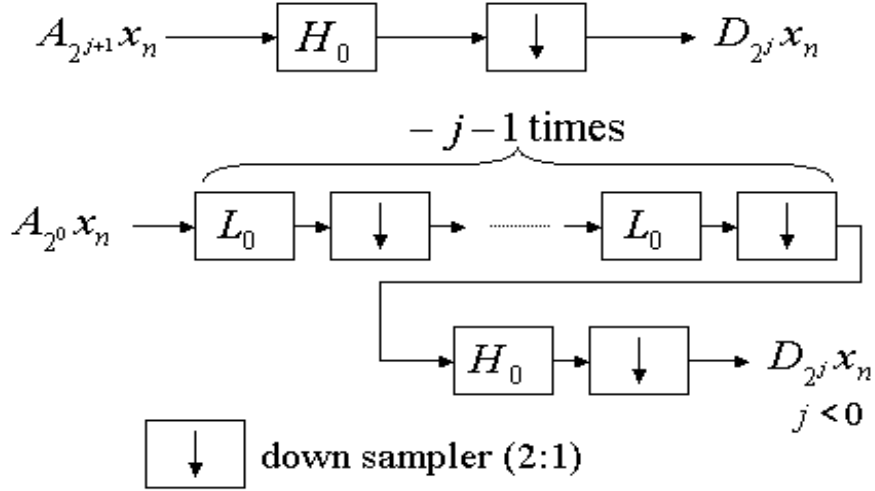


図 3.2: 離散表現  $D_{2^j}x_n$  と  $A_{2^{j+1}}x_n$  の間の関係 (上段), 離散表現  $A_{2^0}x_n$  から  $D_{2^j}x_n, j < 0$  を得る方法 (下段)

### 3.6 直交ウェーブレット表現からの合成

$V_2 = V_1 \oplus O_1$  であるから,  $\phi(2t) \in V_2$  は,  $V_1$  の正規直交基底  $\{\phi(t-n) \mid n \in \mathbf{Z}\}$ , 及び  $O_1$  の正規直交基底  $\{\psi(t-n) \mid n \in \mathbf{Z}\}$  の線形結合の和

$$\phi(2t-k) = 2^{-1/2} \sum_{n \in \mathbf{Z}} l_1(k-2n)\phi(t-n) + 2^{-1/2} \sum_{n \in \mathbf{Z}} h_1(k-2n)\psi(t-n) \quad (3.22)$$

として表現できる. 上式において, 展開係数  $l_1(k-2n), h_1(k-2n)$  を  $l_1(n), h_1(n)$  とすることもできるが, こうした場合, 以後の式の導出において変数変換を行う際,  $l_1(n), h_1(n)$  を非整数  $n$  に対して再定義する必要が生じる. これを避けるために, 便宜上, 展開係数を  $l_1(k-2n), h_1(k-2n)$  とおいた. また, 式 (3.22) の右辺で,  $2^{-1/2}$  を乗じているが, これも便宜上のことである.

さて, 展開係数  $l_1(k-2n), h_1(k-2n)$  は,  $\{\phi(t-n), \psi(t-n) \mid n \in \mathbf{Z}\}$  の正規直交性から

$$l_1(k-2n) = 2^{1/2} \int_{-\infty}^{\infty} \phi^*(t-n)\phi(2t-k) dt \quad (3.23)$$

$$h_1(k-2n) = 2^{1/2} \int_{-\infty}^{\infty} \psi^*(t-n)\phi(2t-k) dt \quad (3.24)$$

と求められる. 式 (3.22) において,  $t$  を  $2^j t$  とおき, 両辺で  $x(t)$  との内積をとり, 式 (3.10), (3.17) を適用すれば, 次式が得られる.

$$A_{2^{j+1}}x_n = \sum_{k \in \mathbf{Z}} l_1(n-2k)A_{2^j}x_k + \sum_{k \in \mathbf{Z}} h_1(n-2k)D_{2^j}x_k \quad (3.25)$$

このように, ある関数  $x(t)$  のある解像度での離散表現と, その解像度空間の直交補空間での離散表現から, 一つ高い解像度での離散表現を得ることを直交ウェーブレット合成 (orthogonal wavelet reconstruction) という.

式 (3.25) の意味を考えてみよう.  $l_1(n), h_1(n)$  をそれぞれ離散時間 FIR フィルタ  $L_1, H_1$  のインパルス応答,  $A_{2^{j+1}}x_n, A_{2^j}x_n, D_{2^j}x_n$  を共に標本化された離散信号とみなせば,  $A_{2^{j+1}}x_n$  は,  $A_{2^j}x_n, D_{2^j}x_n$  の各

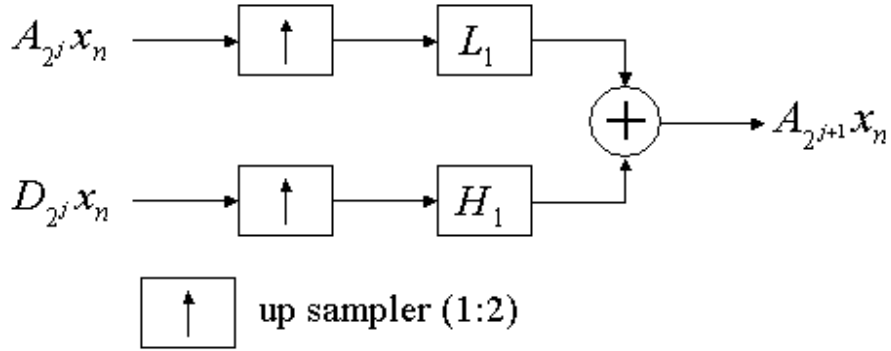


図 3.3: 離散表現  $A_{2^j}x_n$  と  $D_{2^j}x_n$  から離散表現  $A_{2^{j+1}}x_n$  を得る方法

標本点の間に 0 を詰め，それぞれ，FIR フィルタ  $L_1, H_1$  に通した後，和をとったものとなっている（図 3.3 参考）。

各解像度空間，およびその直交補空間での離散表現の関係，すなわち，式 (3.13)，(3.21)，(3.25) は，実は，工学において古くから知られているサブバンドフィルタ (subband filter) におけるサブバンド分解あるいは合成の処理に等しいことが知られている．具体的には，ある解像度の離散表現とその解像度の直交補空間の離散表現を一つ高い解像度の離散表現から計算する処理は，低域通過フィルタ  $L_0$ ，高域通過フィルタ  $H_0$  から構成されるサブバンドフィルタによるサブバンド分解に等しく，ある解像度の離散表現を一つ低い解像度の離散表現とその解像度の直交補空間の離散表現から構成する処理は，低域通過フィルタ  $L_1$ ，高域通過フィルタ  $H_1$  から構成されるサブバンドフィルタによるサブバンド合成に等しい．3 層サブバンドフィルタを用いたサブバンド分解，サブバンド合成の流れを図 3.4 に示す．

ある解像度，例えば  $2^0 = 1$  での  $x(t)$  の離散表現が得られたならば，より低い解像度での離散表現，あるいはその直交補空間の離散表現は，サブバンド分解により求められるから，結局，計算機を使用することの多いデジタル信号処理の分野では，スケーリング関数  $\phi(t)$  やウェーブレット  $\psi(t)$  の波形よりもサブバンドフィルタのフィルタ係数  $l_0(t)$ ， $l_1(t)$ ， $h_0(t)$ ， $h_1(t)$  の方がより重要であるともいえる．

### 3.7 サブバンドフィルタとの関係

式 (3.12)，式 (3.20)，式 (3.23)，式 (3.24) で与えられる FIR フィルタ  $L_0, H_0, L_1, H_1$  のそれぞれのインパルス応答関数

$$l_0(n), h_0(n), l_1(n), h_1(n)$$

の関係を調べよう．式 (3.23)，式 (3.24) で  $n = 0$  とおき，両辺の複素共役をとった式と式 (3.12)，式 (3.20) をそれぞれ比較することにより， $l_0(n)$  と  $l_1(n)$ ，及び  $h_0(n)$  と  $h_1(n)$  の間の関係式

$$l_1(n) = l_0^*(-n) \quad (3.26)$$

$$h_1(n) = h_0^*(-n) \quad (3.27)$$

が得られる．また， $V_1$  と  $O_1$  が互いに直交するから，それらの正規直交基底  $\{\phi(t-n) \mid n \in \mathbf{Z}\}$  と  $\{\psi(t-n) \mid n \in \mathbf{Z}\}$  もまた互いに直交関係にある，すなわち，

$$\int_{-\infty}^{\infty} \phi^*(t-m)\psi(t)dt = 0, \quad \forall m \in \mathbf{Z} \quad (3.28)$$

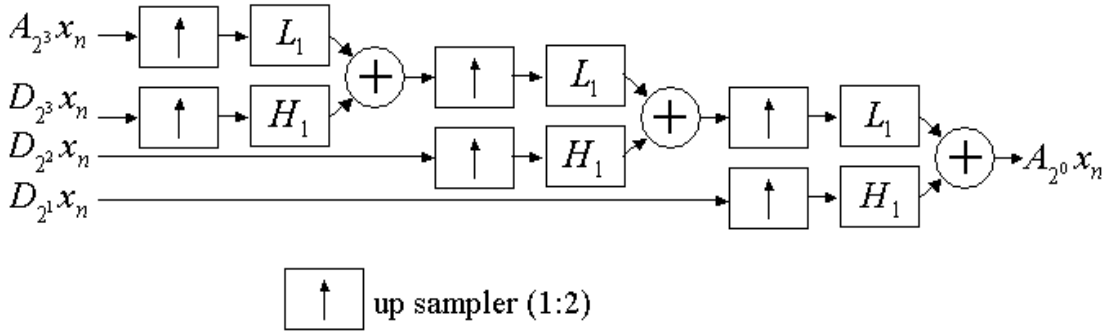
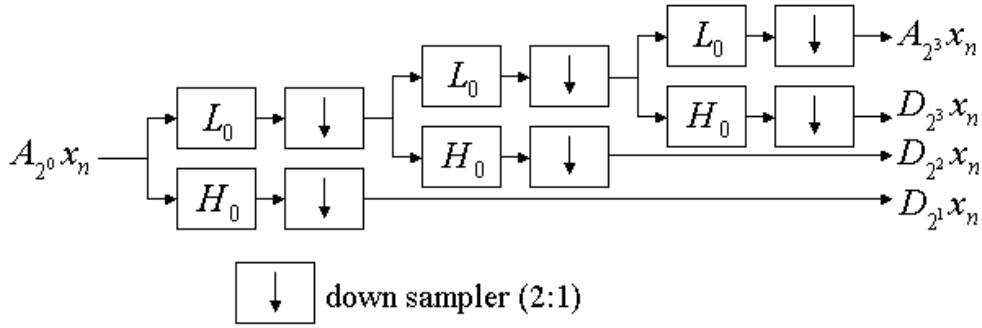


図 3.4: 3 層サブバンドフィルタによるサブバンド分解 (上段) とサブバンド合成 (下段)

が成り立つ．式 (3.11),(3.19) を式 (3.28) に代入すれば，

$$\sum_{n \in \mathbf{Z}} \sum_{k \in \mathbf{Z}} l_0(n) h_0(k) \int_{-\infty}^{\infty} \phi^*(t - 2m + n) \phi(t + k) dt = 0, \quad \forall m \in \mathbf{Z} \quad (3.29)$$

が得られる．式 (3.29) は， $k \neq n - 2m$  ではその積分値が 0 になるため， $k = n - 2m$  の場合についてのみに総和を考えればよい．したがって，条件

$$\sum_{n \in \mathbf{Z}} l_0(n) h_0(n - 2m) = 0, \quad \forall m \in \mathbf{Z} \quad (3.30)$$

が得られる．式 (3.30) を満たす一つの解は，

$$h_0(n) = (-1)^{1-n} l_0(1 - n) \quad (3.31)$$

であり，同様に  $l_1(n)$  と  $h_1(n)$  の間の関係式

$$h_1(n) = (-1)^{1-n} l_1(1 - n) \quad (3.32)$$

も得られる．すなわち，フィルタ  $L_0$  と  $H_0$ ，及び  $L_1$  と  $H_1$  が互いにミラーフィルタの関係にあることが直交ウェーブレットを構成するための一つの十分条件となる．

### 3.8 Daubechies の直交ウェーブレット

一つの多重解像度近似に対し，スケーリング関数，ウェーブレットが一意に対応し，スケーリング関数とウェーブレットは，信号  $x(t)$  の各解像度での離散表現の間の関係を与えるサブバンドフィルタを規定す

る．これらは、互いに深く関係し、いずれかを定めることにより、残りはほぼ自動的に決定される．Meyer は、前に述べたように、滑らかさの立場からスケーリング関数、ウェーブレットを設計した．しかし、デジタル信号を扱う限り、解像度の変換はすべてサブバンドフィルタにより実現されるから、スケーリング関数  $\phi(t)$  やウェーブレット  $\psi(t)$  の波形よりもサブバンドフィルタ  $L_0, L_1, H_0, H_1$  の方がより重要である．そこで、Daubechies は、こうしたサブバンドフィルタの立場から直交ウェーブレットの設計を試みた．以下、Daubechies による直交ウェーブレットの導出過程の概要を紹介しよう．

デジタルフィルタリングを実現する際の計算コストの立場からは、サブバンドフィルタのフィルタ係数  $l_0(t), l_1(t), h_0(t), h_1(t)$  は、非ゼロ要素の個数が少ない方が望ましい．そこで、Daubechies は、まず、すべての自然数  $N$  に対し、次の条件を満たす  $\{l_0(n) \mid n \in \mathbb{Z}\}$  が存在することを示した．ここで、 $N$  は、Daubechies のパラメータと呼ばれている．

$$\left. \begin{aligned} l_0(n) &= 0, \quad n < 0, \quad n > 2N \\ \sum_{k=0}^{2N-1} l_0(n) &= 2^{1/2} \\ \sum_{k=0}^{2N-1} l_0^2(n) &= 1 \\ \sum_{k=0}^{2N-1} l_0(n)l_0(n-2m) &= 0, \quad m \neq 0 \end{aligned} \right\} \quad (3.33)$$

式 (3.33) を満足する  $l_0(n)$  のうち、Daubechies が設計した  $N = 1, \dots, 10$  の係数を表 3.1 に示す．さらに、Daubechies は、式 (3.33) を満足する  $l_0(n), N \geq 2$  に対し、式 (3.11) は連続かつコンパクト・サポートで  $\int_{-\infty}^{\infty} \phi(x) dx = 1$  となる解  $\phi(t)$  を唯一持ち、こうした  $\phi(x)$  から得られるウェーブレット  $\psi(x)$  は  $N-1$  次モーメントまでが 0 になることを示した．こうして設計されるウェーブレットは、Daubechies の直交ウェーブレットと呼ばれている．ウェーブレットを用いたサブバンド分解、サブバンド合成は、それぞれウェーブレット変換、逆ウェーブレット変換と呼ばれている．

演習 5：音声などの適当な信号に対し、適当な  $N$  の Daubechies の直交ウェーブレットにより、ウェーブレット変換を行ってみよう．また、ウェーブレット変換した結果を逆ウェーブレット変換することにより、元の信号が復元されることを確認しよう．

解答：

表 3.1: Daubechies のフィルタ  $L_0$  のインパルス応答  $l_0(n)$

$N$	$n$	$l_0(n)$	$N$	$n$	$l_0(n)$
1	0	$2^{-1/2}$	7	13	0.000353713800
1	1	$2^{-1/2}$	8	0	0.054415842243
2	0	0.482962913145	8	1	0.312871590914
2	1	0.836516303738	8	2	0.675630736297
2	2	0.224143868042	8	3	0.585354683654
2	3	-0.129409522551	8	4	-0.015829105256
3	0	0.332670552950	8	5	-0.284015542962
3	1	0.806891509311	8	6	0.000472484574
3	2	0.459877502118	8	7	0.128747426620
3	3	-0.135011020010	8	8	-0.017369301002
3	4	-0.085441273882	8	9	-0.044088253931
3	5	0.035226291882	8	10	0.013981027917
4	0	0.230377813309	8	11	0.008746094047
4	1	0.714846570553	8	12	-0.004870352993
4	2	0.630880767930	8	13	-0.000391740373
4	3	-0.027983769417	8	14	0.000675449406
4	4	-0.187034811719	8	15	-0.000117476784
4	5	0.030841381836	9	0	0.038077947364
4	6	0.032883011667	9	1	0.243834674613
4	7	-0.010597401785	9	2	0.604823123690
5	0	0.160102397974	9	3	0.657288078051
5	1	0.603829269797	9	4	0.133197385825
5	2	0.724308528438	9	5	-0.293273783279
5	3	0.138428145901	9	6	-0.096840783223
5	4	-0.242294887066	9	7	0.0148540749338
5	5	-0.032244869585	9	8	0.030725681479
5	6	0.077571493840	9	9	-0.067632829061
5	7	-0.006241490213	9	10	0.000250947115
5	8	-0.012580751999	9	11	0.022361662124
5	9	0.003335725285	9	12	-0.004723204758
6	0	0.111540743350	9	13	-0.004281503682
6	1	0.494623890398	9	14	0.001847646883
6	2	0.751133908021	9	15	0.000230385764
6	3	0.315250351709	9	16	-0.000251963189
6	4	-0.226264693965	9	17	0.000039347320
6	5	-0.129766867567	10	0	0.026670057901
6	6	0.097501605587	10	1	0.188176800078
6	7	0.027522865530	10	2	0.527201188932
6	8	-0.031582039318	10	3	0.688459039454
6	9	0.000553842201	10	4	0.281172343661
6	10	0.004777257511	10	5	-0.249846424327
6	11	-0.001077301085	10	6	-0.195946274377
7	0	0.077852054085	10	7	0.127369340336
7	1	0.396539319482	10	8	0.093057364604
7	2	0.729132090846	10	9	-0.071394147166
7	3	0.469782287405	10	10	-0.029457536822
7	4	-0.143906003929	10	11	0.033212674059
7	5	-0.224036184994	10	12	0.003606553567
7	6	0.071309219267	10	13	-0.010733175483
7	7	0.080612609151	10	14	0.001395351747
7	8	-0.038029936935	10	15	0.001992405295
7	9	-0.016574541631	10	16	-0.000685856695
7	10	0.012550998556	10	17	-0.000116466855
7	11	0.000429577973	10	18	0.000093588670
7	12	-0.001801640704	10	19	-0.000013264203

## 第4章 ケプストラム解析とヒルベルト変換

### 4.1 パワーケプストラム

ある信号  $x(t)$  が複数の信号の和

$$x(t) = \sum_{n=1}^N x_n(t) \quad (4.1)$$

で構成されており、各信号  $x_n(t)$  の周波数帯域が十分離れていれば、信号をフーリエ変換することにより、それぞれ独立したスペクトルとして観測することが可能である。したがって、こうした信号を分離するためには、それぞれのスペクトルを抽出する帯域通過フィルタを設計し、信号  $x(t)$  をフィルタリングすればよい。しかし、信号  $x(t)$  が複数の信号の畳み込み積分

$$x(t) = x_1(t) * x_2(t) * \cdots * x_N(t)$$

で表現される場合には、フーリエ変換しても、それは、それぞれの信号のフーリエ変換の積

$$X(\omega) = \prod_{n=1}^N X_n(\omega) \quad (4.2)$$

で表されるため、通常のフィルタリングでは分離することができない。ただし、 $X(\omega) = \mathcal{F}[x(t)]$ 、 $X_n(\omega) = \mathcal{F}[x_n(t)]$  とする。しかし、式 (4.2) は対数をとれば積が和になることから、このことを利用すれば、フィルタリングによる信号の分離が可能になると思われる。こうした解析を、ケプストラム解析 (cepstrum analysis) という。ケプストラムとは、スペクトル (spectrum) の spec を逆順にした造語である。

フーリエスペクトル  $X(\omega)$  は一般に複素数であることから、ケプストラム解析を行う、すなわち各信号  $x_n(t)$  の波形を求めるためには、位相情報を保持するために、その複素対数をとる必要がある。フーリエスペクトル  $X(\omega)$  の複素対数の逆フーリエ変換は、複素ケプストラム (Complex Cepstrum) と呼ばれている。しかし、 $x(t)$  のパワースペクトル  $|X(\omega)|^2$  から各信号  $x_n(t)$  のパワースペクトル  $|X_n(\omega)|^2$  を分離することが目的である場合には、パワースペクトル  $|X(\omega)|^2$  の実対数の逆フーリエ変換として定義されるパワーケプストラム (Power Cepstrum) を用いるのが便利である。まず、パワーケプストラムから説明してゆこう。

式 (4.2) の絶対値の2乗は、

$$|X(\omega)|^2 = \prod_{n=1}^N |X_n(\omega)|^2 \quad (4.3)$$

であり、両辺の実対数を取り、2で割ると、

$$\log |X(\omega)| = \sum_{n=1}^N \log |X_n(\omega)| \quad (4.4)$$

となり、さらに逆フーリエ変換することにより、パワーケプストラムは、

$$x_{pc}(t) \equiv \mathcal{F}^{-1}[\log |X(\omega)|] = \sum_{n=1}^N \mathcal{F}^{-1}[\log |X_n(\omega)|] \quad (4.5)$$

となる．パワーケプストラムは，パワースペクトルの対数を逆フーリエしたものであるから，時間軸  $t$  での表現である．しかし，もとの時間軸と区別するために，周波数 (frequency) の  $fre$  と  $que$  を入れ換えて，ケフレンシ (quefrequency) と呼ぶ．

さて，ケフレンシ上で  $x_{pc}(t)$  の帯域が分離されていれば，ケフレンシ上のフィルタリングを行い，再びフーリエ変換することにより，それぞれのパワースペクトル  $|X_n(\omega)|^2$  を分離して観測することが可能になる．ケフレンシ上のフィルタ (filter) を， $fil$  を逆順にしてリフタ (lifter) と呼び，その処理をリフタリングと呼ぶ．周波数軸では，高域 (high)，低域 (low) と呼ぶのに対応して，ケフレンシ軸では，時間に相当する軸であることから，それぞれ，長 (long)，短 (short) と呼ぶ．

パワーケプストラムを利用したパワースペクトル分離の一例を示そう．信号  $x(t)$  が，ある信号  $y(t)$  とその遅れ時間  $s$  のエコー  $ay(t)$  の和

$$x(t) = y(t) + ay(t - s), \quad |a| < 1 \quad (4.6)$$

で表現されている場合を考える．これは， $y(t)$  がインパルス応答特性  $h(\tau) = \delta(\tau) + a\delta(\tau - s)$  を持つシステムに入力され，その出力が  $x(t) = h(t) * y(t)$  であると考えられることができる．我々は， $y(t)$  を  $x(t)$  から分離して観測したいのであり，これはしばしば発生する問題である． $y(t)$  の非ゼロである範囲が  $s$  に比べ長い場合には，時間軸で  $x(t)$  から  $y(t)$  を分離することは不可能である．また，フーリエ変換したとしても， $y(t)$  と  $y(t - s)$  は，同じパワースペクトルを持つため，周波数軸でも分離は不可能である．こうした  $y(t)$  のパワースペクトルを求めたい場合に，パワーケプストラムは非常に便利である．

さて， $x(t)$ ， $h(t)$ ， $y(t)$  のそれぞれのフーリエ変換を  $X(\omega)$ ， $H(\omega)$ ， $Y(\omega)$  とすれば，式 (4.6) のフーリエ変換は，時間軸をシフトした信号のフーリエ変換の性質を利用すれば，

$$X(\omega) = Y(\omega)H(\omega) = Y(\omega)(1 + ae^{-i\omega s}) \quad (4.7)$$

となり，そのパワースペクトルは，

$$\begin{aligned} |X(\omega)|^2 &= |Y(\omega)|^2 |1 + ae^{-i\omega s}|^2 \\ &= |Y(\omega)|^2 (1 + ae^{-i\omega s})(1 + ae^{i\omega s}) \\ &= |Y(\omega)|^2 (1 + a^2 + 2a \cos(\omega s)) \end{aligned} \quad (4.8)$$

となる．上式の実対数は

$$\log |X(\omega)|^2 = \log |Y(\omega)|^2 + \log(1 + a^2) + \log\left(1 + \frac{2a}{1 + a^2} \cos(\omega s)\right) \quad (4.9)$$

となる．ここで， $|a| < 1$  であることから， $\left|\frac{2a}{1 + a^2}\right| < 1$  である．そこで  $\log(1+x)$  の Taylor 展開が  $-1 < x \leq 1$  で  $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$  となることを利用すれば，上式は，

$$\log |X(\omega)|^2 = \log |Y(\omega)|^2 + \log(1 + a^2) + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \left(\frac{2a}{1 + a^2}\right)^n \cos^n(\omega s) \quad (4.10)$$

と表される．上式，右辺第 2 項は，逆フーリエ変換すれば，ケフレンシ  $t = 0$  で局在し，右辺第 3 項は，さらに  $\cos(\omega s)$ ， $\cos(2\omega s)$ ， $\cos(3\omega s)$ ， $\dots$  に展開できるので，逆フーリエ変換すれば，エコー時間  $s$  の整数倍のケフレンシに局在する．すなわち，パワーケプストラム  $x_{pc}(t)$  は，

$$x_{pc}(t) = y_{pc}(t) + A_0\delta(t) + \sum_{n=1}^{\infty} A_n(\delta(t - ns) + \delta(t + ns)) \quad (4.11)$$

と表されることになる． $A_n$ ,  $n \in \mathcal{Z}$  は，エコーにより出現したピークの係数である．そこで，このピークをリフタにより除去し，フーリエ変換することより， $y(t)$  のパワースペクトル  $|Y(\omega)|^2$  を  $|X(\omega)|^2$  から分離することが可能になる．また，ピークの間隔を測ることにより，エコー時間  $s$  を推定できる．

ここまでは，連続時間で定義される信号  $x(t)$  を考えてきたが，標本化された離散時間信号に対するケプストラム解析はどのようになるのだろうか．離散フーリエ変換は，フーリエ変換の基本的な性質をすべて有しているので，原理的には，離散時間系列に対しては離散フーリエ変換を用いてケプストラム解析を行うことができる．しかし，信号を離散化する際のサンプリング定理に起因して新たな問題が発生する．まず，最初の問題は，信号  $x(t)$  のパワースペクトルに対し，非線形変換である対数を取った場合にも，周波数軸のサンプリング定理を満たしている必要があることである．周波数軸のサンプリング間隔は，元々の信号  $x(t)$  の時間軸での幅に逆比例するので，ケプストラム解析では，時間軸のサンプリング定理を満たすように信号  $x(t)$  を標本化するだけでなく，あらかじめ十分な長さで信号  $x(t)$  を切り出しておかなければならない．十分な長さで信号  $x(t)$  を切り出すことができない場合には， $x(t)$  の後ろに適当な長さの 0 を詰めてもよい（ゼロ詰め）．

実際，エコーを含む信号のパワーケプストラムは，式 (4.11) の右辺第 3 項からも明らかのように，いくらでも長い (long) ケフレンシ成分を含むことになる． $|a|$  が 1 に近いほど， $2a/(1+a^2)$  の絶対値が 1 に近くなることから，第 3 項の級数の収束が遅くなる．したがって，元々の信号  $x(t)$  が，この級数が十分に 0 に収束する長さを有していない場合には，ケフレンシ軸でのエイリアジグが生じることになり，リフタによるエコー成分により発生したピークを除去することが困難になる．

演習 6：信号

$$y(t) = \cos\left(2\pi\frac{2}{21}t\right)e^{-\frac{t}{8}}, \quad t = 0, 1, \dots, 23$$

が，線形インパルス応答

$$h(t) = \begin{cases} 1, & t = 0 \\ 0.7, & t = 10 \\ 0, & \text{others} \end{cases}$$

を持つエコーを起こすシステムに入力され， $x(t) = y(t) * h(t)$  が観測されるものとする． $x(t)$  が与えられたものとして，パワーケプストラムを利用してエコーの成分を除去し，元々の信号  $y(t)$  のパワースペクトルを推定してみよう．もちろん，推定の際には， $h(t)$  が分からないものとする．また， $x(t)$  の系列長が  $N = 64, 128, 256$  となるようにゼロ詰めを行い，系列長を長くとることにより，ケフレンシ軸でのエイリアジグの影響が少なくなり，信号  $y(t)$  のパワースペクトルの復元精度が改善することを確認してみよう．

解答：図 4.1(a) は，エコーがかかる前の信号，すなわちエコーシステムへの入力  $y(t)$  であり，そのパワースペクトルを同図 (b) に示す．(c) は，エコーがかかった信号  $x(t)$  であり，我々の目的は，(c) が与えられて，これから (b) を復元することである．エコーがかかった観測信号  $x(t)$  に  $N = 128$  点までゼロをつめて，求められたパワースペクトルを (d) に示す．(d) に示したパワースペクトルの対数を取り，逆フーリエ変換して得られるパワーケプストラムを (e) に青色の実線で示す．これを見ると，10 標本点間隔で，インパルス状のピークがあることが分かる．これが 10 標本点分遅れるエコーの影響である．そこで，このピークをその前後 2 点の平均値で置き換えたものを (e) に赤色の実線で示す．これをフーリエ変換し，指数をとったものを (f) に示す．これが，原信号  $y(t)$  のパワースペクトルの復元値である．(b) に示した原信号の真のパワースペクトルがよく復元されていることが分かる．また，エコーがかかった観測信号  $x(t)$  に対し，系列長が  $N = 64, 128, 256$  点になるようにゼロ詰めしてケプストラム解析を行った際のそれぞれのパワーケプストラムと復元された原信号のパワースペクトルを図 4.2 に示す．



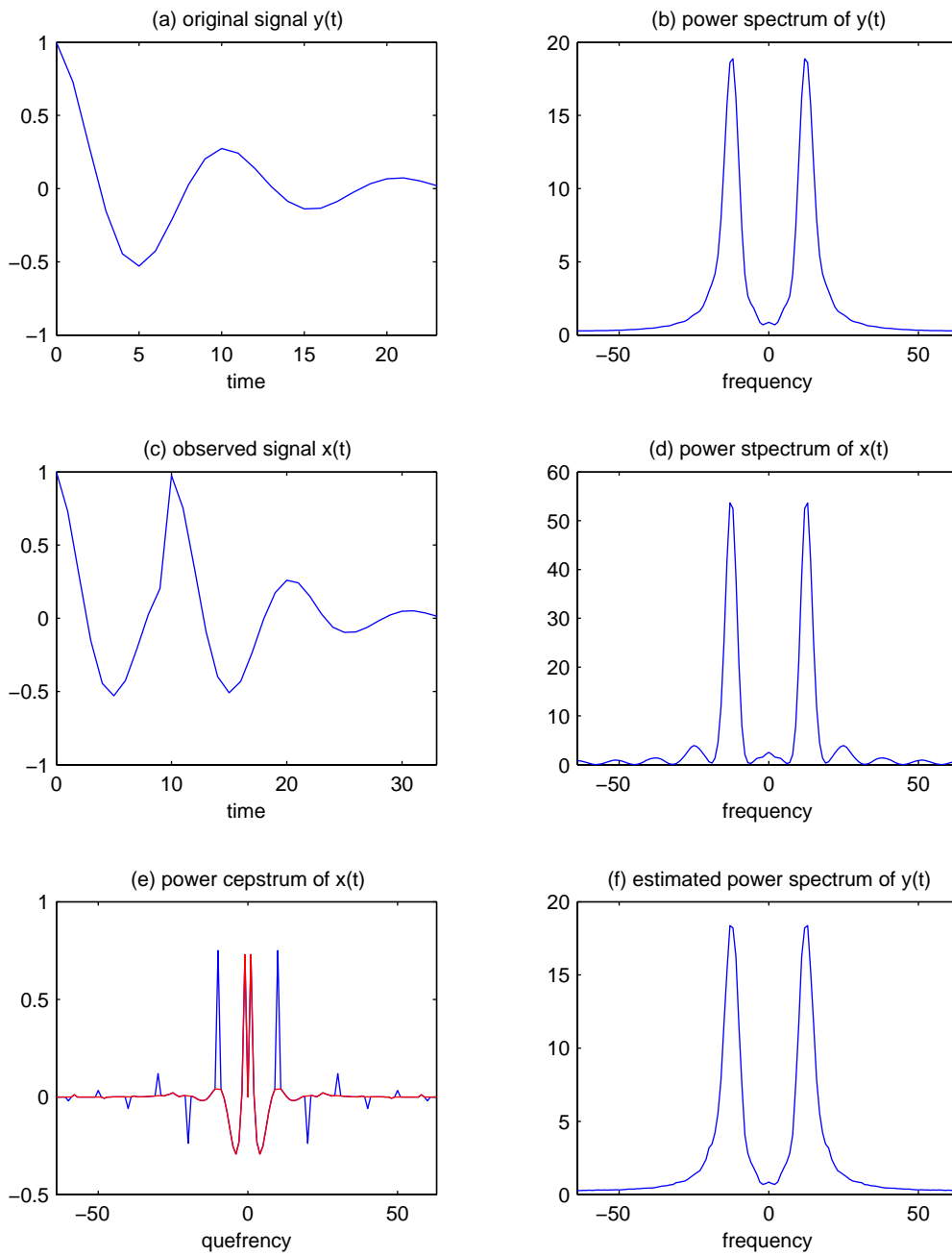


図 4.1: 演習 6 の結果 : パワーケプストラムを利用してエコーを除去した信号のパワースペクトルの推定

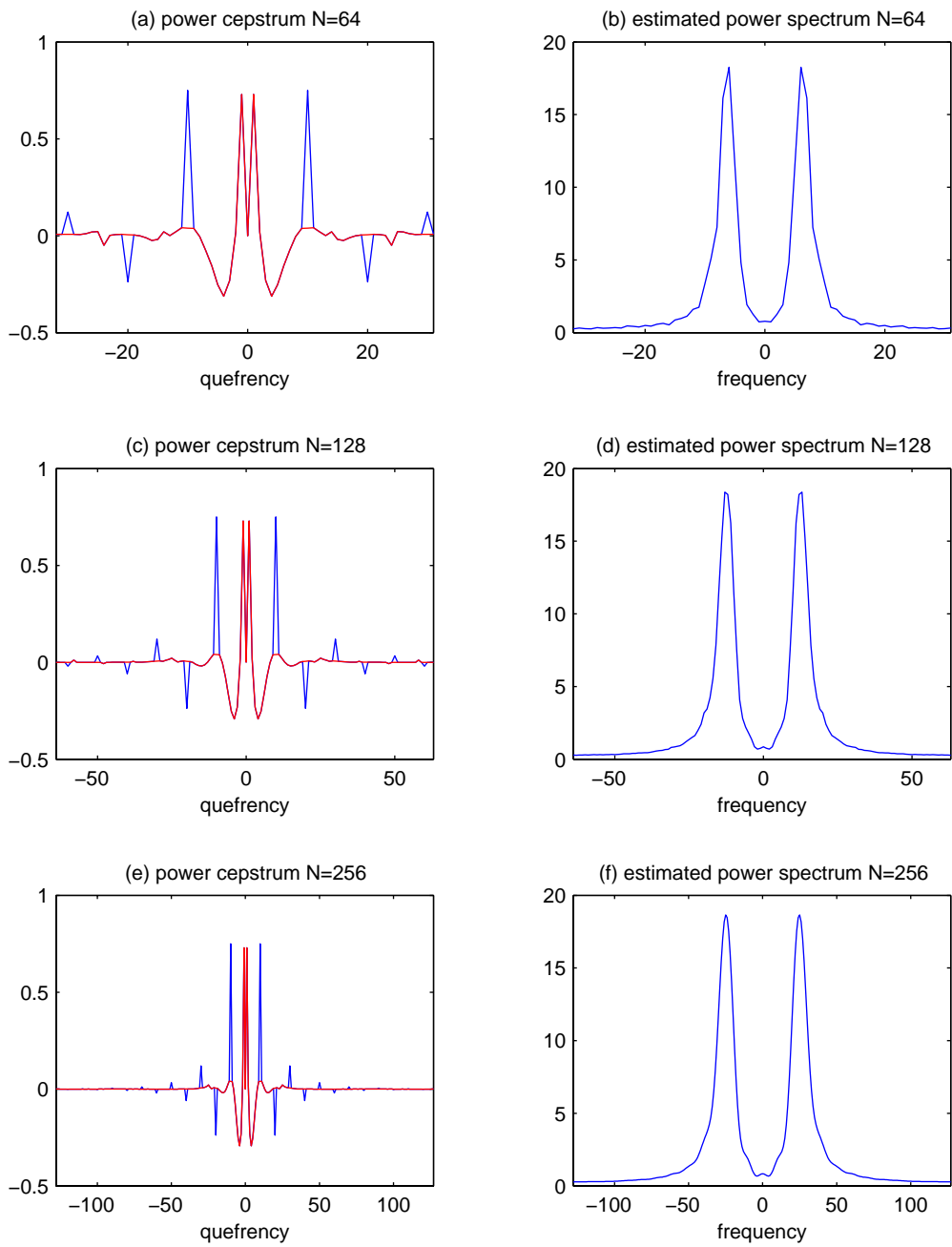


図 4.2: ケフレンシ上のエイリアジングの影響

## 4.2 複素ケプストラム

パワーケプストラムをリフタ処理後，フーリエ変換し，指数をとることにより，分離された波形のパワースペクトルが得られる．しかし，位相情報が失われているため，分離された波形そのものを見ることはできない．したがって，分離した波形を見るためには，位相情報を含めてケフレンシ上に展開しなければならない．

信号  $x(t)$  のフーリエ変換は，一般に複素数であるから，その対数は複素対数となる．ここで，複素数  $x$  の複素対数を

$$\log(x) \equiv \log|x| + i \arg(x) \quad (4.12)$$

と定義すれば， $x(t)$  の複素ケプストラムは

$$x_{cc}(t) \equiv \mathcal{F}^{-1}[\log \mathcal{F}[x(t)]] = \mathcal{F}^{-1}[\log|X(\omega)| + i \arg(X(\omega))] \quad (4.13)$$

と定義される．計算機を用いて複素ケプストラムを求める場合，注意しなければならない点がある．計算機では，位相  $\arg(X(\omega))$  を求めると，主値  $(-\pi, \pi)$  で値を返すようになっている．これは，真の位相  $\arg(X)$  に対して，主値に直す非線形変換を行うことに相当し，こうした非線形変換を行うと，線形加算性が失われるので，正しい複素ケプストラムは求められない．そこで，計算機を用いて複素ケプストラムを求める場合，主値として求められた位相を真の位相に変換する，位相アンラップ（位相戻し）を行う必要がある．

パワーケプストラム  $x_{pc}(t)$  と複素ケプストラム  $x_{cc}(t)$  の関係を調べてみよう． $x(t)$  のパワーケプストラム  $x_{pc}(t)$  は，

$$\begin{aligned} x_{pc}(t) &= \mathcal{F}^{-1}[\log|X(\omega)|^2] \\ &= \mathcal{F}^{-1}[\log(X(\omega)X^*(\omega))] \\ &= \mathcal{F}^{-1}[\log X(\omega) + \log X^*(\omega)] \end{aligned} \quad (4.14)$$

と表される．さて，ここで，フーリエ変換の性質

$$X^*(\omega) = X(-\omega)$$

を利用すれば， $x(t)$  のパワーケプストラム  $x_{pc}(t)$  は，

$$x_{pc}(t) = x_{cc}(t) + x_{cc}(-t) \quad (4.15)$$

となる．つまり，パワーケプストラムは，複素ケプストラムを用いて表すことができる．また，上式より，パワーケプストラムは，複素ケプストラムの奇関数成分が失われ，偶関数成分が2倍になり，結果として偶関数になることが分かる．

先に示したエコーを含む信号  $x(t) = y(t) + ay(t-s)$ ， $|a| < 1$  の複素ケプストラムを求めてみよう． $x(t)$  のフーリエ変換の複素対数をとれば，

$$\log X(\omega) = \log Y(\omega) + \log(1 + ae^{-i\omega s}) \quad (4.16)$$

となるが，右辺第2項を Taylor 展開して

$$\log X(\omega) = \log Y(\omega) + \sum_{n=1}^{\infty} \frac{a^n (-1)^{n+1}}{n!} e^{-in\omega s} \quad (4.17)$$

となる．したがって， $x(t)$  の複素ケプストラム  $x_{cc}(t)$  は，上式の逆フーリエ変換により

$$\begin{aligned} x_{cc}(t) &= \mathcal{F}^{-1}[\log X(\omega)] \\ &= y_{cc}(t) + \sum_{n=1}^{\infty} \frac{a^n (-1)^{n+1}}{n!} \delta(t - ns) \end{aligned} \quad (4.18)$$

となる．すなわち，エコー時間  $s$  を持つ信号の和の複素ケプストラムは， $s$  の整数倍の時刻でインパルスを生じる．したがって，このインパルスをリフタにより除去し，フーリエ変換し，指数をとることより， $y(t)$  のスペクトル  $Y(\omega)$  を  $X(\omega)$  から分離できる．さらに， $Y(\omega)$  は位相情報を持っているので，これを逆フーリエ変換すれば， $y(t)$  を復元することが可能になる．

演習 7： 信号

$$y(t) = \cos\left(2\pi \frac{2}{21} t\right) e^{-\frac{t}{8}}, \quad t = 0, 1, \dots, 23$$

が，線形インパルス応答

$$h(t) = \begin{cases} 1, & t = 0 \\ 0.7, & t = 10 \\ 0, & \text{others} \end{cases}$$

を持つエコーを起こすシステムに入力され， $x(t) = y(t) * h(t)$  が観測されるものとする． $x(t)$  が与えられたものとして，複素ケプストラムを利用してエコーの成分を除去し，元々の信号  $y(t)$  を推定してみよう．もちろん，推定の際には， $h(t)$  が分からないものとする．ただし， $x(t)$  の系列長が  $N = 64$  となるようにゼロ詰めを行うものとする．

解答： 図 4.3(a) は，エコーがかかる前の信号，すなわちエコーシステムへの入力  $y(t)$  であり，そのフーリエスペクトルを同図 (b) に示す．(c) は，エコーがかかった信号  $x(t)$  であり，我々の目的は，(c) が与えられて，これから (a) を復元することである．エコーがかかった観測信号  $x(t)$  に  $N = 64$  点までゼロをつめて，求められたフーリエスペクトルを (d) に示す．(d) に示したフーリエスペクトルの複素対数を取り，逆フーリエ変換して得られる複素ケプストラムを (e) に青色の実線で示す．これを見ると，10 標本点間隔で，インパルス状のピークがあることが分かる．これが 10 標本点分遅れるエコーの影響である．そこで，このピークをその前後 2 点の平均値で置き換えたものを (e) に赤色の実線で示す．これをフーリエ変換し，指数をとったものを (f) に示す．さらに，逆フーリエ変換して得られる原信号の復元値を (g) に示す．(a) がよく復元されていることがわかる．

### 4.3 ケプストラム解析と最小(最大)位相性，線形位相性

式 (4.9)，あるいは式 (4.16) において， $|a| > 1$  の場合はどうなるのだろうか．式 (4.16) を変形し，

$$\log X(\omega) = \log Y(\omega) + \log\left(ae^{-i\omega s}(1 + a^{-1}e^{i\omega s})\right) \quad (4.19)$$

とすれば， $|a^{-1}| < 1$  になるので，Taylor 展開を行うことが可能になる．Taylor 展開を行えば，

$$\log X(\omega) = \log Y(\omega) + \log a - i\omega s + \sum_{n=1}^{\infty} \frac{a^{-n}(-1)^{n+1}}{n!} e^{in\omega s} \quad (4.20)$$

となり，逆フーリエ変換することにより，複素ケプストラムは，

$$x_{cc}(t) = y_{cc}(t) + \delta(t) \log a + \mathcal{F}^{-1}[-i\omega s] + \sum_{n=1}^{\infty} \frac{a^{-n}(-1)^{n+1}}{n!} \delta(t + ns) \quad (4.21)$$

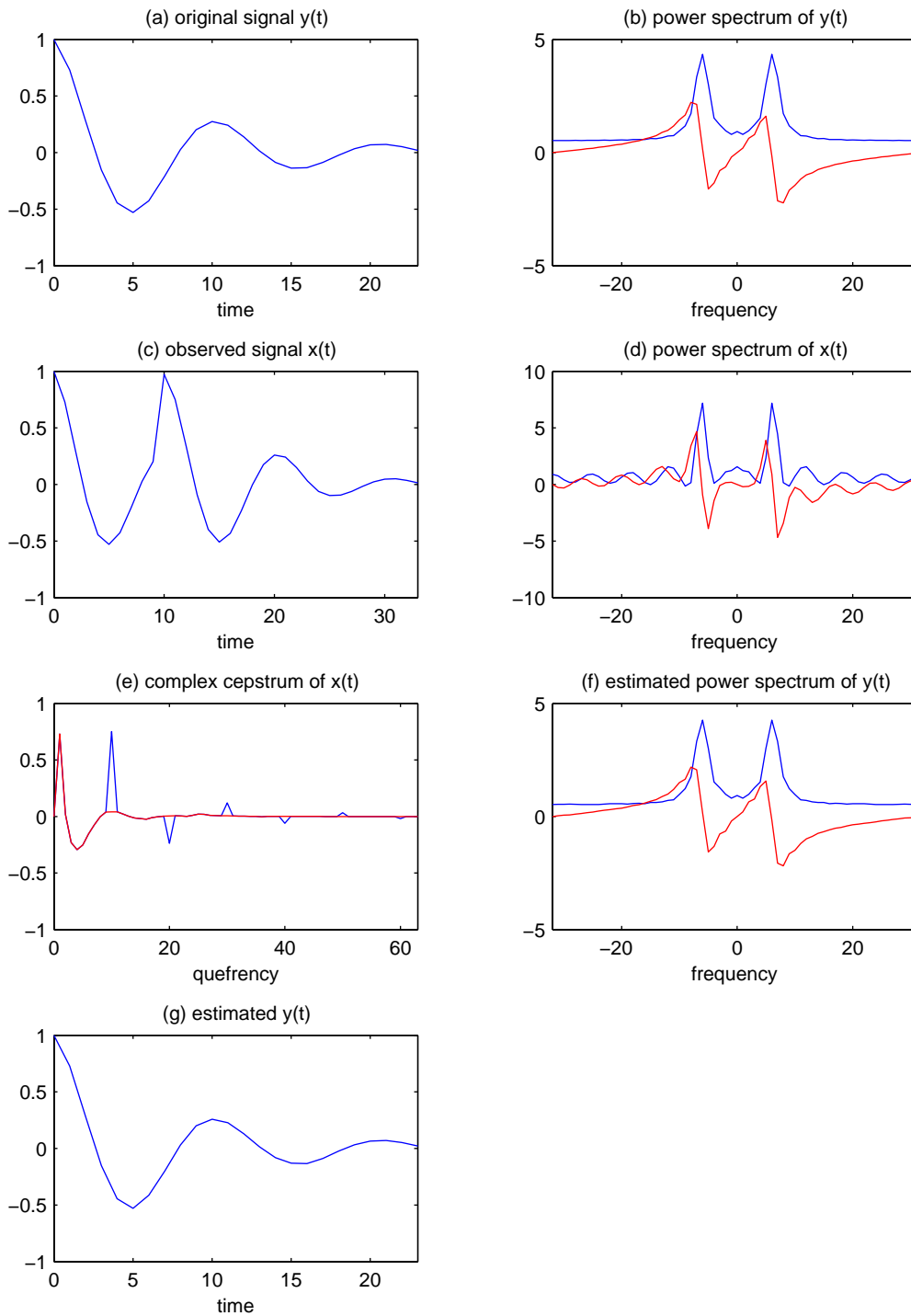


図 4.3: 演習 7 の結果 : 複素ケプストラムを利用してエコーを除去した信号の推定

となる．上式と式 (4.18) を比較すれば， $|a| > 1$  となったことにより，上式右辺第 2,3 項が増えている．第 3 項は，線形位相項であることから，この項の影響を考えるために，系列の最小・最大位相性を思い出し， $x(t)$  の極，零点の立場から，解析してみよう．

離散時間信号  $x(t)$  の  $z$  変換  $X(z)$  を極，零点を用いて

$$X(z) = \frac{Az^r \prod_{i=1}^{n_a} (1 - a_i z^{-1}) \prod_{i=1}^{n_b} (1 - b_i z)}{\prod_{i=1}^{n_c} (1 - c_i z^{-1}) \prod_{i=1}^{n_d} (1 - d_i z)} \quad (4.22)$$

と表現する．ただし，各係数は

$$|a_i|, |b_i|, |c_i|, |d_i| < 1 \quad \forall i$$

とする．したがって， $a_i, c_i$  は，それぞれ  $z$  平面の単位円内にある零点，極を表し，また， $b_i, d_i$  は，それぞれ  $z$  平面の単位円外にある零点，極を表す．また， $A$  は正の定数とし， $z^r$  は， $r < 0$  ならば純粋な遅れ， $r > 0$  ならば純粋な進みを表す線形位相成分となる．さて，式 (4.22) の対数をとれば

$$\begin{aligned} \log X(z) &= \log A + r \log z + \sum_{i=1}^{n_a} \log(1 - a_i z^{-1}) + \sum_{i=1}^{n_b} \log(1 - b_i z) \\ &\quad - \sum_{i=1}^{n_c} \log(1 - c_i z^{-1}) - \sum_{i=1}^{n_d} \log(1 - d_i z) \end{aligned} \quad (4.23)$$

となる．上式は， $\log(1+x)$  の Taylor 展開が，

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \quad \text{for } |x| < 1 \quad (4.24)$$

となることを利用すれば，

$$\begin{aligned} \log X(z) &= \log A + r \log z - \sum_{i=1}^{n_a} \sum_{n=1}^{\infty} \frac{a_i^n}{n} z^{-n} - \sum_{i=1}^{n_b} \sum_{n=1}^{\infty} \frac{b_i^n}{n} z^n \\ &\quad + \sum_{i=1}^{n_c} \sum_{n=1}^{\infty} \frac{c_i^n}{n} z^{-n} + \sum_{i=1}^{n_d} \sum_{n=1}^{\infty} \frac{d_i^n}{n} z^n \end{aligned} \quad (4.25)$$

となる．ここで，

$$\begin{aligned} \sum_{n=1}^{\infty} (\cdot) &= \sum_{n=-\infty}^{\infty} (\cdot) u(n-1) \\ \text{where, } u(n) &= \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \end{aligned} \quad (4.26)$$

を利用すれば，

$$\begin{aligned} \log X(z) &= \log A + r \log z - \sum_{i=1}^{n_a} \sum_{n=-\infty}^{\infty} \frac{a_i^n}{n} z^{-n} u(n-1) - \sum_{i=1}^{n_b} \sum_{n=-\infty}^{\infty} \frac{b_i^n}{n} z^n u(n-1) \\ &\quad + \sum_{i=1}^{n_c} \sum_{n=-\infty}^{\infty} \frac{c_i^n}{n} z^{-n} u(n-1) + \sum_{i=1}^{n_d} \sum_{n=-\infty}^{\infty} \frac{d_i^n}{n} z^n u(n-1) \end{aligned} \quad (4.27)$$

となり，上式右辺第 4,6 項において  $n$  を  $-n$  に置き換えれば，

$$\begin{aligned} \log X(z) = & \log A + r \log z - \sum_{n=-\infty}^{\infty} \sum_{i=1}^{n_a} \frac{a_i^n}{n} u(n-1) z^{-n} + \sum_{n=-\infty}^{\infty} \sum_{i=1}^{n_b} \frac{b_i^{-n}}{n} u(-n-1) z^{-n} \\ & + \sum_{n=-\infty}^{\infty} \sum_{i=1}^{n_c} \frac{c_i^n}{n} u(n-1) z^{-n} - \sum_{n=-\infty}^{\infty} \sum_{i=1}^{n_d} \frac{d_i^{-n}}{n} u(-n-1) z^{-n} \end{aligned} \quad (4.28)$$

となる．ここで，上式を逆  $z$  変換することにより， $z$  変換が式 (4.22) で与えられる離散系列  $x(n)$  の複素ケプストラムは，次式で書けることになる．

$$\begin{aligned} x_{cc}(n) = & \delta(n) \log A + \frac{r(-1)^n}{n} \{u(n-1) + u(-n-1)\} - \sum_{i=1}^{n_a} \frac{a_i^n}{n} u(n-1) + \sum_{i=1}^{n_b} \frac{b_i^{-n}}{n} u(-n-1) \\ & + \sum_{i=1}^{n_c} \frac{c_i^n}{n} u(n-1) - \sum_{i=1}^{n_d} \frac{d_i^{-n}}{n} u(-n-1) \end{aligned} \quad (4.29)$$

さらに， $n$  が負，0，正の各場合に分けて表せば

$$x_{cc}(n) = \begin{cases} \frac{r(-1)^n}{n} + \sum_{i=1}^{n_b} \frac{b_i^{-n}}{n} - \sum_{i=1}^{n_d} \frac{d_i^{-n}}{n} & \text{for } n < 0 \\ \log A & \text{for } n = 0 \\ \frac{r(-1)^n}{n} - \sum_{i=1}^{n_a} \frac{a_i^{-n}}{n} + \sum_{i=1}^{n_c} \frac{c_i^{-n}}{n} & \text{for } n > 0 \end{cases} \quad (4.30)$$

となる．

式 (4.30) において， $a_i, b_i, c_i, d_i$  を含む項は，信号そのものの性質を表している．一方， $\frac{r(-1)^n}{n}$  は，線形位相項による影響であり，標本点毎に正負が交番する波形となる．正負が交番する成分を持つ波形に対してはノッチリフタをかけにくいことから，線形位相項は，あきらかにケプストラム解析を困難にする要因となり，あらかじめ除去しておくことが望まれる．線形位相項を除去するためには，解析する信号を離散フーリエ変換した後，その位相スペクトルから  $\omega$  に関して線形な成分を最小 2 乗法により推定して除去すれば良い．

式 (4.30) より，複素ケプストラム  $x_{cc}(n)$  は，信号  $x(n)$  が最小位相系列ならば， $n < 0$  で 0，最大位相系列ならば， $n > 0$  で 0 になる．最小位相系列あるいは，最大位相系列の場合，複素ケプストラムは，パワーケプストラムから，それぞれ

$$x_{cc}(n) = \begin{cases} x_{pc}(n), & \text{for } n > 0 \\ 0.5x_{pc}(n), & \text{for } n \neq 0 \\ 0, & \text{for } n < 0 \end{cases} \quad (4.31)$$

$$x_{cc}(n) = \begin{cases} 0, & \text{for } n > 0 \\ 0.5x_{pc}(n), & \text{for } n \neq 0 \\ x_{pc}(n), & \text{for } n < 0 \end{cases} \quad (4.32)$$

として算出できる．

ケプストラム解析では，スペクトル解析と同様に，切り出された有限長の系列を扱うことから，窓関数が必要になる．スペクトルが窓関数を用いることによりひずむのと同様，ケプストラムも窓関数によりひずむことになる．したがって，解析対象に応じて適切な窓関数を選択しなければならない．例えば，ケプストラム解析をエコーの検出等に用いる場合には，次式で示される指数窓 (exponential window) が有用である．

$$w(t) = \begin{cases} \alpha^n, 0 < \alpha < 1 & \text{for } 0 \leq t \leq N-1 \\ 0, & \text{others} \end{cases} \quad (4.33)$$

信号  $x(t)$  に、この窓関数  $w(t)$  をかけた離散時間系列の  $z$  変換は、 $z$  変換の性質から、 $X(\alpha^{-1}z)$  となる。これにより、エコーの大きさ  $a$  は、エコー時間を  $s$  として、 $a\alpha^s$  に変化したことになり、エリアジングの影響を少なく抑えることができる。通常、 $\alpha$  は、0.96 から 0.99 程度の値が用いられる。

演習 8：信号

$$y(t) = \cos\left(2\pi\frac{2}{21}t\right)e^{-\frac{t}{8}}, \quad t = 0, 1, \dots, 23$$

が、線形インパルス応答

$$h(t) = \begin{cases} 1, & t = 0 \\ 2, & t = 10 \\ 0, & \text{others} \end{cases}$$

を持つエコーを起こすシステムに入力され、 $x(t) = y(t) * h(t)$  が観測されるものとする。 $x(t)$  が与えられたものとして、複素ケプストラムを利用してエコーの成分を除去し、元々の信号  $y(t)$  を推定してみよう。もちろん、推定の際には、 $h(t)$  が分からないものとする。ただし、 $x(t)$  の系列長が  $N = 64$  となるようにゼロ詰めを行うものとする。また、 $h(t)$  が最大位相性を持ち、 $x(t)$  が線形位相項を持つことに注意しよう。

解答：図 4.4(a) は、エコーがかかる前の信号、すなわちエコーシステムへの入力  $y(t)$  であり、そのフーリエスペクトルを同図 (b) に示す。(c) は、エコーがかかった信号  $x(t)$  であり、我々の目的は、(c) が与えられて、これから (a) を復元することである。エコーがかかった観測信号  $x(t)$  に  $N = 64$  点までゼロをつめて、求められたフーリエスペクトルを (d) に示す。(d) に示したフーリエスペクトルの複素対数を取り、逆フーリエ変換して得られる複素ケプストラムを (e) に青色の実線で示す。これを見ると、10 標本点間隔で、インパルス状のピークがあることが分かる。これが 10 標本点分遅れるエコーの影響である。そこで、このピークをその前後 2 点の平均値で置き換えたものを (e) に赤色の実線で示す。これをフーリエ変換し、指数をとったものを (f) に示す。さらに、逆フーリエ変換して得られる原信号の復元値を (g) に示す。(a) がよく復元されていることがわかる。 $h(t)$  が最大位相系列であることから、エコーの成分であるエコー時間 10 の整数倍のケフレンシに現れるインパルス状のピークは、正のケフレンシでは現れないことに注意しよう。演習 7 では、 $h(t)$  が最小位相系列であることから、図 4.3(e) では、エコーによるピークは 0, 10, 20, 30, ... に現れたが、図 4.4(e) では、0, -10, -20, -30, ... に現れることになる。離散時間系列を対象にしていることから、ケプストラムも周期性を持つので、(e) において、原点を左右のどちらにとっても問題ないが、エコーによるピークの発生場所を分かりやすくするため、右端を原点にとって描いてある。

## 4.4 ヒルベルト変換

連続時間信号  $x(t)$  のフーリエ変換を  $X(\omega)$  として、

$$\hat{X}(\omega) = -i\text{sgn}(\omega)X(\omega) \quad (4.34)$$

は、周波数領域における  $X(\omega)$  のヒルベルト変換 (Hilbert transform) という。ただし、

$$\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$$

とする。一方、時間領域での  $x(t)$  のヒルベルト変換  $\hat{x}(t)$  は、 $\hat{X}(\omega)$  の逆フーリエ変換として定義される。 $-i\text{sgn}(\omega)$  の逆フーリエ変換が  $1/\pi t$  であることから、フーリエ変換の性質より、

$$\hat{x}(t) = \mathcal{F}^{-1}[\hat{X}(\omega)] = \mathcal{F}^{-1}[-i\text{sgn}(\omega)X(\omega)]$$



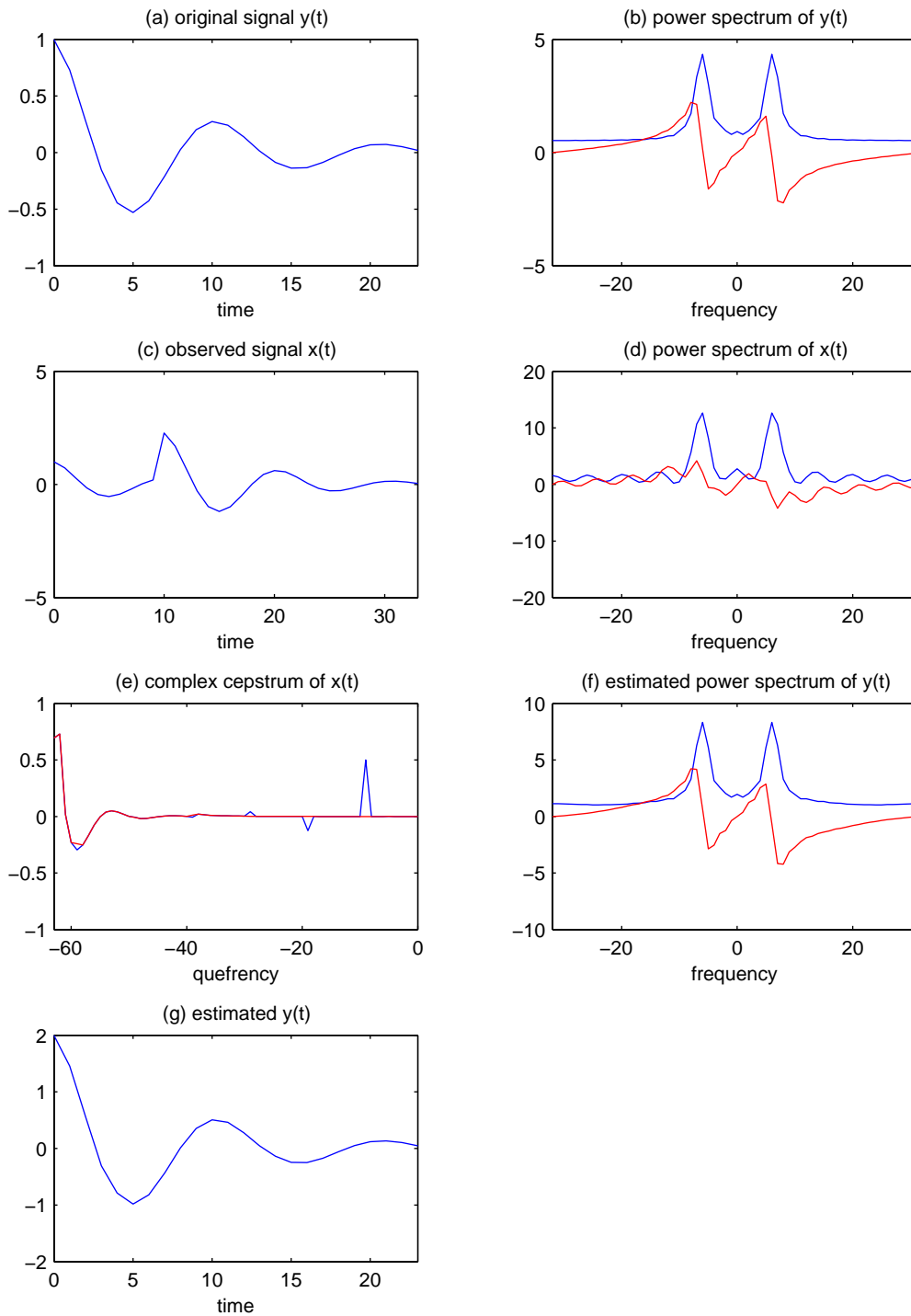


図 4.4: 演習 8 の結果：線形位相項があり，最大位相系列である信号のケプストラム解析

表 4.1: ヒルベルト変換の例

$x(t)$	$\hat{x}(t)$
$\hat{y}(t)$	$-y(t)$
$y(t)e^{i\omega t}$	$-iy(t)e^{i\omega t}$
$y(t)\cos(\omega t)$	$y(t)\sin(\omega t)$
$y(t)\sin(\omega t)$	$-y(t)\cos(\omega t)$

$$\begin{aligned}
 &= \mathcal{F}^{-1}[-i\text{sgn}(\omega)] * \mathcal{F}^{-1}[X(\omega)] = \frac{1}{\pi t} * x(t) \\
 &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{4.35}
 \end{aligned}$$

となる .

基本的なヒルベルト変換の例を表 4.1 に示そう .  $X(\omega)$  と  $\hat{X}(\omega)$  は , 位相の変化があるのみであるから ,  $x(t)$  と  $\hat{x}(t)$  のパワースペクトル , 自己相関関数は等しい . また ,  $x(t)$  と  $\hat{x}(t)$  は ,

$$\begin{aligned}
 \int_{-\infty}^{\infty} x(t) \cdot \hat{x}^*(t) dt &= \int_{-\infty}^{\infty} \mathcal{F}^{-1}[X(\omega) * \hat{X}^*(-\omega)] dt \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(\omega - \omega') \hat{X}^*(-\omega') d\omega' e^{i\omega t} d\omega dt \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t) e^{i\omega' t} dt \hat{X}^*(-\omega') d\omega' \\
 &= \int_{-\infty}^{\infty} X(-\omega') \hat{X}^*(-\omega') d\omega' \\
 &= \int_{-\infty}^{\infty} (-i)\text{sgn}(\omega') X(-\omega') X^*(-\omega') d\omega' \\
 &= i \int_{-\infty}^{\infty} \text{sgn}(\omega) |X(\omega)|^2 d\omega \\
 &= 0 \tag{4.36}
 \end{aligned}$$

となることから , 互いに直交している . 但し , 上式において ,  $\cdot^*$  は ,  $\cdot$  の共役を意味する .

## 4.5 解析信号

信号  $x(t)$  が実数値をとるものとし , そのヒルベルト変換を  $\hat{x}(t)$  とする .  $x(t)$  とそのヒルベルト変換  $\hat{x}(t)$  をそれぞれ実部 , 虚部とする

$$x_+(t) = x(t) + i\hat{x}(t) \tag{4.37}$$

$$x_-(t) = x(t) - i\hat{x}(t) \tag{4.38}$$

は , 解析信号と呼ばれている .  $x_+(t)$  と  $x_-(t)$  は本質的には同じ性質を持っており , どちらか一方 , 通常は  $x_+(t)$  が用いられる . 解析信号  $x_+(t)$  ,  $x_-(t)$  のそれぞれのフーリエ変換  $X_+(\omega)$  ,  $X_-(\omega)$  は ,

$$X_+(\omega) = X(\omega) + i(-i\text{sgn}(\omega))X(\omega) = X(\omega) + \text{sgn}(\omega)X(\omega) \tag{4.39}$$

$$X_-(\omega) = X(\omega) - i(-i\text{sgn}(\omega))X(\omega) = X(\omega) - \text{sgn}(\omega)X(\omega) \tag{4.40}$$

となることから，

$$X_+(\omega) = \begin{cases} 2X(\omega), & \omega > 0 \\ X(\omega), & \omega = 0 \\ 0, & \omega < 0 \end{cases} \quad (4.41)$$

$$X_-(\omega) = \begin{cases} 0, & \omega > 0 \\ X(\omega), & \omega = 0 \\ 2X(\omega), & \omega < 0 \end{cases} \quad (4.42)$$

となる．したがって，解析信号は，片側スペクトルを持つ，具体的には， $x_+(t)$  は，負の周波数成分がなく， $x_-(t)$  は，正の周波数成分がない信号になる．複素指数信号  $e^{\pm i\omega t}$  は，

$$e^{\pm i\omega t} = \cos(\omega t) \pm i \sin(\omega t) \quad (4.43)$$

で表されることから，解析信号の一つであることが分かる．

$x_+(t)$  のフーリエ変換が  $X_+(\omega)$  であるから，フーリエ変換の性質を利用すれば， $X_+(t)$  のフーリエ変換は  $2\pi x_+(-\omega)$  となる． $X_+(t)$  が実数をとる信号とすれば，そのフーリエ変換は，実部が偶関数，虚部が奇関数となるので，

$$\begin{aligned} \mathcal{F}[X_+(t)] &= 2\pi x_+(-\omega) \\ &= 2\pi(x(-\omega) + i\hat{x}(-\omega)) \\ &= 2\pi(x(\omega) - i\hat{x}(\omega)) \end{aligned} \quad (4.44)$$

となる． $X_+(t)$  は，式 (4.41) より， $t < 0$  で 0 となる．そのフーリエ変換の実部と虚部は，互いにヒルベルト変換の関係にある． $X_+(t)$  を線形インパルス応答  $h(t)$  と考えれば，線形インパルス応答が  $t < 0$  で 0 となる，つまりシステムが因果律を満たすならば，その周波数伝達関数の実部と虚部は，互いに，ヒルベルト変換の関係にあることが分かる．

離散時間信号  $x(t)$  の解析信号  $x_+(t)$  は， $x(t)$  を離散フーリエ変換することにより  $X(\omega)$  を得，それを式 (4.41) に適用することにより  $X_+(\omega)$  を得，それを逆離散フーリエ変換することにより求めることができる．ただし，この際， $x(t)$  の系列長  $N$  が偶数の場合，

$$X_+(\omega) = \begin{cases} 2X(\omega), & \omega \in \{1, 2, \dots, N/2 - 1\} \\ X(\omega), & \omega \in \{0, N/2\} \\ 0, & \omega \in \{N/2 + 1, \dots, N - 1\} \end{cases}$$

となり， $x(t)$  の系列長  $N$  が奇数の場合，

$$X_+(\omega) = \begin{cases} 2X(\omega), & \omega \in \{1, 2, \dots, (N - 1)/2\} \\ X(\omega), & \omega = 0 \\ 0, & \omega \in \{(N - 1)/2 + 1, \dots, N - 1\} \end{cases}$$

となることに注意しよう．

ところで， $\cos(\omega t)$  を搬送波とする信号  $m(t)$  の振幅変調 (Amplitude Modulation) は，

$$x(t) = m(t) \cos(\omega t)$$

として行われる． $m(t)$  は  $x(t)$  の包絡線と呼ばれる．受信端では， $x(t)$  を受信し，からその包絡線  $m(t)$  を抽出することにより，復調が行われる．こうした振幅変調とヒルベルト変換の関係を見てみよう．

振幅変調された信号  $x(t)$  のヒルベルト変換は，表 4.1 より，

$$\hat{x}(t) = m(t) \sin(\omega t)$$

であるから， $x(t)$  の解析信号は，

$$\begin{aligned}x_{\pm}(t) &= m(t) \cos(\omega t) \pm i m(t) \sin(\omega t) \\ &= m(t)e^{\pm i\omega t}\end{aligned}\tag{4.45}$$

となる．解析信号の絶対値をとれば，

$$|x_{\pm}(t)| = |m(t)|$$

となる． $m(t) \geq 0, \forall t$  ならば，振幅変調を受けた信号  $x(t)$  の復調は，その信号の解析信号の絶対値をとることに他ならない． $m(t) \geq 0, \forall t$  という条件は，振幅変調を行う際の一般的な条件である．

また，振幅変調された信号  $x(t)$  の解析信号  $x_+(t) = x(t) + i\hat{x}(t)$  の位相

$$\phi_m(t) \equiv \tan^{-1} \frac{\hat{x}(t)}{x(t)}\tag{4.46}$$

は，瞬時位相と呼ばれ，瞬時位相の時間微分

$$\omega_m(t) \equiv \frac{d}{dt} \phi_m(t)\tag{4.47}$$

は，瞬時周波数と呼ばれる．

演習 9：搬送波を  $\sin(\omega_c t)$ ， $\omega_c = 2\pi \cdot 0.02$  として， $m(t) = 0.3 \sin(\omega_m t) + 0.5$ ， $\omega_m = 2\pi \cdot 0.004$  を振幅変調し，それからヒルベルト変換を行い解析信号を求めることにより復調を行ってみよう．ただし，時刻  $t$  は離散化されており， $t = 0, 1, \dots, 999$  をとるものとする．

解答：搬送波を図 4.5(a)，信号  $m(t)$  を同図 (b) に示す．振幅変調された信号  $x(t)$  とそのヒルベルト変換を同図 (c) の青色と赤色の実線で示す．これらは， $x(t)$  の解析信号の実部と虚部に一致する．また，解析信号の絶対値として復調された信号  $m(t)$  を同図 (d) に示す．

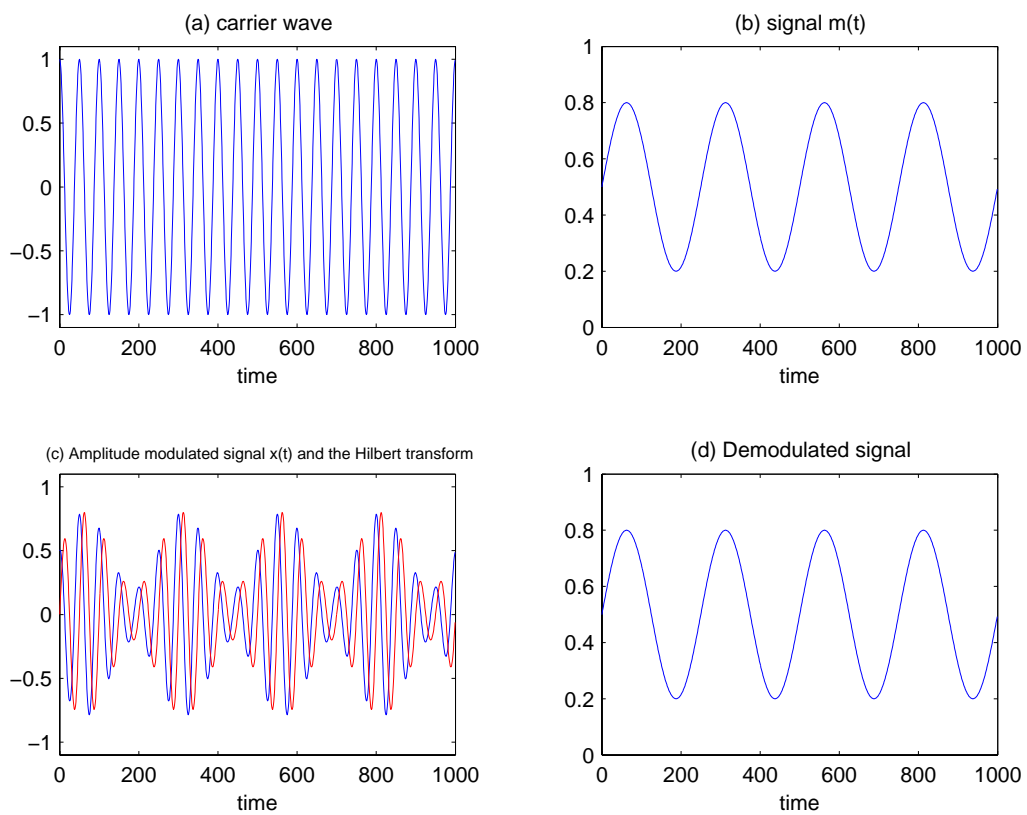


図 4.5: 演習 9 の結果：解析信号を用いた振幅変調された信号の復調