通信会議における遠方音声収音のための残響抑圧方式の研究

古家, 賢一

https://doi.org/10.15017/459055

第6章

実数データに対するDFTの計算精度における有限レジスタ長の影響

6.1 まえがき

第4章でSemi-blind残響抑圧を提案し，第5章でSemi-blind残響抑圧処理の高速化手法を報告した．ここまでの検討により，一般的な数万円の安価なパソコンでリアルタイム動作させることができる残響抑圧処理が可能となった．この処理の高速化には，FFT（高速フーリエ変換）を利用しているが，さらに，実数データのFFT計算に関してはさらに高速化できる$N/2$点ＦＦＴ法や高速ハートレー変換（FHT）を利用できる．

一方，ＦＦＴの計算を計算機上で実現する上で，計算量の他に考慮すべき要因として，計算機の有限レジスタ長の影響による計算精度の問題がある．計算機でＦＦＴを計算した場合には，どの程度の計算誤差があるか常に把握しておく必要がある．一般に，レジスタ長が短くなれば計算時間は短くなるが，計算精度は悪くなるというトレードオフの関係が存在し，ＦＦＴの計算時間を考える上でもレジスタ長と計算精度の関係を明かにすることは重要である．

残響抑圧処理システムを一般家庭に普及させていくためには，今後さらに数千円程度のDSPチップでの実現が必要である．DSPチップは，通常，パソコンなどと同じ浮動小数点演算が計算できるものは数万円するが，有限レジスタ長

の固定小数点演算するものは数千円と安価である．但し，有限レジスタ長の固定小数点演算の場合，十分に演算精度について設計しないと計算誤差により，期待する効果がまったく得られなくなる．

残響抑圧処理では，FFT 計算以外でも共役勾配法などの計算を行っている．これらすべての計算における計算精度について解析することは困難であるが，計算の大部分を占める FFT 計算での計算精度が推定できれば，パソコンなどで浮動小数点演算を行いそれとの相対比較で，どの程度のレジスタ長の固定小数点演算が必要か見積もることができる．

本章では，残響抑圧処理での計算機のレジスタ長と計算精度の関係を明かにするため，理論解析及び実験をおこない，FHT 及び $N/2$ 点ＦＦＴを用いたＮ点ＦＦＴの計算方法について計算精度について評価する．以下，第 6.2 節では実数データの FFT 計算手法として離散的ハートレー変換，高速ハートレー変換について述べる．第 6.3 節では，ＦＨＴにおけるレジスタ長と計算精度の関係を理論的に解析し，第 6.4 節で実験および比較をおこない，第 6.5 節でそれらの結果のまとめを述べる．

6.2 実数データに対するFFTの計算手法

一般的には，離散的フーリエ変換（ＤＦＴ）の高速計算はデータ数の増加とともにますます必要となる．N 個のデータに対するＤＦＴをその定義式通りに計算すると，N^2 に比例した計算量が必要である．この数は N が大きなときには膨大となり，高速の計算機を用いても長い計算時間を必要とする．このＤＦＴを効率よく高速に計算する周知のアルゴリズムとして高速フーリエ変換（ＦＦＴ）がある [60]．ＦＦＴを用いれば，計算量は $N \log N$ に比例し，ＤＦＴ計算の大幅な高速化が実現できる．

われわれが扱う信号は，一般に音響信号，画像信号などの直接的には実数データであることが多い．ところがＦＦＴでは複素データに対してＤＦＴを計算するようになっており，実数データに対してＦＦＴを適用する場合には，実数データを実部に持ち虚部がすべて零である複素データを作らなければならない．虚部がすべて零であるということはその部分は情報を持っておらず冗長なデータとなっている．このためＦＦＴで行われる計算の約半分が冗長な計算となってしまう．

実数データに対するＦＦＴの冗長な計算を除くため次の方法が提案されている [61]．N （偶数）点の実数データに対する N 点ＤＦＴを計算するのに，データを $N/2$ 個づつに分割しそれぞれを実部，虚部に持つ $N/2$ 点の複素データをつくり，それに対し $N/2$ 点ＦＦＴを適用し，それから N 点ＤＦＴを計算する．この方法だと虚部を零にして計算するという無駄なことを行わないので，計算量は通常のＦＦＴの約半分になる．

一方，Bracewell は実数データに対する変換としてＤＦＴに類似した離散的ハートレー変換（ＤＨＴ）を提案し [62]，その高速計算アルゴリズムである高速ハートレー変換（ＦＨＴ）を用いてＤＦＴが高速に計算できることを示した [63]．ＦＨＴではもともと実数データに対する変換であるＤＨＴを用いているため複素データを扱う必要がなくＦＦＴのような冗長な計算を行わずにすむ．また，Sorensen らは，ＦＨＴの計算量について詳細に検討し，ＦＦＴの約半分の計算量でＤＦＴを計算できることを示した [64],[65].

6.2.1 離散的ハートレイ変換（DHT）

連続時間での実数関数 $x(t)$ に対してハートレイ変換は (6.1) 式により定義される [66].

$$H(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x(t)\mathrm{cas}\omega t \; dt \tag{6.1}$$

ここで，$\mathrm{cas}\,\omega t = \cos\omega t + \sin\omega t$ である．その逆変換は，(6.2) 式となる．

$$x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H(\omega)\mathrm{cas}\,\omega t \; d\omega \tag{6.2}$$

この連続時間でのハートレイ変換に対応して，離散時間でのハートレイ変換，つまり離散的ハートレイ変換は実数列 $x(n)$ に対して (6.3) 式で与えられる [62].

$$H(k) = \sum_{n=0}^{N-1} x(n)\mathrm{cas}\,(2\pi nk/N) \tag{6.3}$$
$$(k = 0, 1, 2, \cdots, N-1)$$

その逆変換は，

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} H(k)\mathrm{cas}(2\pi nk/N) \tag{6.4}$$
$$(n = 0, 1, 2, \cdots, N-1)$$

である．$x(n)$ の離散的ハートレイ変換 $H(k)$ と離散的フーリエ変換 $X(k)$ の間には，次式の関係があり

$$X(k) = \frac{1}{2}\{H(k) + H(N-k)\} - \frac{j}{2}\{H(k) - H(N-k)\} \tag{6.5}$$

特に，パワースペクトル $|X(k)|^2$ のみが必要なときには，さらに簡単に $H(k)$ から

$$|X(k)|^2 = \frac{1}{2}\{H^2(k) + H^2(N-k)\} \tag{6.6}$$

と計算できる．

6.2.2 高速ハートレイ変換（ＦＨＴ）アルゴリズム

高速ハートレイ変換アルゴリズムにも，ＦＦＴ同様，種々の方式があるが [64]，ここでは，基数2の場合の時間間引き・入力側ビット逆順序アルゴリズム [63] について述べる．

長さ $N=2^\nu$（ν:正の整数）の実数列 $x(n)$ のＤＨＴを計算することを考える．このとき，$x(n)$ を次式のように二つに分割すると，

$$\begin{aligned} a(m) &= x(2m) \\ b(m) &= x(2m+1) \\ &\qquad (m=0,1,2,\cdots,N/2-1) \end{aligned} \tag{6.7}$$

$x(n)$ のＤＨＴは (6.8) 式のように分解できる．

$$\begin{aligned} H(k) &= \sum_{n=0}^{N-1} x(n)\mathrm{cas}\,(2\pi nk/N) \\ &= \sum_{m=0}^{N/2-1} \{a(m)\mathrm{cas}\,(2\pi mk/N) + b(m)\mathrm{cas}\,(2\pi(2m+1)k/N)\} \\ &= \sum_{m=0}^{N/2-1} \{a(m)\mathrm{cas}\,(2\pi mk/(N/2)) + b(m)\{\cos(2\pi k/N)\mathrm{cas}\,(2\pi mk/(N/2)) \\ &\qquad + \sin(2\pi k/N)\mathrm{cas}\,(-2\pi mk/(N/2))\}\} \\ &= H_a(k) + H_b(k)\cos(2\pi k/N) + H_b(N-k)\sin(2\pi k/N) \end{aligned} \tag{6.8}$$

ただし，$H_a(k), H_b(k)$ はそれぞれ $a(m), b(m)$ のＤＨＴであり，$k \geq N/2$ での $H_a(k), H_b(k)$ の値は $0 \leq k \leq N/2-1$ での値を繰り返すものとする．(6.8) 式は，N 個のデータに対するＤＨＴが $N/2$ 個のデータに対するＤＨＴから求められることを示している．$H(k)$ を (6.3) 式から直接計算した場合，N^2 回の乗算が必要であるが，(6.8) 式を用いて計算すれば $(2N+N^2/2)$ 回の乗算ですむ．したがって，N が大きなときには，計算量を半分にすることができる．さらに，この分解を繰り返していくと $N=2^\nu$ に対して ν 回の分解がおこなえる．各段での乗

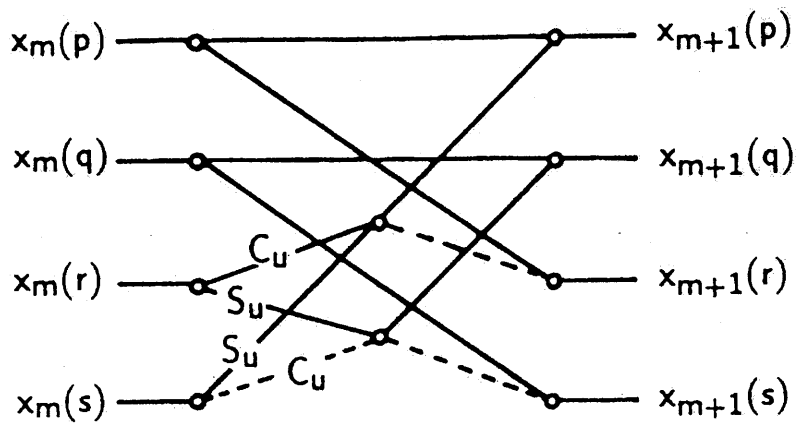

図 6.1: 時間間引き形FHTアルゴリズムのバタフライ．実線：加算，破線：減算

算回数は$2N$回なので，全体の乗算回数は$2N\log_2 N$回となり直接計算した場合に比べ大幅な計算量の削減ができる．

各段での演算は，(6.9) 式のこれもバタフライ演算と呼ばれる基本演算によって構成されている．

$$\begin{aligned} x_{m+1}(p) &= x_m(p) + C_u x_m(r) + S_u x_m(s) \\ x_{m+1}(q) &= x_m(q) + C_u x_m(r) + S_u x_m(s) \\ x_{m+1}(r) &= x_m(p) + C_u x_m(r) + S_u x_m(s) \\ x_{m+1}(s) &= x_m(q) + C_u x_m(r) + S_u x_m(s) \end{aligned} \tag{6.9}$$

ただし，添字の$m, m+1$はそれぞれm段目，$m+1$段目の出力配列，p, q, r, sは各配列での位置を示し，$C_u = \cos(2\pi u/N), S_u = \sin(2\pi u/N)$である．($m = 0$は入力配列を$m = \nu$は出力配列を示す.) これが行われる様子を図 6.1 に示す．このバタフライ演算は，一時的に中間配列を要するが定位置計算となっている．つまり，(6.9) 式によって計算された結果を，それらを計算するのに用いた入力配列にもどしても，他のバタフライ演算には影響を与えない構造になっている．また，出力を配列を順序通りに配列させたまま (6.8) 式の分解をこのバタフライ

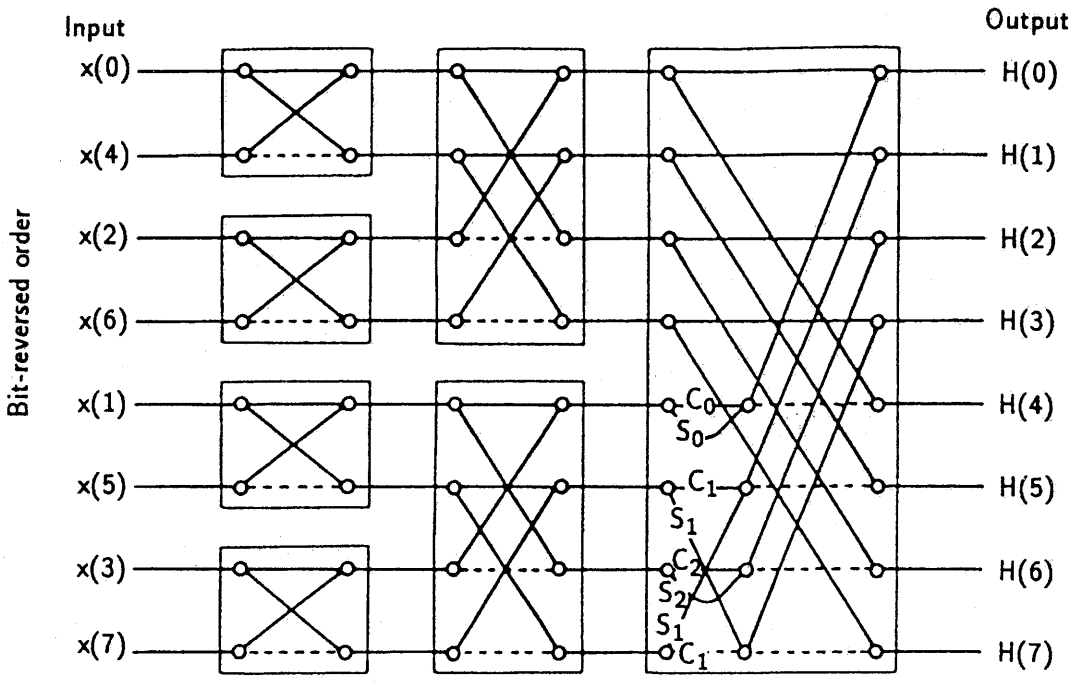

図 6.2: $N = 8$ のときの高速ハートレー変換フローダイアグラム実線：加算，破線：減算

演算を用いておこなう場合，入力配列はビット逆順序並び換えをおこなわなければならない．ビット逆順序並び換えとは，n 番目の入力データ $x(n)$ において n が二進数 $b_1 b_2 \cdots b_{i-1} b_i$ で表現されているときに，$b_i b_{i-1} \cdots b_2 b_1$ となる番号の位置へ $x(n)$ を移動することである．

図 6.2 に $N = 8$ の場合のＦＨＴアルゴリズムのフローグラフを示す．図 6.2 をみるとわかるように，１段目，２段目においては $-1, 1$ の乗算のみとなるので，実際には加減算のみで計算は行われる．また，図 6.1 でのバタフライ演算では次式の三角関数の対称性

$$\begin{aligned} \cos(2\pi(k + N/2)/N) &= -\cos(2\pi k/N) \\ \sin(2\pi(k + N/2)/N) &= -\sin(2\pi k/N) \end{aligned} \tag{6.10}$$

を利用して $x_{m+1}(k)$ と $x_{m+1}(k + N/2)$ の計算における乗算を同時に行っている

ので，各段での乗算はN回ですんでいる（図6.2では8回）．したがって，図6.1のバタフライ演算を用いて分解すれば，ＦＨＴでの乗算回数は最大$N\log_2 N$回となる．

N個のデータに対してN点ＤＦＴを基数2のN点ＦＦＴで求める場合，$(N/2)\log_2 N$回の複素乗算が，つまり$2N\log_2 N$回の実数乗算が必要である．それに対し，ＦＨＴを利用してＤＦＴを求める場合，ＤＨＴの計算で$N\log_2 N$回，(6.5)式を用いてＤＨＴからＤＦＴを計算するのに$2N$回，したがって全体で$N(\log_2 N+2)$回の実数乗算ですむ．これは，Nが大きな場合，ＦＨＴを利用してＤＦＴを求めれば，通常のN点ＦＦＴを用いたときの約半分の計算時間ですむことを示している．

6.3　ＦＨＴにおける有限レジスタ長の影響の解析

6.3.1　固定小数点演算ＦＨＴの場合（スケーリングなし）

レジスタ長$b+1$ビットの固定小数点演算ＦＨＴで実数列$x(n)$のＤＨＴ$H(k)$を計算するときの計算誤差について考える．ＦＨＴでは，各段ごとに(6.9)式のバタフライ演算がおこなわれる．m段目出力から$m+1$段目出力のp番要素を計算するときの計算誤差を$\varepsilon(m,p)$とする．固定小数点演算では，丸め誤差は乗算のみで生じ，加算では生じないこと，また計算誤差は加算的であることを考え，さらに乗算での丸め誤差に対して次の仮定をする．

- 丸め誤差は，$-(1/2)2^{-b}$から$(1/2)2^{-b}$の範囲
  に一様分布する．したがって，その分散は$2^{-2b}/12$である．
- 誤差は互いに無相関である．
- すべての，誤差は入力と無相関である．

この仮定は多少乱暴であるが，信号が音声などの複雑な場合には信号と誤差との間の相関は減少し，誤差も互いに無相関となることがわかっている [67]．このとき，$\varepsilon(m,p)$ の二乗平均は，

$$E[\varepsilon(m,p)] = \frac{2^{-2b}}{6} \tag{6.11}$$

となる．ここで $E[\cdot]$ は期待値を表す．

出力 $H(k)$ での計算誤差を考えるときには各段でおきた誤差の伝搬を考えなければならない．つまり，(6.9) 式中の $x_{m+1}(m)$ の計算において誤差はその段の乗算によるものだけでなく $x_m(p), x_m(r), x_m(s)$ をとおして伝搬してくる誤差がある．$F_m(p), F_m(r), F_m(s)$ をそれぞれ，その伝搬してきた誤差とすると $x_{m+1}(p)$ での誤差の総和 $F_{m+1}(p)$ は，

$$F_{m+1} = \varepsilon(m,p) + F_m(p) + C_u F_m(r) + S_u F_m(s) \tag{6.12}$$

である．その二乗平均は，

$$E[F_{m+1}^2(p)] = \frac{2^{-2b}}{6} + E[F_m^2(p)] + C_u^2 E[F_m^2(r)] + S_u^2 E[F_m^2(s)] \tag{6.13}$$

となる．入力信号には誤差が無い，つまり

$$F_0(p) = F_0(r) = F_0(s) = 0 \tag{6.14}$$

とすると，(6.13) 式より漸化的に導かれる $E[F_m^2(p)]$ は，その要素の位置 p に無関係となる．

$$E[F_m^2(p)] = E[F_m^2(r)] = E[F_m^2(s)] \equiv E[F_m^2] \tag{6.15}$$

(6.15) 式より (6.13) は，

$$\begin{aligned} E[F_{m+1}^2] &= \frac{2^{-2b}}{6} + E[F_m^2](1 + C_u^2 + S_u^2) \\ &= \frac{2^{-2b}}{6} + 2E[F_m^2] \end{aligned} \tag{6.16}$$

(6.14) 式の条件のもとで (6.16) 式を解くと，

$$E[F_m^2] = \frac{2^{-2b}}{6}(2^m - 1) \tag{6.17}$$

となり，出力 $H(k)$ での誤差 F_ν の二乗平均は，

$$E[F_\nu^2] = \frac{2^{-2b}}{6}(N - 1) \tag{6.18}$$

となる．

ＤＨＴを固定小数点演算で計算するときには，オーバーフローが生じる場合がある．このオーバーフローが生じないためには，$|H(k)| \leq 1$ である必要がある．このためには，

$$|x(n)| \leq \frac{1}{N} \tag{6.19}$$

であれば十分である．出力での誤差対信号比を導くため，$-1/N$ と $1/N$ との間を一様分布する白色雑音を入力信号 $x(n)$ とする．このとき，$x(n)$ が (6.19) 式を満たすことに注意する．この $x(n)$ の二乗平均は，

$$E[x^2(n)] = \frac{1}{3N^2} \tag{6.20}$$

であり，そのＤＨＴ $H(k)$ の二乗平均は，

$$\begin{aligned} E[H^2(k)] &= E[\{\sum_{n=0}^{N-1} x(n)\mathrm{cas}\,(2\pi nk/N)\}^2] \\ &= \sum_{n=0}^{N-1} E[x^2(n)]\mathrm{cas}^{\,2}(2\pi nk/N) \\ &= E[x^2(n)] \sum_{n=0}^{N-1} \mathrm{cas}^{\,2}(2\pi nk/N) \\ &= \frac{1}{3N} \end{aligned} \tag{6.21}$$

(6.18),(6.21) 式より出力での誤差対信号比は，

$$\frac{E[F_\nu^2]}{E[H^2(k)]} = 2^{-2b-1}N(N-1) \tag{6.22}$$

とかける．したがって，出力での誤差対信号比は，データ数 N の二乗に比例し，レジスタ長が１ビット増えると dB 換算で 6dB 小さくなる．

6.3.2 固定小数点演算ＦＨＴの場合（スケーリングあり）

次に，固定小数点演算のＦＦＴにも用いられている各段のバタフライにおいて，1/2 のスケーリングを行う方法について考える．この方法は，各段のバタフライにおいてオーバーフローを避けるために 1/2 を乗じる操作を行うため，入力信号として，

$$|x(n)| \leq 1 \tag{6.23}$$

を入力できるので，各段のバタフライでスケーリングを行わない場合に比べて，計算精度の点で有利である．しかも，1/2 を乗ずる操作は二進数では 1 ビットのシフトで実現されるのでそれほど計算時間も増えない．この場合の $\varepsilon^2(m,p)$ は，(6.9) 式のバタフライ演算にさらに 1/2 を乗ずる操作が加わるので 3 回の乗算の丸め誤差の和となる．

$$E[\varepsilon^2(m,p)] = \frac{2^{-2b}}{4} \tag{6.24}$$

このとき，$x_{m+1}(p)$ の計算誤差 $F_{m+1}(p)$ は，

$$F_{m+1}(p) = \varepsilon(m,p) + \frac{1}{2}F_m(p) + \frac{C_u}{2}F_m(r) + \frac{S_u}{2}F_m(s) \tag{6.25}$$

である．その二乗平均は，

$$E[F^2_{m+1}(p)] = \frac{2^{-2b}}{4} + \frac{1}{4}E[F^2_m(p)] + \frac{C_u^2}{4}E[F^2_m(r)] + \frac{S_u^2}{4}E[F^2_m(s)] \tag{6.26}$$

と漸化的に表せる．(6.14) 式と同様に入力信号には誤差がないとすると，$E[F^2_{m+1}(p)]$ は，その位置 p に無関係となり (6.26) 式は，

$$\begin{aligned} E[F^2_{m+1}] &= \frac{2^{-2b}}{4} + \frac{1}{4}E[F^2_m](1 + C_u^2 + S_u^2) \\ &= \frac{2^{-2b}}{4} + \frac{1}{2}E[F^2_m] \end{aligned} \tag{6.27}$$

となる．これを解くと次式となる．

$$E[F^2_m] = 2^{-2b-1}(1 - \frac{1}{2^m}) \tag{6.28}$$

したがって，出力 $H(k)$ での誤差 F_ν の二乗平均は，

$$E[F_\nu^2] = 2^{-2b-1}(1 - \frac{1}{N}) \tag{6.29}$$

となる．出力での誤差対信号比を導くため，-1 と 1 との間で一様分布する白色雑音 $x(n)$ を入力する．このとき，各段のバタフライ演算で $1/2$ のスケーリングを行っているのでオーバーフローをおこさないことに注意する．出力 $H(k)$ の二乗平均は (6.21) 式と同じであるから，出力での誤差対信号比は，

$$\frac{E[F_\nu^2]}{E[H^2(k)]} = 2^{-2b-1}3(N-1) \tag{6.30}$$

であり，N が大きなときには N に比例し，レジスタ長が 1 ビット増えると dB 換算で誤差対信号比は 6dB 小さくなる．

6.3.3 浮動小数点演算ＦＨＴの場合

仮数部レジスタ長 $b+1$ ビットの浮動小数点演算ＦＨＴで白色雑音 $x(n)$ のＤＨＴを求める場合について考える．浮動小数点演算での誤差は仮数にのみ影響を与える．したがって，浮動小数点演算では絶対誤差よりも，信号に対する相対誤差のほうが重要である．その相対誤差に対して，次の仮定をする．

- 誤差は，-2^{-b} から 2^{-b} の間を一様分布する白色雑音
  である．したがって，その分散は $\frac{2^{-2b}}{3}$ である．
- 誤差は互いに無相関である．
- 入力信号と誤差とは無相関である．

(6.9) 式のバタフライ演算での $x_{m+1}(p)$ の計算において乗算で生じる誤差 $\varepsilon_1(m,p)$ の二乗平均は，

$$E[\varepsilon_1^2(m,p)] = \frac{1}{3}2^{-2b}\{E[C_u^2 x_m^2(r)] + E[S_u^2 x_m^2(s)]\}$$

$$= \frac{1}{3}2^{-2b}E[x_m^2(p)] \tag{6.31}$$

となる．但し，入力信号が白色雑音であることから $E[x_m^2(p)] = E[x_m^2(r)] = E[x_m^2(s)]$ である．加算で生じる誤差 $\varepsilon_2(m,p)$ の二乗平均は，

$$\begin{aligned} E[\varepsilon_2^2(m,p)] &= \frac{1}{3}2^{-2b}\{E[x_m^2(p)] + (E[C_u^2x_m^2(r)] + E[S_u^2x_m^2(s)])\} \\ &\quad + \frac{1}{3}2^{-2b}\{E[C_u^2x_m^2(r)] + E[S_u^2x_m^2(s)]\} \\ &= 2^{-2b}E[x_m^2(p)] \end{aligned} \tag{6.32}$$

となる．したがって，$x_{m+1}(p)$ の計算で生じた誤差 $\varepsilon(m,p)$ の二乗平均は，

$$\begin{aligned} E[\varepsilon^2(m,p)] &= E[\varepsilon_1^2(m,p)] + E[\varepsilon_2^2(m,p)] \\ &= \frac{4}{3}2^{-2b}E[x_m^2(p)] \end{aligned} \tag{6.33}$$

となる．また，$x(n)$ が白色雑音であることから，

$$E[x_m^2(p)] = 2^m E[x^2(n)] \tag{6.34}$$

となる．$x_{m+1}(p)$ の計算での伝搬してきた誤差も含めた誤差の総和 $F_{m+1}(p)$ は，

$$F_{m+1}(p) = \varepsilon(m,p) + F_m(p) + C_uF_m(r) + S_uF_m(s) \tag{6.35}$$

であり，この式と (6.33),(6.34) 式より誤差の総和 $F_{m+1}(p)$ の二乗平均は，

$$E[F_{m+1}^2(p)] = \frac{4}{3}2^{-2b}2^m E[x^2(n)] + E[F_m^2(p)] + C_u^2E[F_m^2(r)] + S_u^2E[F_m^2(s)] \tag{6.36}$$

となる．さらに，(6.14) 式と同様に入力信号には誤差は含まれないとすると，$E[F_{m+1}^2(p)]$ はその位置 p に無関係となり，(6.36) 式は，

$$\begin{aligned} E[F_{m+1}^2] &= \frac{4}{3}2^{-2b}2^m E[x^2(n)] + E[F_m^2](1 + C_u^2 + S_u^2) \\ &= \frac{4}{3}2^{-2b}2^m E[x^2(n)] + 2E[F_m^2] \end{aligned} \tag{6.37}$$

となる．これを解くと，

$$E[F_m^2] = \frac{2}{3}2^{-2b}m2^m E[x^2(n)] \tag{6.38}$$

となり，出力 $H(k)$ での誤差 F_ν の二乗平均は，

$$\begin{aligned} E[F_\nu^2] &= \frac{2}{3}2^{-2b}\nu 2^\nu E[x^2(n)] \\ &= \frac{2}{3}2^{-2b}\nu N E[x^2(n)] \end{aligned} \tag{6.39}$$

となる．出力 $H(k)$ の二乗平均が，

$$E[H^2(k)] = NE[x^2(n)] \tag{6.40}$$

であるので，出力での誤差対信号比は，

$$\frac{E[F_\nu^2]}{E[H^2(k)]} = \frac{2}{3}2^{-2b}\nu \tag{6.41}$$

となり，この式から浮動小数点演算ＦＨＴの場合には，出力での誤差対信号比が $\nu = \log_2 N$ に比例すること，および仮数部レジスタ長が 1 ビット増えると dB 換算で誤差対信号比は 6dB 小さくなることがわかる．

6.3.4 ＤＨＴからＤＦＴの計算での誤差

ＦＨＴを利用してＤＦＴを計算する場合，ＦＨＴのでの計算誤差に加えＤＨＴからＤＦＴを計算するときに生じる誤差を考えなければならない．$x(n)$ のＤＦＴ $X(k)$ での計算誤差を $F(k)$ とする．固定小数点演算を用いた場合，ＤＨＴからＤＦＴの計算に実部と虚部でそれぞれ一回ずつの乗算が必要なことから，$F(k)$ の二乗平均はＦＨＴからの計算誤差に乗算 2 回分の誤差が加わる．

$$E[|F(k)|^2] = E[F_\nu^2] + \frac{1}{6}2^{2b} \tag{6.42}$$

出力 $X(k)$ は,

$$E[|X(k)|^2] = \frac{1}{3N} \tag{6.43}$$

となるので，各段のバタフライ演算で 1/2 のスケーリングをしないときの出力での誤差対信号比は (6.18),(6.42) ,(6.43) 式から,

$$\begin{aligned} \frac{E[|F(k)|^2]}{E[|X(k)|^2]} &= \frac{N}{2}(N-1)2^{-2b} + \frac{N}{2}2^{-2b} \\ &= \frac{N^2}{2}2^{-2b} \end{aligned} \tag{6.44}$$

となる．スケーリングをしたときには，(6.29),(6.42),(6.43) 式から,

$$\frac{E[|F(k)|^2]}{E[|X(k)|^2]} = \frac{3N}{2}(1-\frac{1}{N})2^{-2b} + \frac{N}{2}2^{-2b} \tag{6.45}$$

となる.

浮動小数点演算ＦＨＴを利用した場合の $F(k)$ の二乗平均は,

$$E[|F(k)|^2] = E[F_\nu^2] + \frac{2}{3}2^{-2b}E[H^2(k)] \tag{6.46}$$

出力 $X(k)$ の二乗平均は,

$$E[|X(k)|^2] = NE[x^2(n)] = E[H^2(k)] \tag{6.47}$$

であるので，出力での誤差対信号比は (6.39), (6.46),(6.47) 式から,

$$\begin{aligned} \frac{E[|F(k)|^2]}{E[|X(k)|^2]} &= \frac{2}{3}2^{-2b}\nu + \frac{2}{3}2^{-2b} \\ &= \frac{2}{3}2^{-2b}(\nu+1) \end{aligned} \tag{6.48}$$

となる.

6.4 計算誤差の測定

6.4.1 測定方法

計算誤差の測定には，メインCPUが8086で，実装メモリ1Mバイト，演算プロセッサ8087を持つ16ビットコンピュータAICOM16を用いた．プログラムは，C言語を用いて作成され，演算時のレジスタ長はソフトウェア上でのビット演算を用いて操作した．今回の測定では，固定小数点演算のレジスタ長を8ビットから24ビット，浮動小数点演算の仮数部レジスタ長を8ビットから24ビットまで変化させ，8バイト浮動小数点演算（仮数部レジスタ長56ビット）で計算した値を真値と考えてそれとの差を計算誤差として測定した．固定小数点演算については各段バタフライで1/2スケーリングを行う場合と行わない場合の二通りを測定した．また，入力信号の性質の違いによる計算誤差をみるため，入力信号として$-1/N$と$1/N$の間で一様分布する白色雑音と正弦波の二種類を用いた．

6.4.2 測定結果

図6.3，図6.4，図6.5にＦＨＴを用いてＤＦＴを計算した場合の計算誤差を示す．図6.3は固定小数点演算ＦＨＴにおいて各段バタフライでスケーリングを行わない場合，図6.4は固定小数点演算ＦＨＴにおいて各段バタフライでスケーリングを行う場合，図6.5は浮動小数点演算ＦＨＴの場合である．各図の(a),(b)は，それぞれ白色雑音入力の場合，正弦波入力の場合である．参考のため，ＦＨＴと同じ条件で従来の$N/2$点ＦＦＴを用いて計算した場合の測定結果も併せて示してある．

図6.3，図6.4，図6.5から共通して，レジスタ長を4ビット増やすことによっ

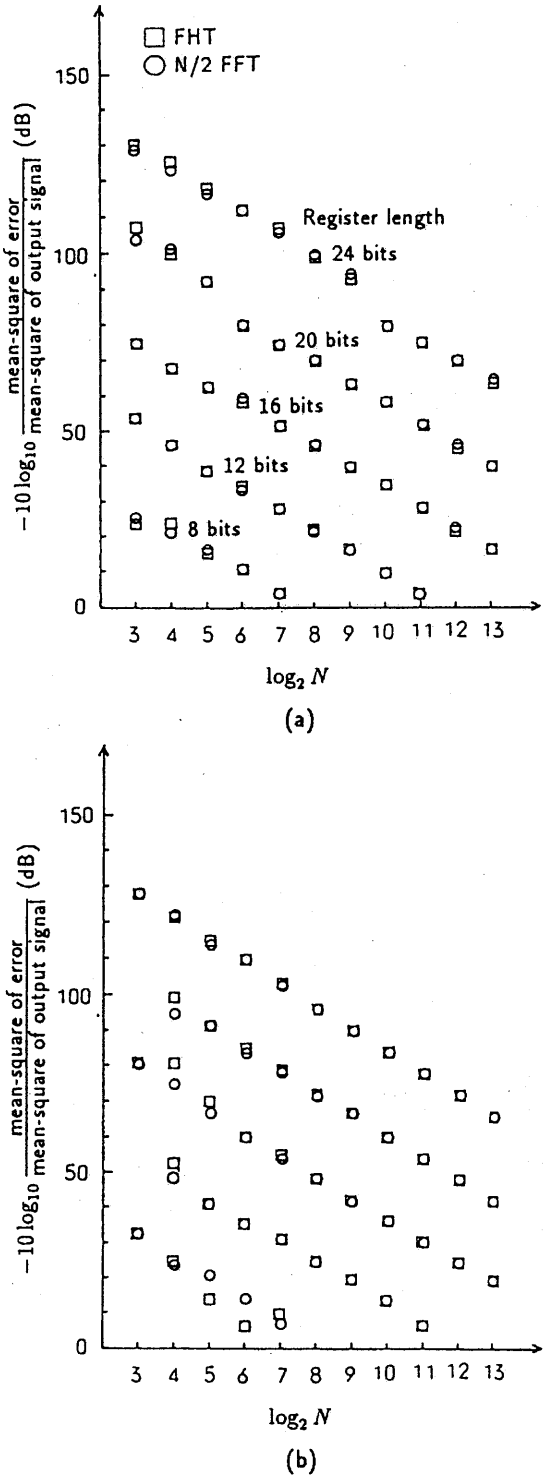

図 6.3: 固定小数点 FHT，固定小数点 $N/2$FFT を用いた DFT 計算における計算誤差　(a) 白色雑音入力，(b) 正弦波入力

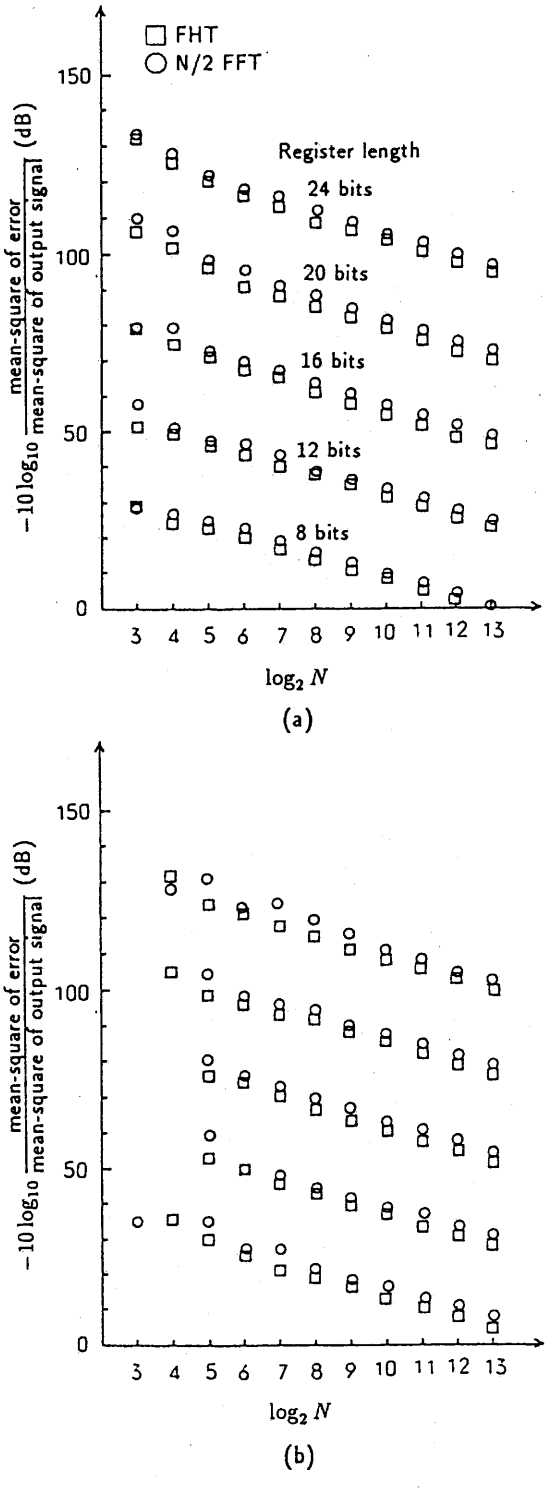

図 6.4: 各バタフライで1/2のスケーリングのある固定小数点FHT，固定小数点N/2FFTを用いたDFT計算における計算誤差　(a) 白色雑音入力，(b) 正弦波入力

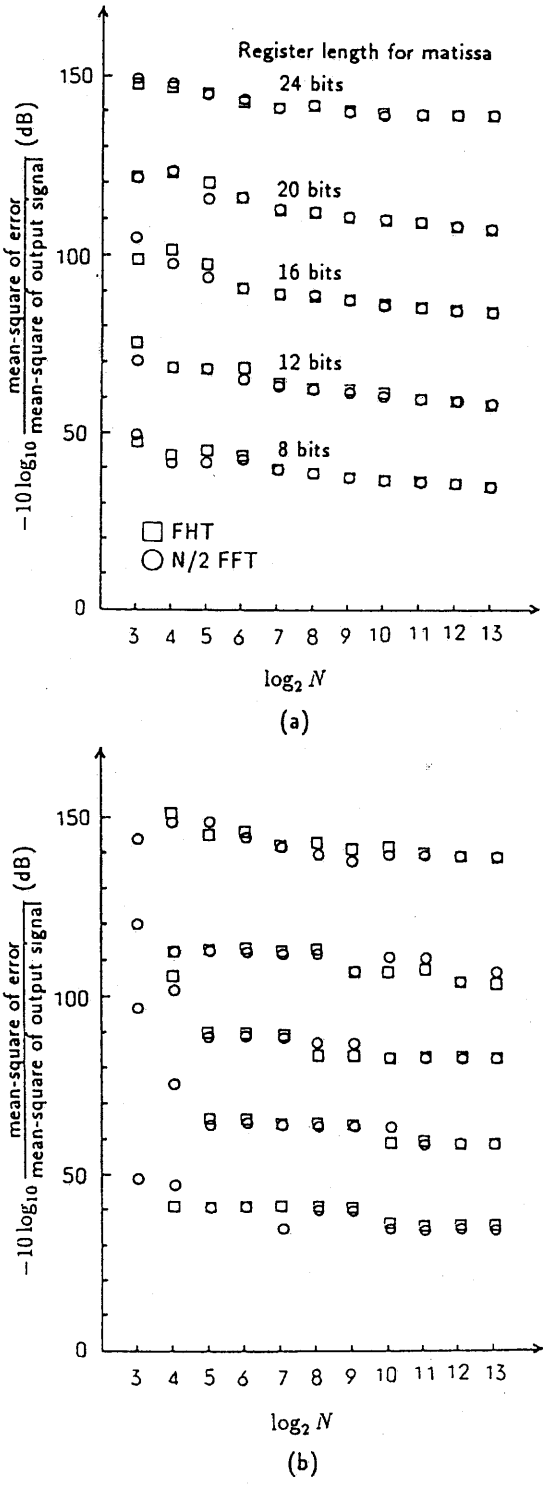

図 6.5: 浮動小数点 FHT，浮動小数点 $N/2$FFT を用いた DFT 計算における計算誤差　(a) 白色雑音入力，(b) 正弦波入力

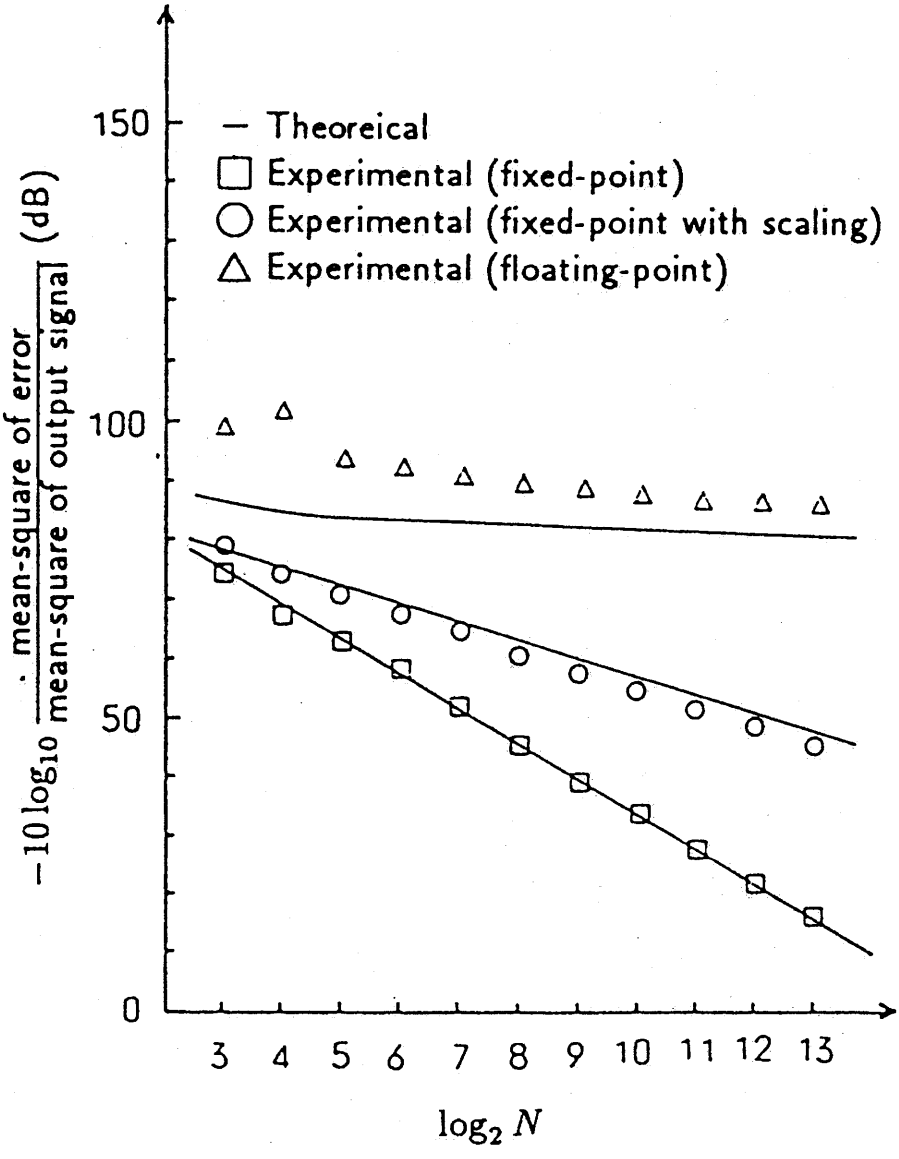

図 6.6: FHT を用いた DFT 計算における計算誤差の理論値と実測値　固定小数点レジスタ長 16 ビット，浮動小数点仮数部レジスタ長 16 ビット

て，24dB の誤差対信号比の改善が行われる，つまり 1 ビット当り 6dB の改善が行われることがわかる．一方，誤差対信号比の劣化の傾向は演算方法により異なり，固定小数点演算ＦＨＴで各段バタフライにおいてスケーリングを行わない場合はデータ数 N の二乗に比例し，各段バタフライにおいてスケーリングを行う場合にはデータ数 N そのものに比例する．浮動小数点演算ＦＨＴの場合には，誤差対信号比の劣化の傾向は，固定小数点演算の場合に比べ非常に緩やかであり，およそ $\log N$ に比例している．以上のことは，(6.44), (6.45), (6.48) 式と一致する．

白色雑音を入力とした場合と正弦波を入力とした場合を比較すると，浮動小数点演算のときの正弦波入力の場合において若干ばらつきがあるが，両者はおよそ同じ傾向をしめしており，我々が通常扱う信号の範囲では入力の性質の違いによる計算誤差の違いはそれほど大きくないと考えられる．

従来の $N/2$ 点ＦＦＴを用いた方法と比べると，固定小数点演算ＦＨＴでスケーリングなしの場合と浮動小数点演算ＦＨＴの場合には，両者の差はほとんどない．固定小数点演算ＦＨＴでスケーリングありの場合には，$N/2$ 点ＦＦＴを用いた方法の方が約 2dB ほど計算誤差が少ない．

入力信号が白色雑音で浮動小数点演算仮数部レジスタ長 16 ビットおよび固定小数点演算レジスタ長 16 ビットの場合の計算誤差の測定値と理論値を図 6.6 に示す．この図より浮動小数点演算の場合には測定値の方が計算誤差が小さくなっているが，定性的な傾向は一致している．固定小数点演算の場合には測定値と理論値はほぼ一致していることがわかる．

6.5 むすび

本章では，実数データに対するFHT（高速ハートレー変換）及び$N/2$点ＦＦＴを用いたＮ点ＦＦＴの計算方法における演算レジスタ長と計算精度の関係を明かにするため，計算誤差の統計的扱いによる理論解析をおこない理論式を導き，それを実験により検証した．計算誤差の理論値と測定値は，固定小数点演算を用いた場合にはよく一致した．浮動小数点演算を用いた場合には，その定性的な傾向は一致した．また，ＦＨＴを用いてＤＦＴを計算する方法と従来の$N/2$点ＦＦＴを用いた方法との計算精度はほぼ同じであることがわかった．

解析および実験結果から，浮動小数点演算では，2^5以上ではデータ数の増加に伴う計算誤差の増加は僅かである．一方，固定小数点演算ではスケーリングを行ったとしても，データ数の増加による計算誤差の増加は無視できないことが分かる．

例えば，5章で述べたパソコンによるリアルタイム残響抑圧システムは，仮数部24bit浮動小数点で実現されているが，これと同程度の計算精度を得るためには，式（6.45），式（6.48）から，

$$\frac{3N}{2}(1-\frac{1}{N})2^{-2b}+\frac{N}{2}2^{-2b}=\frac{2}{3}2^{-2(24-1)}(\nu+1) \tag{6.49}$$

となる．この式をbについて解くと，

$$b=21-\frac{1}{2}\log_2(\nu+1)+\frac{1}{2}\log_2 3(4N-3) \tag{6.50}$$

となる．例えば，FFTデータ数$N=2^{16}$のとき固定小数点演算に必要なレジスタ長は，式（6.50）から計算すると，$b\approx 28$となる．したがって，固定小数点演算に必要なレジスタ長は$b+1=29$bitとなる．以上，残響抑圧処理に必要な固定小数点演算のレジスタ長を見積もることができた．