Vol.2021-HPC-182 No.10
2021/12/7

少数のレゾルベントを用いるフィルタ対角化法について

村上 弘[1,a)]

概要：実対称定値一般固有値問題の近似固有対を解くために，（単一ではなくて）少数のレルベントの線形結合のチェビシェフ多項式であるフィルタを用いる．我々は既に，レゾルベントのシフトに複素数を用いることで，このようなフィルタを構築する一般的な方法を既に導出しており，良い伝達特性が得られている．しかし，固有値が固有値分布の下端付近にある固有対だけを解く場合で，レゾルベントのシフトを実数に制限したい場合には，その方法は適用できない．そこで本研究では，固有値が下端付近の固有対を解く場合について，良好な特性を持つフィルタをシフトが実数であるレゾルベントを 2 つ用いて構成する方法を示し，さらにこれらのフィルタを用いて近似固有対を解いた実験例をいくつか紹介する．

キーワード：フィルタ，対角化法，固有値問題，レゾルベント，多項式，実数シフト，伝達関数

On Filter Diagonalization Methods Which Use a Small Number of Resolvents

Abstract: We use a filter which is a Chebyshev polynomial of a linear combination of (not a single but) a small number of resolvents to solve approximate eigenpairs of real symmetric-definite generalized eigenproblems. We have already derived a general method to construct such filters with good transfer properties by using complex numbers for the shifts of resolvents. However, the method is not available when all shifts are real numbers even we desire the use of real shifts to solve only lower-exterior eigenpairs. Therefore, in this study, to solve lower-exterior eigenpairs we construct filters whose transfer properties are good by using two resolvents whose shifts are real numbers, and show by using those filters some experiments to solve approximate eigenpairs.

Keywords: filter, diagonalization, eigenproblem, resolvent, polynomial, real shift, transfer function

1. はじめに

いま実対称定値一般固有値問題 (1)（行列 A と B は実対称で，B は正定値）の固有対 $(\lambda, \mathbf{v})$ であってその固有値 λ が指定された区間 $[a, b]$ にあるものを近似して求めることにする．

$$A\,\mathbf{v} = \lambda B\,\mathbf{v} \tag{1}$$

本報告では，そのためのフィルタとして少数 k 個のレゾルベントの線形結合の作用の Chebyshev 多項式の作用 (2) を採用する．

$$\mathcal{F} = g_{\mathrm{s}} T_n(\mathcal{Y}) \tag{2}$$

[1] 東京都立大学・数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan University
[a)] mrkmhrsh@tmu.ac.jp

ここで作用素 $\mathcal{Y}$ はシフトが ρ_i のレゾルベント $\mathcal{R}(\rho_i)$, $i = 1, 2, \ldots, k$ と恒等作用素 I の線形結合であるとする（式 (3)）．

$$\mathcal{Y} = c_\infty I + \sum_{i=1}^{k} c_i \mathcal{R}(\rho_i)\,. \tag{3}$$

我々は既に文献 [29] において，レゾルベントのシフトとして複素数を使用するのであれば，レゾルベントの数を増すことで系統的にフィルタの伝達関数 $f(\lambda)$ の形状を良くできる方法を示し，その方法ではレゾルベントの数 k をいくつにしても数式に数値を入れて同じように計算することで各レゾルベントのシフトとその線形結合の係数が求められることを示した．特に，すべてのシフトが虚数になる場合は，区間 $[a, b]$ の位置を自由に設定できるので，中間固有値を持つ固有対でも困難なく求めることができる．

しかし応用上は，実対称定値一般固有値問題の求めたい

固有対はその固有値が固有値分布で端付近にあるものだけのことがよくある．そのような場合には，すべてのレゾルベントのシフト ρ_i を実数に制限すれば，フィルタを作用させる計算で必要になる記憶量と演算量を減らせる可能性が出てくる．（なぜならば，計算を複素数で行う場合には実数で行う場合に比べて必要な記憶の量は 2 倍になる．また演算については，複素数の加算は実数の加算 2 つで構成され，複素数の乗算は通常の方法では実数の乗算 4 つと実数の加算 2 つで構成される．行列分解や前進後退代入などの計算を行う場合に，加算と乗算が乗加算の形でほぼ 1 対 1 の割合で含まれているとし，また計算機の演算装置も実数の加算器と乗算器が 1 対 1 の割合で備わっていて実数の加算と乗算は演算の手間が同じと仮定すれば，計算を複素数で行う場合は実数で行う場合に比べて演算の手間は 4 倍になる．さらに今回は一般固有値問題を実対称定値の場合に限っているが，複素エルミート定値の一般固有値問題も実数の固有値だけを持ち，その場合に実対称定値の場合とほぼ同様のフィルタ構成の議論ができるが，レゾルベントを実現するための連立 1 次方程式の係数行列はシフトを実数に採れば複素エルミートになるが，シフトを虚数に採ると特別な対称性を持たない複素行列になるので，対称性を利用した行列分解の技法は使えず，シフトを実数に採る場合と比べて行列分解の計算量や分解の記憶量などが増える．）なお，レゾルベントのシフトとして実数を用いる場合には，シフトと一致もしくは近接する固有値が存在すると，必要な固有対全体についてフィルタによる伝達率の最大最小比が極端に大きくなり，精度の限られた数値と演算を用いる通常の計算では結果の精度が大きく失われるリスクがある [10]．そこでもしも必要な固有対の固有値が固有値分布の下端付近にある場合には，すべてのシフトを最小固有値よりも小さい値にとることにより，そのような一致や近接から生じるリスクを事前に回避することができる．しかしシフトの範囲を複素数から実数に制限するということは，選択の範囲を狭めていることであるから，シフトを実数から選ぶことに限定した場合に達成可能であるフィルタの伝達関数の特性は，シフトを複素数の範囲から選べる場合に比べて劣ったものになる．

我々はこれまでに，単一のレゾルベントから構成された極めて簡易なフィルタを扱ってきた（文献 [11], [12], [13], [14], [23]）．単一のレゾルベントを用いるのであれば，たとえばレゾルベントの作用を実現する連立 1 次方程式を直接法で行列分解を用いて解くと仮定するとき，分解も単一となり，複数のレゾルベントを用いるフィルタに比べて分解に掛かる計算量が少なくなる．さらに分解された行列を保持することで右辺だけが異なる連立 1 次方程式の組を解く処理を多項式の次数に等しい回数繰り返すのに掛かる計算量も少なくできるが，その分解された行列を保持するための記憶量も複数のものを用いる場合に比べて単一のレゾルベントを用いる場合には少なくなる．そのため，大規模な問題で行列分解を格納する記憶量が制約となる状況では，単一のレゾルベントを用いたフィルタには利点がある．しかし複数のレゾルベントで構成されたフィルタに比べて，単一のレゾルベントで構成されたフィルタにはその伝達関数の特性をあまり良くできないことが難点である．たとえば伝達関数の通過域での値の変動比を抑えながら遷移域の幅を狭くすることは難しい（しかもシフトを実数に制限する場合は，複素数をシフトとして選べる場合に比べてフィルタの伝達関数の特性は良くない）．しかし我々は，そのようにあまり良くない特性を持つフィルタであっても，再直交化とフィルタを組み合わせた処理を数回反復することにより，不変部分空間の基底の近似が改良されて，必要な固有対の近似精度を高めることができることを示した（文献 [22], [24], [25], [27]）．ここまでを読むと，単一のレゾルベントから構成したフィルタを用いて問題をうまく解けるのであれば，あえて複数のレゾルベントを用いて構成をする必要は無いものと思われるであろう．しかしそれでも，もしも小規模な並列度のシステム上で計算を行う場合に，もしも少数個のレゾルベントに対応する連立 1 次方程式の係数行列の分解処理の結果を主記憶に保持することができて，さらに行列分解や分解後の前進後退代入に必要な処理もそれぞれ少数で並行してほぼ独立に計算を行える場合であれば，以下のような利点が生じる可能性がある（この事情は，フィルタとして少数のレゾルベントの線形結合の実部の Chebyshev 多項式を用いた場合（文献 [26], [28], [29]）と同様である．ただしこれらの文献はシフトは実数に制限せずに複素数から選べる場合についてである）．

フィルタとして「少数」のレゾルベントの線形結合の作用の Chebyshev 多項式を用いた場合を，単一のレゾルベントの Chebyshev 多項式を用いた場合と比べると：

- Chebyshev 多項式の次数 n が減れば，（少数 k 個のレゾルベントの処理はそれぞれ並行に処理できると仮定すれば）フィルタを適用する処理の中でレゾルベントを逐次的に n 回反復して適用する部分（行列分解の後に，前進後退代入を n 回逐次的に繰り返して実現する）の経過時間が減る．
- 伝達関数の遷移域の幅が狭くなれば，フィルタを適用するベクトルの数を減らせる．
- フィルタの伝達特性の閾値を向上できて，もしもその結果としてフィルタを 1 回適用した段階で既に，近似固有対の精度が用途に対する要求を満たすことができるのなら，フィルタを反復して近似を改良する処理を省ける．

そこで固有値が固有値分布の下端付近にある固有対を近似して求める場合について，使用するフィルタはシフトが実数のレゾルベント少数の線形結合の Chebyshev 多項式とする場合について，その特性をなるべく良くする構成法の

導出を試みる．

我々は以前の文献 [15] において，フィルタをレゾルベント 2 つの線形結合の（実部の）Chebyshev 多項式として構成するのに，2 つのシフト両方を実数にする場合と虚数にする場合（ただし複数共役性を用いてシフトの虚部が正のものが 2 つ）のそれぞれについて考察を行った．その文献より後の我々の一連の研究（文献 [16], [17], [18], [19], [20], [21], [26], [28], [29]）において，アナログ電気回路におけるフィルタの設計手法を模倣することにより，フィルタの伝達関数を最良近似理論に現れる有理関数を利用する関数合成の手法で設計する方法を示した．その方法では，シフトは複素数であり，レゾルベントの数をいくつにしても数式に数値を入れて計算をすれば具体的にフィルタを決定できる．そうして少数 3〜4 個のレゾルベントを用いて優れた特性を持つフィルタが構成できることを示し，その確認のための実験も行なった [29]．しかしその方法はシフトを実数に制限する場合は使えないので，別の新しい方法が必要になる．そこで本報告ではまず（複数である数の最初は 2 つであるから），シフトが実数のレゾルベント 2 つで構成されるフィルタについて考究を行った（個数を 3 つあるいは 4 つの程度にまで範囲を拡げることは，今後の研究課題である）．

1.1 フィルタを用いた固有値問題の解法の概要

N 次の実対称定値一般固有値問題 (1) に対して，フィルタを用いて固有対を近似して求める方法の概要は以下のようになる（文献 [2], [3]）．

(1) まず，線形の作用素で，固有値が区間 $[a,b]$ にある固有ベクトルは良く通過させるが，固有値がその区間から離れた固有ベクトルは強く阻止するものをフィルタ $\mathcal{F}$ として用意する．
(2) そうしてランダムな N 次ベクトルを m 個生成し，それらを B-正規直交化して得られる m 個のベクトルの組 X を作成する．それはベクトルを列として並べた $N\times m$ 行列としても扱えて，$X^TBX = I$ である．
(3) 次にベクトルの組 X に線形作用素であるフィルタを適用して新たなベクトルの組 $Y \leftarrow \mathcal{F}X$ を作る．これもまた $N\times m$ 行列として扱える．
(4) そうして（ベクトルの組である X と Y の情報のほかにフィルタの伝達特性も考慮に入れて）Y の列ベクトルの線形結合の組をうまく構成して「区間 $[a,b]$ の近傍の固有値すべてに対応する不変部分空間」の基底を近似するベクトルの組 Z を作る．
(5) 式 (1) の一般固有値問題に対応する Rayleigh-Ritz 法を Z に適用して，得られた Ritz 対を近似固有対とする．

1.2 フィルタの概要

実対称定値一般固有値問題 (1) の固有対であって固有値が区間 $[a,b]$ にあるものを近似して求める．そのために用いるフィルタ $\mathcal{F}$ は線形の作用素であって，固有値が区間 $[a,b]$ にある固有ベクトルは良く伝達するが，固有値がその区間から離れた固有ベクトルは強く阻止するものとなるようにうまく構成する．フィルタ $\mathcal{F}$ をレゾルベントの線形結合やレゾルベントの多項式で構成した場合には，任意の固有対を $(\lambda,\mathbf{v})$ とするとき，式 (4) が成立する．ここで $f(\lambda)$ はフィルタ $\mathcal{F}$ の伝達関数と呼ばれ，λ の有理関数になる．

$$\mathcal{F}\mathbf{v} = f(\lambda)\mathbf{v} \tag{4}$$

求めたい固有対の固有値は固有値分布の下端を含む区間 $[a,b]$ にある（a が最小固有値 λ_0 以下である）場合には，その区間 $\lambda\in[a,b]$ を標準区間 $t\in[0,1]$ と対応させる線形変換 (5) により，固有値 λ に対する正規化座標 t を定義する（**図 1**）．

$$t = \frac{\lambda - a}{b - a}\,. \tag{5}$$

そうして式 (6) により引数が正規化座標 t の伝達関数 $g(t)$ を定義する．

$$g(t) \equiv f(\lambda) \tag{6}$$

伝達関数 $g(t)$ の概形の例を**図 2** に示す．伝達関数の形状についての 3 つのパラメタ μ，g_{p}，g_{s} は，伝達関数 $g(t)$ は通過域 $t\in[0,1]$ では最大値が 1 で最小値は g_{p} であり，遷移域 $t\in(1,\mu)$ においては g_{p} よりも小で g_{s} よりも大であり，阻止域 $t\in[\mu,\infty)$ では大きさ（絶対値）が g_{s} 以下であることを意味する．すると本報告では $g(t)$ が $t\in[0,\infty)$ において連続関数であることを要請するので，2 つの条件 $g(1) = g_{\mathrm{p}}$ と $g(\mu) = g_{\mathrm{s}}$ を満たすことが必要である．しかしこの図 2 のグラフのように，$g(0) = 1$ であることまでは必ずしも必要ではない．後で述べる「方式 I」の伝達関数は 2 つの条件 $g(0) = 1$ と $g'(0) = 0$ を追加し，また「方式 II」の伝達関数では 3 つの条件 $g(0) = g_{\mathrm{p}}$ と $g(t_p) = 1$ と $g'(t_p) = 0$，ただし t_p は通過域内部のある点の座標である，を追加している．）

1.3 単一のレゾルベントの多項式によるフィルタ

レゾルベントの作用を与える連立 1 次方程式を係数行列の分解を用いて直接法で解くことにするとき，フィルタを複数のレゾルベントの線形結合にする場合（文献 [2], [3], [4], [5], [6], [9]）に比べてフィルタを単一のレゾルベントの多項式にする場合には（文献 [7], [8], [11], [12], [13]），行列の分解が複数ではなくて単数になることが大きな利点である．（連立 1 次方程式を直接法ではなくて反復法を用いて解く場合にも，反復あたりの収束率向上のための前処理として係数行列の不完全

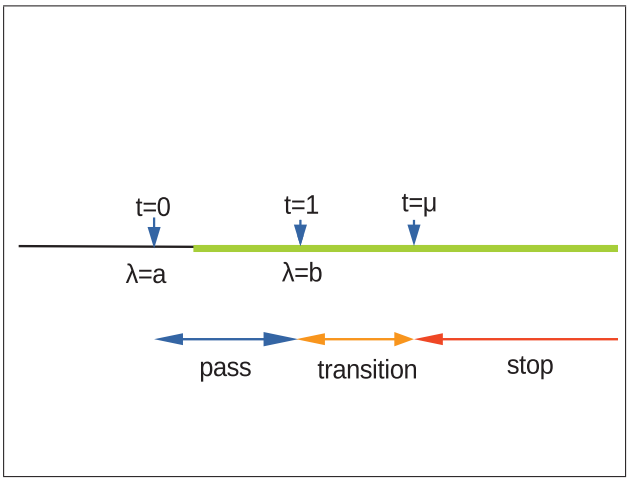

図 1 （下端固有値用）固有値 λ の区間 $[a, b]$ と正規化座標 t
通過域 $t \in [0, 1]$；遷移域 $t \in (1, \mu)$；阻止域 $t \in [\mu, \infty)$
Fig. 1 The interval $[a, b]$ of eigenvalue and normalized coordinate t (for lower-end eigenvalues).
passband $t \in [0, 1]$; transitionband $t \in (1, \mu)$; stopband $t \in [\mu, \infty)$.

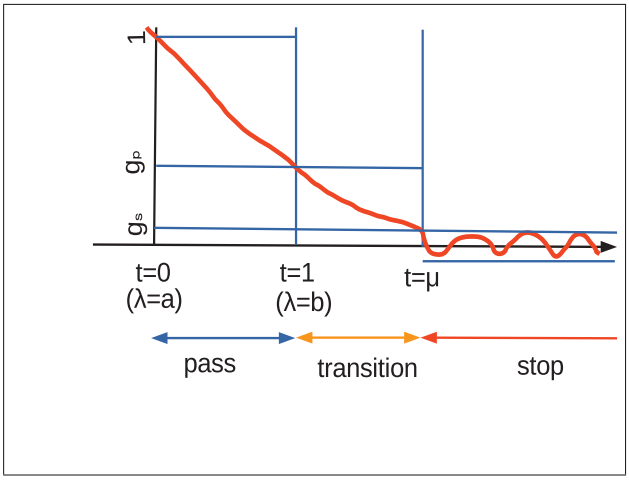

図 2 （下端固有値用）伝達関数 $g(t)$ の概形の例
Fig. 2 Conceptual shape of transfer function $g(t)$ (for lower-end eigenvalues).

分解を用いるのであれば，複数ではなくて単一の行列を不完全分解することになる）．レゾルベントの作用の n 次多項式をベクトル m 個の組に対して適用するには，係数行列が共通で右辺ベクトルだけが異なる m 通りの連立 1 次方程式の組を解く操作を逐次的に n 回だけ繰り返す必要がある．その際に，係数行列の分解を最初に 1 度行ってその結果を保持することができれば，その後の毎回の連立 1 次方程式の組を解く処理では行列分解を新たにせずに保持してある行列分解の結果を利用してベクトルの組に対する前進後退代入だけで計算を行える．

フィルタを単一のレゾルベントの「多項式」として構成する場合には，最小 2 乗法に類似する方法で（通過域と阻止域の両方で折り合いをみながら）フィルタの伝達特性をなるべく都合良くするようにうまく「多項式」を最適化して決めることが望ましいと考えられる（文献 [7], [8], [11]）．しかし，数値的に最適化した結果である「多項式」の係数は数値を列挙したものとしてだけ得られる．Chebyshev 多項式を採用した「簡易な設計法」（文献 [12]）で得られるフィルタでは，数値的な最適化による「多項式」の調整をせずに阻止域における良い減衰特性を容易に実現できるが，他方で通過域における伝達率の最大最小比を抑える機能は持たない．通過域における伝達率の最大最小比が大きいと，必要とする固有対相互の近似精度のばらつきも大きい可能性があり，極端な場合には必要な固有対の幾つかが欠落してしまう可能性もある．そこで我々は以前に，フィルタは単一のレゾルベントの多項式とするが，その多項式として n 次 Chebyshev 多項式を用いる従来の簡易設計型の方法を拡張して，多項式を n 次以下の複数の Chebyshev 多項式の線形結合で表す方法を試みた．その方法では最小 2 乗法の定式化により，多項式の線形結合の係数を通過域における伝達率の最大最小比を抑えるように決める（文献 [13], [14]）．その場合の係数は数値最適化の結果であって数値を列挙して与えられるものになる．そのように拡張した方法は考察の段階であり，まだ実験実証を行っていない．

1.4 少数のレゾルベントの多項式によるフィルタ

本報告では，固有値が下端付近の固有対を求めるためのフィルタとして，実数をシフトとするレゾルベントを 2 つ用いた簡易型，つまりそれらレゾルベント 2 つの線形結合の Chebyshev 多項式であるものを扱う．これは実数をシフトとするレゾルベントの数を 1 つから 2 つに増やすことにより，フィルタの伝達関数の自由度を増して，通過域における最大最小比を小さくすることあるいは遷移域の幅を狭くすることを狙ったものである．

これまでのところ，シフトが実数であるレゾルベントを単数ではなくて複数用いた簡易型，すなわちレゾルベントの線形結合の Chebyshev 多項式としてうまくフィルタの構成を示すことができたのはレゾルベントが 2 つの場合だけである．レゾルベントを少数の範囲で（たとえば 3 つ，4 つ程度に）増やした場合の簡易型のフィルタの構成法の導出については，今後の課題である．

フィルタを「レゾルベントの線形結合」（あるいはその実部）の多項式ではなくて，「レゾルベントの線形結合」（あるいはその実部）の形そのものを用いる方法もある（文献 [1], [2], [10]）．しかしその場合にはフィルタによる不要なベクトルに対する減衰能力を多項式を用いて強化しないので，多項式を用いる場合に比べてレゾルベントを多く用いる必要がある．特にレゾルベントのシフトを実数に制限して，固有値の存在する区間の中にシフトを配置する場合には，フィルタの伝達関数は固有値の存在する区間のなかに極を持ち有界ではないので，シフトと一致もしくは近接する固有値があれば，それに対応する固有ベクトルがフィルタでろ過して得られたそれぞれのベクトルに強く拡大さ

れて含まれ，それ以外の必要な固有値を持つ固有ベクトルは覆い隠されるために情報の精度が落ちる．そのことを事後に検出して対処する方法が提案されている [10]．たとえば固有値と一致や近接したシフトを持つレゾルベントだけを除外して線形結合を取り直してフィルタの構成を修正する方法である．しかしそれは，用いるレゾルベントの数が 1 個や数個のようなごく少数ではなくて 10 個や 20 個のような多数である場合を想定している．本報告では，レゾルベントに対応する連立 1 次方程式は直接法で行列分解を用いて解くと仮定してして，まずは行列分解に要する計算量を減らすためにフィルタを構成するレゾルベントの数をそれほど多くしないで少数とする場合について考えている．そうして，固有値がシフトと一致もしくは近接することをあらかじめ避けるために，下端付近の固有値を持つ固有対を求める場合に問題を限定して，採用するレゾルベントのシフトは最小固有値よりも小さい実数にとることにしている．そうしてさらに少数のレゾルベントの線形結合だけでは必要な固有値を含む区間から離れた固有値を持つ固有ベクトルに対する十分な減衰能力が得られないので，能力を強化するために線形結合の Chebyshev 多項式として構成されたフィルタにしている．

2. 極が実数 2 つの簡易型の伝達関数の設計法

フィルタが相異なる実数をシフトに持つ 2 つのレゾルベントの多項式であるとき，その伝達関数は相異なる 2 つの実数だけを（多重の）極として持つ有理関数になる．以下では，その関数形を制限した簡易型の設計法についてだけ考察する．

これまで Chebyshev 多項式を用いた簡易型の設計法では，単一の実数の極だけを持つ伝達関数 $g(t)$ を式 (7) の形に制限してきた（単一の極は実数 $-\sigma$ で，n 位である）．

$$g(t) \equiv g_{\mathrm{s}} T_n(y(t)),\ y(t) \equiv 2x(t) - 1,\ x(t) \equiv \frac{\mu + \sigma}{t + \sigma}\,. \tag{7}$$

このような形に制限された伝達関数は，（n を増せば）阻止域における伝達率を容易に微小にできるが，通過域における伝達率の最大最小比を小さくすることは（遷移域の幅 $\mu - 1$ を大きくしなければ）できない．そこでこの簡易設計の手法を極が単一の実数だけである場合から相異なる 2 つの実数だけの場合に拡張し，それにより増えた自由度を利用して，通過域における伝達率の最大最小比を小さくすることを試みる．

そのためには，新たに式 (8) で表される実有理関数 $x(t)$ を採用する．これは無限遠での値が零であり，相異なる負の実数 2 つだけを極として持つ．以下では，これら相異なる 2 つの極の符号を逆にした値をそれぞれ実数 σ_1 と σ_2 にしており，$\sigma_1 > \sigma_2 > 0$ であるとする．

$$x(t) \equiv \frac{\alpha_1}{t + \sigma_1} - \frac{\alpha_2}{t + \sigma_2}\,. \tag{8}$$

いま $\mu > 1$ であるとする．式 (8) の関数 $x(t)$ は $t \geq 0$ では連続である．そうして x_H と x_L は $x_H > x_L > 1$ を満たす実数の未知数として，阻止域 $t \in [\mu, \infty)$ では $x(t)$ は値が 1 以下で正であり，遷移域 $t \in (1, \mu)$ では $x(t)$ は値が x_L 未満で 1 よりも大きいとし，通過域 $t \in [0, 1]$ では $x(t)$ は最大値が x_H で最小値は x_L であると仮定する．すると関数 $x(t)$ の連続性から，$x(\mu) = 1$ と $x(1) = x_L$ であることが必要である．そうして伝達関数 $g(t)$ は $x(t)$ の多項式に Chebyshev 多項式を用いた簡易型であるとして，これまでと同様に式 (9) を採用する．

$$g(t) \equiv g_{\mathrm{s}} T_n(y(t)),\ y(t) \equiv 2x(t) - 1\,. \tag{9}$$

すると阻止域において $x(t)$ の値は 1 以下で正と仮定したので，阻止域において式 (9) の $g(t)$ の大きさ $|g(t)|$ は g_{s} 以下になる．さらに通過域 $t \in [0, 1]$ における $g(t)$ の最大値と最小値がそれぞれ 1 と g_{p} であると仮定すると，$x_H > x_L > 1$ より $2x_H - 1 > 2x_L - 1 > 1$ であること，Chebyshev 多項式は引数が 1 以上で単調増加性を持つこと，通過域において $x(t)$ の最大値は x_H で最小値は x_L と仮定したこと，を併せると式 (10) の 2 つの関係が得られる．

$$\begin{cases} 1 \ = g_{\mathrm{s}} T_n(2x_H - 1)\,, \\ g_{\mathrm{p}} = g_{\mathrm{s}} T_n(2x_L - 1)\,. \end{cases} \tag{10}$$

さらに z が実数の場合の恒等式 (11) を用いると，式 (10) を双曲線関数を用いて表した式 (12) が得られる．

$$\cosh^{-1}(2z^2 - 1) = 2\cosh^{-1}|z| \tag{11}$$

$$\begin{cases} x_H = \cosh^2\Big(\dfrac{1}{2n}\cosh^{-1}\dfrac{1}{g_{\mathrm{s}}}\Big)\,, \\ x_L \ = \cosh^2\Big(\dfrac{1}{2n}\cosh^{-1}\dfrac{g_{\mathrm{p}}}{g_{\mathrm{s}}}\Big)\,. \end{cases} \tag{12}$$

すると 3 つのパラメタ g_{p}，g_{s}，n の組が指定された場合に，式 (12) の中の各式の右辺をそれぞれ計算すれば，$x(t)$ の通過域 $t \in [0, 1]$ における最大値 x_H と最小値 x_L が決まる．

4 つのパラメタ μ，g_{p}，g_{s}，n が指定されたときに，式 (8) の $x(t)$ に含まれている 4 つの値である σ_1，σ_2，α_1，α_2（ただし $\sigma_1 > \sigma_2 > 0$）がすべて実数としてうまくとれることが伝達関数 $g(t)$ およびフィルタの構成には必要であり，そうでなければ構成は不可能である．

2.1 「方式 I」：通過域の左端で最大かつ停留になる伝達関数

いま通過域 $[0, 1]$ における伝達率の最大最小比を小さくすることを狙って，伝達関数は通過域の左端である原点 $t = 0$ において最大値 1 をとり，しかもそこで値が停留する（微分の値が零になる）という条件を課して構成してみる．（伝達関数が原点 $t = 0$ において最大でかつ平坦な特

性を持てば，固有値が指定区間の下側にある固有ベクトルほど良く通過するフィルタとなり，得られる固有値の精度は固有値が下側のものであるほど良くなることが期待できる．必要な固有対の固有値が下側から順に少数だけの場合には，通常そのような性質は最も望ましいものであろう．）

いま 4 つのパラメタの値の組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ を指定するとき，それに対応する伝達関数が実現可能である場合には，以下に示す手順により式 (8) の $x(t)$ に含まれる 4 つの値 σ_1，σ_2，α_1，α_2 が決定できる．

まず式 (12) により，x_L と x_H の値を求めておく．そうして $t=\mu$ における値の条件 $x(\mu)=1$，$t=1$ における値の条件 $x(1)=x_L$，さらに通過域の左端 $t=0$ における値の条件 $x(0)=x_H$ とそこで停留になるという条件 $\left.\frac{d}{dt}x(t)\right|_{t=0}=0$ の全部で 4 つを順番に並べて数式で表すと，式 (13) になる．

$$\left\{\begin{array}{lcl} 1 &=& \dfrac{\alpha_1}{\mu+\sigma_1}-\dfrac{\alpha_2}{\mu+\sigma_2},\\ x_L &=& \dfrac{\alpha_1}{1+\sigma_1}-\dfrac{\alpha_2}{1+\sigma_2},\\ x_H &=& \dfrac{\alpha_1}{\sigma_1}-\dfrac{\alpha_2}{\sigma_2},\\ 0 &=& -\dfrac{\alpha_1}{\sigma_1^2}+\dfrac{\alpha_2}{\sigma_2^2}\,. \end{array}\right. \tag{13}$$

既知である 3 つの値 x_L，x_H，μ を用いてこの連立方程式 (13) を解いて 4 つの未知数 σ_1，α_1，σ_2，α_2 を実数の範囲で求める．そのような解が存在すれば，指定されたパラメタの 4 つ組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ に対応する「方式 I」の伝達関数 $g(t)$ は実現可能であり，そうでなければ実現不可能である．

4 つの未知数を実数の範囲で求めるための具体的な手順の導出の記述は長くなるのでここでは省略して（付録 §A.4.1 に移動），以下には結果だけを述べる．

2.1.1 「方式 I」の場合の設計法

4 つのパラメタ $(\mu, g_{\rm p}, g_{\rm s}, n)$ が与えられたとき，以下の式 (14) を順に計算する（ここで $x'_H = x_H - 1$ の意味である）．

$$\left\{\begin{array}{lcl} x_H &\leftarrow& \cosh^2\Big(\dfrac{1}{2n}\cosh^{-1}\dfrac{1}{g_{\rm s}}\Big),\\ x'_H &\leftarrow& \sinh^2\Big(\dfrac{1}{2n}\cosh^{-1}\dfrac{1}{g_{\rm s}}\Big),\\ x_L &\leftarrow& \cosh^2\Big(\dfrac{1}{2n}\cosh^{-1}\dfrac{g_{\rm p}}{g_{\rm s}}\Big),\\ p &\leftarrow& \dfrac{x_H}{x'_H}\times\mu^2,\\ q &\leftarrow& \dfrac{x_H}{x_H-x_L},\\ S_1 &\leftarrow& \dfrac{p-q}{\mu-1}-(\mu+1),\\ S_2 &\leftarrow& \mu+\dfrac{\mu\,q-p}{\mu-1},\\ D &\leftarrow& S_1^2-4S_2. \end{array}\right. \tag{14}$$

そうして，3 つの条件 $S_1>0$，$S_2>0$，$D>0$ をすべて満たしていれば構成は可能であり，そうでなければ不可能である．そうして，構成が可能な場合には，以下の式 (15) を上から順に計算する（なお $\sigma_1-\sigma_2=\sqrt{D}$ であることを用いている）．

$$\left\{\begin{array}{lcl} \sigma_1 &\leftarrow& \dfrac{1}{2}\,(S_1+\sqrt{D}\,),\\ \sigma_2 &\leftarrow& \dfrac{S_2}{\sigma_1},\\ C &\leftarrow& \dfrac{x_H}{\sqrt{D}},\\ \alpha_1 &\leftarrow& C\sigma_1^2,\\ \alpha_2 &\leftarrow& C\sigma_2^2. \end{array}\right. \tag{15}$$

このようにして式 (8) の $x(t)$ が含む実数値の組 $(\sigma_1, \alpha_1, \sigma_2, \alpha_2)$ を決定できる．ここで $\sigma_1 > \sigma_2 > 0$ であり，さらに $\alpha_1 > \alpha_2 > 0$ になることも示せる．

2.2 「方式 II」：通過域の両端で値の等しい伝達関数

「方式 II」では通過域における伝達関数の最大最小比を小さくすることを容易にする手段として，通過域の両端で伝達関数の値が等しいという条件 $g(0)=g(1)=g_{\rm p}$ を課すことにする．そうして伝達関数は通過域の内部のある 1 点 $t=t_p$（$0<t_p<1$）において最大値 1 をとるとする．

簡易設計による伝達関数の関数形は「方式 I」の場合と同じく (8) と (9) で与えられるとする（$\sigma_1>\sigma_2>0$ も仮定する）．そうしてパラメタの値の 4 つ組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ を指定して，その組を持つ伝達関数が実現可能であれば，式 (8) の $x(t)$ が含む 4 つの実数値 σ_1，σ_2，α_1，α_2 を以下の手順で具体的に決定できる．

まず前節 2.1 と同様に，通過域における $x(t)$ の最小値 x_L と最大値 x_H の値はそれぞれ，3 つのパラメタ $g_{\rm p}$，$g_{\rm s}$，n の値から式 (12) を計算して求める．

すると $x(t)$ の満たすべき条件は $x(0)=x_L$ と $x(1)=x_L$ と $x(\mu)=1$，それと最大点における条件 $x(t_p)=x_H$ と $\left.\frac{d}{dt}x(t)\right|_{t=t_p}=0$ の全部で 5 つであり，それらの条件を上から順に並べて式 (16) が得られる．

$$\left\{\begin{array}{lcl} x_L &=& \dfrac{\alpha_1}{\sigma_1}-\dfrac{\alpha_2}{\sigma_2},\\ x_L &=& \dfrac{\alpha_1}{1+\sigma_1}-\dfrac{\alpha_2}{1+\sigma_2},\\ 1 &=& \dfrac{\alpha_1}{\mu+\sigma_1}-\dfrac{\alpha_2}{\mu+\sigma_2},\\ x_H &=& \dfrac{\alpha_1}{t_p+\sigma_1}-\dfrac{\alpha_2}{t_p+\sigma_2},\\ 0 &=& -\dfrac{\alpha_1}{(t_p+\sigma_1)^2}+\dfrac{\alpha_2}{(t_p+\sigma_2)^2}\,. \end{array}\right. \tag{16}$$

3 つの値 x_L，x_H，μ を与えてこの連立方程式 (16) を解いて，5 つの未知数 σ_1，α_1，σ_2，α_2，t_p を実数の範囲で求める．それが可能である場合は，指定されたパラメタの 4 つ組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ に対応する「方式 II」の伝達関数 $g(t)$ は実現可能であり，そうでなければ不可能である．

4 つの未知数を実数の範囲で求めるための具体的な手順の導出の記述はかなり長くなるのでここでは省略して（付録 §A.4.2 に移動），以下に結果だけを述べる．

2.2.1 「方式 II」の場合の設計法

4 つのパラメタ $(\mu, g_{\rm p}, g_{\rm s}, n)$ が与えられたとき，以下の式 (17) を上から順に計算する．（ここで $x'_H = x_H - 1$, $x'_L = x_L - 1$ の意味である．）

$$\left\{\begin{array}{ll} x_H & \leftarrow \cosh^2\big(\frac{1}{2n}\cosh^{-1}\frac{1}{g_{\rm s}}\big), \\ x'_H & \leftarrow \sinh^2\big(\frac{1}{2n}\cosh^{-1}\frac{1}{g_{\rm s}}\big), \\ x_L & \leftarrow \cosh^2\big(\frac{1}{2n}\cosh^{-1}\frac{g_{\rm p}}{g_{\rm s}}\big), \\ x'_L & \leftarrow \sinh^2\big(\frac{1}{2n}\cosh^{-1}\frac{g_{\rm p}}{g_{\rm s}}\big), \\ \kappa & \leftarrow \dfrac{\mu}{\mu-1}, \\ \zeta_0 & \leftarrow 1 - \dfrac{x_L}{x'_L}\times\dfrac{x'_H}{x_H}\times\kappa, \\ \zeta_1 & \leftarrow -2\times\dfrac{\kappa}{x'_L}\times\dfrac{x_H - x_L}{x_H}, \\ \zeta_2 & \leftarrow (\kappa - \dfrac{x_L}{x'_L}\times\dfrac{x'_H}{x_H})\,\kappa, \\ D_1 & \leftarrow \zeta_1^2 - 4\zeta_0\zeta_2\,. \end{array}\right. \tag{17}$$

そうして，$\zeta_2 \le 0$ であれば構成不能として終了し，$\zeta_2 > 0$ の場合にはさらに次の式 (18) を上から順に計算する．

$$\left\{\begin{array}{ll} S_2 & \leftarrow \dfrac{2\,\zeta_2}{-\zeta_1 + \sqrt{D_1}}, \\ S_1 & \leftarrow (1+S_2)\sqrt{\dfrac{x_L}{x_H}}, \\ D_2 & \leftarrow S_1^2 - 4S_2\,. \end{array}\right. \tag{18}$$

次に，2 次方程式 $w^2 - S_1 w + S_2 = 0$ が相異なる正の実数（$z_1 > z_2 > 0$）を持つための必要十分条件「$S_1 > 0$ かつ $S_2 > 0$ かつ $D_2 \equiv S_1^2 - 4S_2 > 0$」を調べて，満たしていなければ構成は不可能なので終了する．満たしている場合は，式 (19) を用いて 2 次方程式の相異なる 2 つの正の実数解 z_1 と z_2 の値（$z_1 > z_2 > 0$）を計算する．

$$\left\{\begin{array}{ll} z_1 & \leftarrow \dfrac{1}{2}\,(S_1 + \sqrt{D_2})\,, \\ z_2 & \leftarrow \dfrac{S_2}{z_1}\,. \end{array}\right. \tag{19}$$

このとき $z_1 \ge 1$ であれば（条件 $1 > z_1 > z_2 > 0$ を満たせないので），構成は不可能であり終了する．それ以外の場合には構成は可能であり，式 (20) を上から順に計算する．（ここでは，$\sigma_1 - \sigma_2 = (1+\sigma_1)(1+\sigma_2)S_1\sqrt{D_2}$ となることを用いている．）

$$\left\{\begin{array}{ll} \sigma_1 & \leftarrow \dfrac{z_1^2}{(1-z_1)(1+z_1)}, \\ \sigma_2 & \leftarrow \dfrac{z_2^2}{(1-z_2)(1+z_2)}, \\ C & \leftarrow \dfrac{x_L}{(1+\sigma_1)(1+\sigma_2)S_1\sqrt{D_2}}, \\ \alpha_1 & \leftarrow C\sigma_1(1+\sigma_1)\,, \\ \alpha_2 & \leftarrow C\sigma_2(1+\sigma_2)\,. \end{array}\right. \tag{20}$$

以上の手順により，構成が可能な場合には，式 (8) の $x(t)$ が含む値の組 $(\sigma_1, \alpha_1, \sigma_2, \alpha_2)$ を決定できる．ここで $\sigma_1 > \sigma_2 > 0$ であり，$\alpha_1 > \alpha_2 > 0$ も示せる．

2.3 「方式 I」と「方式 II」で構成された関数 $g(t)$ の振る舞いの確認

まず式 (8) の $x(t)$ は「方式 I」と「方式 II」のどちらの場合にも阻止域 $t \in [\mu, \infty)$ において $1 \ge x(t) > 0$ を満たすことを確認する（この条件を満たせば，式 (9) の伝達関数の大きさ $|g(t)|$ は阻止域において常に $g_{\rm s}$ 以下となる）．そのためには具体的に構成された $x(t)$ は，「方式 I」と「方式 II」どちらの場合も $\alpha_1 > \alpha_2 > 0$ を満たすことを利用する．

いま式 (8) の変形から式 (21) が得られる．

$$x(t) = \frac{(\alpha_1 - \alpha_2)t + (\alpha_1\sigma_2 - \alpha_2\sigma_1)}{(t+\sigma_1)(t+\sigma_2)}\,. \tag{21}$$

いま $x(0) \ge x_L > 0$ であることを式 (21) にあてはめてみると，σ_1 と σ_2 は正なので，不等式 (22) がなりたつことがわかる．

$$\alpha_1\sigma_2 - \alpha_2\sigma_1 > 0\,. \tag{22}$$

式 (21) とこの不等式 (22)，および σ_1 と σ_2 が共に正であり，また $\alpha_1 > \alpha_2$ であることを使うと，$t \ge 0$ であるときには $x(t) > 0$ であることがわかる．

次に $x(t)$ は「方式 I」と「方式 II」どちらの場合にも，阻止域 $[\mu, \infty)$ で単調減少であることを示すために，$x(t)$ の導関数の式 (23) について調べてみる．

$$\begin{aligned} x'(t) &= -\frac{\alpha_1}{(t+\sigma_1)^2} + \frac{\alpha_2}{(t+\sigma_2)^2} \\ &= -\frac{Q(t)}{(t+\sigma_1)^2(t+\sigma_2)^2}\,. \end{aligned} \tag{23}$$

ここで $Q(t)$ は式 (24) で与えられる t の 2 次式多項式である．

$$Q(t) = (\alpha_1 - \alpha_2)t^2 + 2(\alpha_1\sigma_2 - \alpha_2\sigma_1)t + \alpha_1\sigma_2^2 - \alpha_2\sigma_1^2\,. \tag{24}$$

この 2 次方程式 $Q(t) = 0$ の判別式を求めてみると $4\alpha_1\alpha_2(\sigma_1 - \sigma_2)^2$ となり，その値は「方式 I」と「方式 II」どちらの場合にも必ず正になるので，2 次多項式 $Q(t)$ は必ず相異なる 2 つの実数 t_1 と t_2 を零点として持つことがわかる（その大小関係を $t_1 < t_2$ と決める）．いま 2 次方程式の解と係数の関係から式 (25) が導かれる．

$$t_1 + t_2 = -2 \times \frac{\alpha_1\sigma_2 - \alpha_2\sigma_2}{\alpha_1 - \alpha_2}\,. \tag{25}$$

この式 (25) の値は，「方式 I」と「方式 II」どちらの場合にも負であることが，式 (22) と $\alpha_1 > \alpha_2$ であることからわかる．よって少なくとも t_1 と t_2 のうちのどちらかは値が負である．すると $t_1 < t_2$ であると決めておいたので，t_1

の値は必ず負である．

- 「方式 I」ではそれを構成したときの条件から $x'(0) = 0$ である．つまり $t = 0$ は $Q(t)$ の零点である．すると（t_1 の値は負であることから）$t_2 = 0$ である．主係数が正である 2 次式 $Q(t)$ の値は $t > t_2 = 0$ のときには正であるから，式 (23) により $t > 0$ のときには必ず $x'(t) < 0$ である．よって $x(t)$ は $t > 0$ で単調減少である．
- 「方式 II」ではその構成から $x'(t_p) = 0$（$t_p \in (0, 1)$）である．つまり $t = t_p$ が $Q(t)$ の零点である．すると（t_1 の値は負であるから）$t_2 = t_p$ であることがわかる．主係数が正である 2 次式 $Q(t)$ の値は $t > t_2 = t_p$ のときには正であるから，式 (23) により $t > t_p$ のときには必ず $x'(t) < 0$ となる．よって $x(t)$ は $t_p < t$ のときには単調減少である（同様に $x(t)$ は $0 \le t < t_p$ のときには単調増加であることが示せる）．

さらに，$x(t)$ を構成したときの条件から $x(\mu) = 1$ であり，また式 (8) から $\lim_{t\to\infty} x(t) = 0$ である．よって上記のこととあわせると，「方式 I」と「方式 II」どちらの場合にも，$x(t)$ は阻止域 $t \in [\mu, \infty)$ において単調減少で $(0, 1]$ への全射である

以上の結果をまとめると，「方式 I」と「方式 II」どちらの場合にも阻止域 $t \in [\mu, \infty)$ では $|g(t)| \le g_\mathrm{s}$ であり，さらに遷移域 $t \in (1, \mu)$ では $g(t)$ は単調減少で $[g_\mathrm{s}, g_\mathrm{p}]$ への全射である．そうして，通過域 $t \in [0, 1]$ では「方式 I」の $g(t)$ は単調減少であるが「方式 II」の $g(t)$ は $0 \le t < t_p$ では単調増加で，$t_p < t \le 1$ では単調減少であり，$g(t)$ はどちらの方式でも通過域では $[g_\mathrm{p}, 1]$ への全射である．

2.4 「方式 I」と「方式 II」の設計の関係性

いま通過域における「方式 II」の伝達関数 $g(t)$ の最大点を $t = t_p$ とするとき，「方式 II」の本来の通過域 $[0, 1]$ から $[0, t_p)$ の部分を取り除いた区間である $t \in [t_p, 1]$ の全体を $\widetilde{t} \in [0, 1]$ の全体に写す線形の座標変換 $t = \mathcal{L}(\widetilde{t})$ により $\widetilde{g}(\widetilde{t}) = g(\mathcal{L}(t))$ としたものは「方式 I」の伝達関数とみなせるので，「方式 I」と「方式 II」の 3 つのパラメタ g_p と g_s と次数 n を共通に設定した場合に，「方式 II」において伝達関数の遷移域 $(1, \mu)$ の幅と通過域 $[0, 1]$ の幅の比が $\mu - 1$ であるならば，「方式 I」の場合に対応する比は $\widetilde{\mu} - 1 = \dfrac{\mu - 1}{1 - t_p}$ である．「方式 II」の場合の比の値である $\mu - 1$ と「方式 I」の場合の比の値である $\widetilde{\mu} - 1$，その前者に対する後者の比の値を計算すると $\dfrac{\mu - 1}{\widetilde{\mu} - 1} = 1 - t_p$ であり，1 よりも小さい．

以上のことから，3 つのパラメタ g_p と g_s と次数 n を共通にとるとき，「方式 I」よりも「方式 II」の方がフィルタを実現できる μ の値を小さくできることがわかる．

2.5 パラメタ 4 つのうちの 3 つだけを指定するやり方

フィルタを実現可能なパラメタの 4 つ μ，g_p，g_s，n をすべて直接指定する以外の方法としては，たとえば以下に述べるように，3 つの値だけを直接指定して，残りの 1 つについては探索によりフィルタが実現可能な範囲でなるべく都合の良い値となるように決めることができる．

次数 n を最小にする場合

実用性の観点から探索する次数の上限 n_max をあらかじめ設定して（たとえば 50 とする）次数 n を 1 から始めて n_max まで 1 ずつ増してフィルタを実現可能にする最初のもの（n が最小のもの）を探す．探しても無ければ指定した 3 つのパラメタを持つフィルタは（n が上限 n_max 以下の範囲では）実現不可能である．

g_p を最大にする場合

まず定義から $1 > g_\mathrm{p} > g_\mathrm{s}$ である．g_p の最大値はたとえば 2 分法で精密に決めることもできるが，簡易には本当の最大値ではなくてきりの良い値，たとえば単調に減少する 0.5^j の形に制限してその指数 j を 1 ずつ増やして（$g_\mathrm{p} > g_\mathrm{s}$ の範囲で）フィルタを実現可能にする最初のもの（g_p が大きいもの）を探す．探しても無ければ指定した 3 つのパラメタを持つフィルタの構成を諦める．

g_s を最小にする場合

まず定義から $g_\mathrm{p} > g_\mathrm{s} > 0$ である．g_s の最小値はたとえば 2 分法で精密に決めることもできるが，簡易にはきりの良い値，たとえば単調に減少する 0.5^j の形に制限して，$g_\mathrm{p} > g_\mathrm{s}$ を満たしてさらに現実的な考慮からたとえば丸め誤差の単位を ϵ_MAC とするときに $g_\mathrm{s} > \epsilon_\mathrm{MAC}$ の範囲で j の値を 1 ずつ増やして，フィルタを実現可能にする最後のもの（g_s が小さいもの）を探す．探しても無ければ指定した 3 つのパラメタを持つフィルタの構成を諦める．

μ を最小にする場合

まず定義から $\mu > 1$ である．μ の最小値はたとえば 2 分法で精密に決めることもできるが，簡易にはきりの良い値に制限して，たとえば単調に増加する $1 + 0.05\,j$ の形とし，整数 j を 1 から始めて 1 ずつ増して探索することができる．ただし実用性を考えて μ の値にはある上限を設けて（たとえば $\mu \le 2$）それ以下の範囲でだけ探索をするなどとする．そうしてフィルタを実現可能にする最初のもの（μ が小さいもの）を探す．探しても無ければ指定した 3 つのパラメタを持つフィルタの構成を諦める．

2.6 伝達関数からのフィルタの構成

極として実数 2 つだけを持つ簡易型の伝達関数に対応するフィルタは，シフトが実数であるレゾルベントを 2 つ用いて構成できる（実対称定値一般固有値問題 (1) に対応するシフトが ρ のレゾルベント $\mathcal{R}(\rho)$ は $\mathcal{R}(\rho) \equiv (A - \rho B)^{-1} B$ である）．

いま採用する簡易型の設計では，伝達関数 $g(t)$ は式 (8)

と (9) により与えられる．下端付近の固有値を扱うので，$\lambda \in [a,b]$ と $t \in [0,1]$ の間の線形対応関係は式 (5) で与えられる．すると y を λ の関数として表すと，以下の式 (26) になる．

$$y = \frac{2\ell_1}{\lambda-\rho_1} - \frac{2\ell_2}{\lambda-\rho_2} - 1\,. \tag{26}$$

ここでシフト ρ_k と係数 ℓ_k（$k=1,2$）は式 (27) により与えられる．

$$\rho_k \equiv a-(b-a)\,\sigma_k\,,\quad \ell_k \equiv (b-a)\,\alpha_k \tag{27}$$

いま $\sigma_1 > \sigma_2 > 0$ であることから $\rho_1 < \rho_2 < a$ である．さらに「方式 I」と「方式 II」の場合には $\alpha_1 > \alpha_2 > 0$ になることから，$\ell_1 > \ell_2 > 0$ となることもわかる．

式 (26) の y に対応する線形作用素 $\mathcal{Y}$ は，$\dfrac{1}{\lambda-\rho_k}$ にはレゾルベント $\mathcal{R}(\rho_k)$ を，定数 1 には恒等作用素 I を，それぞれ対応させた式 (28) になる．

$$\mathcal{Y} \equiv 2\ell_1\,\mathcal{R}(\rho_1) - 2\ell_2\,\mathcal{R}(\rho_2) - I \tag{28}$$

そうして伝達関数 $f(\lambda)$ に対応する作用素であるフィルタ $\mathcal{F}$ は，作用素 $\mathcal{Y}$ の多項式として式 (2) で表せる．

ベクトルの組 V に式 (2) の形のフィルタ $\mathcal{F}$ を作用させる計算には，Chebyshev 多項式のもつ 3 項漸化式 $T_0(z)=I$，$T_1(z)=z$，$T_j(z)=2zT_{j-1}(z)-T_{j-2}(z)$（$j\geq 2$）を利用する．具体的には，$\mathcal{Y}$ の j 次 Chebyshev 多項式 $T_j(\mathcal{Y})$ を V に作用させて得られるベクトルの組 $V^{(j)} \equiv T_j(\mathcal{Y})V$ を以下の漸化式 (29) を用いて計算する．

$$\begin{cases} V^{(0)} = V\,, \\ V^{(1)} = \mathcal{Y}\,V\,, \\ V^{(j)} = 2\mathcal{Y}\,V^{(j-1)} - V^{(j-2)} \quad (j\geq 2)\,. \end{cases} \tag{29}$$

すると V から始めて漸化式 (29) により $V^{(n)}$ を求めれば，ベクトルの組 V に式 (2) のフィルタ $\mathcal{F}$ を作用させた結果であるベクトルの組は式 (30) の右辺で与えられる．

$$\mathcal{F}\,V = g_{\rm s}\,V^{(n)} \tag{30}$$

2.7 実数シフトの単一のレゾルベントによるフィルタ（拡張版）の構成

シフトが実数である単一のレゾルベントから Chebyshev 多項式を用いて簡易構成されたフィルタの伝達関数 $g(t)$ として，これまでは式 (7) で表されるものを用いてきた．

今回はここでそれを少し拡張して，式 (31) で表される伝達関数を新たに導入する（これは β の値を -1 に制限した場合には従来のものに帰着する）．

$$g(t) = g_{\rm s}T_n(y(t)),\ \ y(t) = \frac{\alpha}{t+\sigma} + \beta. \tag{31}$$

ただし $\sigma > 0$，$\alpha > 0$ であり，$y(t)$ は非負領域 $t \in [0,\infty)$ において連続かつ単調減少になる．そうしてこの伝達関数の形状のパラメタ $(\mu, g_{\rm p}, g_{\rm s})$（ただし $\mu > 1$，$1 > g_{\rm p} > g_{\rm s} > 0$ である）についての条件として，形状パラメタ μ，$g_{\rm p}$，$g_{\rm s}$ のそれぞれの意味に基づいてこれまでと同じように $g(0)=1$，$g(1)=g_{\rm p}$，$g(\mu)=g_{\rm s}$ であること，および $t\in[\mu,\infty)$ において $g_{\rm s} \geq |g(t)|$ であることを要請する．

すると 4 つのパラメタの組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ を指定したときに，式 (31) の形の伝達関数 $g(t)$ の実現可能性の判定，そうして実現可能であるときに $x(t)$ を決定する 3 つの実数 σ，α，β の値を与える手順は以下のようにまとめられる．

まず，式 (32) を上から順に計算する．

$$\begin{cases} y_H \leftarrow \cosh\big(\dfrac{1}{n}\cosh^{-1}\dfrac{1}{g_{\rm s}}\big), \\ y_L \leftarrow \cosh\big(\dfrac{1}{n}\cosh^{-1}\dfrac{g_{\rm p}}{g_{\rm s}}\big), \\ \sigma \leftarrow \dfrac{(y_L-1)\mu}{(y_H-y_L)\mu-(y_H-1)}, \\ \alpha \leftarrow (y_H-y_L)\sigma\times(\sigma+1), \\ \beta \leftarrow y_L-(y_H-y_L)\sigma. \end{cases} \tag{32}$$

これにより得られた結果が $\sigma > 0$ かつ $\beta \geq -1$ であるときに限って要請を満たすフィルタは実現が可能である（なぜならば $\sigma > 0$ は $y(t)$ が非負領域で定数でない連続関数であることの必要十分条件であり，さらに，$1 > g_{\rm p}$ であることから $y_H > y_L$ であり，$y(0)=y_H$，$y(1)=y_L$ であるから $y(t)$ は非負領域において単調減少関数となり，さらに $y(\mu)=1$ と決めるので，阻止域に於いて $g_{\rm s} \geq |g(t)|$ となるための必要十分条件は $\beta \geq -1$ である）．

以上の手順により，4 つのパラメタの組 $(\mu, g_{\rm p}, g_{\rm s}, n)$ を指定したときに，実数をシフトとする単一のレゾルベントを用いたフィルタでその伝達関数 $g(t)$ が式 (31) で表されるものの実現可能性は判定ができて，実現可能である場合には与えられたベクトルに対してフィルタの作用を与える具体的な手続きを構成できる．また，レゾルベントを 2 つ使う場合と同様に，4 つのパラメタのうちの 3 つの値だけを直接指定して，残り 1 つの値はフィルタを実現可能とする範囲で探索を行い，フィルタの性質が最も望ましくなるように決めることもできる．

式 (31) の有理関数 $y(t)$ を決定する 3 つの実数である σ，α，β の値が求まれば，求めたい固有対の固有値が固有値分布の下端の区間 $[a,b]$ に含まれるものであるとするとき，その区間を通過域とするフィルタ $\mathcal{F}$ は式 (33) で与えられる．

$$\begin{cases} \mathcal{F} \equiv g_{\rm s}T_n(\mathcal{Y}),\ \ \mathcal{Y} \equiv \gamma\mathcal{R}(\rho)+\beta I, \\ \rho = a-(b-a)\sigma,\ \gamma = (b-a)\alpha. \end{cases} \tag{33}$$

3. 実験例

固有値問題の固有対を上述の各種フィルタを用いて近似する実験の例は図や表が多いため，本報告では付録 A.1, A.2, A.3 に移して記述する．

4. おわりに

式 (1) の実対称定値一般固有値問題の固有値が固有値分布の下端付近の固有対を近似して求めるためのフィルタを，シフトが実数のレゾルベント 2 つの線形結合の Chebyshev 多項式として構成する方法を示し，それを簡単な例題に適用してある程度うまく機能することを確認した．

本報告で提案している「方式 I」のフィルタの伝達関数の形状は $t = 0$ で平坦な Butterworth 特性のものであるが，小さい固有値の固有ベクトルほどフィルタを良く通過するので，得られる近似固有対も固有値が小さいものほど精度が良くなる傾向を持つ．この性質は，固有値が固有値分布の下端付近にある固有対だけが必要で，しかも固有値が下側にあるものほど高い精度が必要な用途の場合には適切であるといえる．他方で「方式 II」のフィルタは「方式 I」のものに比べて伝達率の通過域での一様性の向上（最大最小比 $1/g_{\mathrm{p}}$ の低減）を狙ったものであり，フィルタを 4 つのパラメタ μ，g_{p}，g_{s}，n の値の組で指定するとき，「方式 I」の場合に比べて実現可能なパラメタの選択範囲が広くなる．たとえば 4 つのうちで μ 以外の 3 つのパラメタの値が共通である場合には，「方式 II」の方が「方式 I」に比べて μ の値を小さくできる．しかし「方式 II」のフィルタの伝達率は通過域をなるべく広く確保するために，通過域における伝達率の最小値である g_{p} を通過域の両端でとるように構成しているので，固有値が通過域の両端に近い固有対ほど近似が悪くなる傾向を持つ．「方式 II」のフィルタを用いて固有対を近似する場合のこのような性質は，有限要素法などの変分的な離散化で導かれた固有値問題の場合のように，固有値が下端付近にある固有対だけが重要で，しかも固有値が下側にあるものほど高い精度で求めたい場合にはあまり望ましいものではないであろう．

本報告の研究は，計算資源が限られている状況で，必要な記憶量と計算量をなるべく減らして大規模な実対称定値一般固有値問題を解くことを想定するものである．すべての固有対を求めるのではなくて，固有値が限られた範囲にある比較的少数の固有対だけを解くのであるが，しばしば必要とする固有対は固有値が下端付近のものだけであり，その場合には，レゾルベントのシフトとして固有値の存在範囲にない実数を選ぶことは容易にできる．しかもシフトをすべて実数にすれば計算をすべて実数の範囲で行える．それゆえシフトとしては実数だけを用いることにする．実数であるシフトを最小固有値よりも小さい値にとれば，レゾルベントに対応する連立 1 次方程式の係数行列は実対称正定値になり，連立 1 次方程式を直接法で解く場合には行列分解では対称性を保ったままピボットの選択をしなくても計算は常に安定で，係数が帯行列であれば帯幅を保って計算が行える．レゾルベントをたとえば数個程度の極めて少ない数に限ることで，大規模な問題において特に行列分解の計算とその分解結果を主記憶に置いて前進後退代入で用いることが容易になる．フィルタを構成するレゾルベントの数が少なければ実現可能なフィルタの特性を良くできず，それを用いて得られる近似固有対の精度も良くない．しかしシフトが実数のレゾルベントを 1 つだけ用いて必要な計算が行えれば，レゾルベントを用いる解法としては最も省資源なものになりうる．しかし得られる近似固有対の精度は良くない．解の精度は用途に依っては数桁程度で十分なこともありうるが，高い精度が要求される場合には，たとえば直交化処理と組み合わせてフィルタの操作を反復すれば，伝達関数の通過域での一様性が悪くてもあるいは阻止域での大きさがそれほど微小ではなくても近似固有対の精度改良が進むので，直交化付きフィルタの反復による近似固有対の精度改良はとても強力である．他方で本報告の研究は，シフトが実数に制限されたレゾルベントの数を 1 つから 2 つに増やすことにより，簡易構成型のフィルタの伝達特性をどれだけ良くできるであろうか，というものである．

なお今後も，シフトが実数であるレゾルベントを 3 つから 4 つ用いた場合の簡易構成型のフィルタをうまく構成する方法についての検討を進めるつもりである．それに向けての試みの 1 つとして，重心表示型の有理関数補間を利用する方法を付録の節 A.6 に紹介した．

付　　録

A.1 実験について

A.1.1 例題に用いた一般固有値問題

実験の例題に用いた実対称定値一般固有値問題 (1) は，各辺が座標軸に沿った 1 辺の長さ π の 3 次元立方体の内部を領域として，その表面において零ディリクレ境界条件を課したときの（符号反対の）3 次元ラプラシアン $-\Delta$ の固有値問題，それを有限要素法（FEM）で離散化近似して得られるものである．

FEM の要素分割は立方体領域の各辺方向をそれぞれ N_1+1，N_2+1，N_3+1 の等間隔の小区間に分割したもので，要素内の展開基底関数には各辺方向の 3 重線形関数を用いた．この FEM の離散化で得られる行列 A と B の次数は $N=N_1N_2N_3$ となり，（$N_1 \le N_2 \le N_3$ であるとして）行列の帯幅を小さくするように基底関数に適切に番号を付けると，各行列の（対角を含まない）半帯幅（下帯幅）は $w_L = 1+N_1+N_1N_2$ である．

いまの場合に，辺に沿った 3 方向の 1 次元 FEM の展開基底である区分線形関数はそれぞれ N_1 個，N_2 個，N_3 個あり，3 方向の区分線形関数に対してその頂上の位置の増加順につけた順番をそれぞれ i_1，i_2，i_3 とする（$1 \le i_k \le N_k$，$k=1,2,3$）．3 次元 FEM の展開基底である 3 重線形関数は各方向の展開基底の積なので 3 重添字 (i_1,i_2,i_3) により指定できる．3 重線形関数の 3 重添字に対しては $i_1+N_1(i_2-1)+N_1N_2(i_3-1)$ の値により順番を付けて，それに基づいて 3 次元 FEM の係数行列 A や B を組み立てている．

このようにして得られた実対称定値一般固有値問題 (1) に対して，フィルタ対角化法を適用して，固有値 λ が区間 $[a,b]$ に含まれる固有対を近似して求めた．このテスト例題の固有値は簡単な数式で表せるので，値を式に入れて計算すれば容易に固有値の厳密値が求まる．区間 $[a,b]$ に固有値が入る固有対の正しい数も固有値の厳密値を列挙して値の大小順に並べて数えれば求まる．

A.1.2 近似固有対の品質評価に用いた相対残差

計算により求めた各近似固有対の品質の評価には相対残差を用いた．近似固有対 $(\lambda,\mathbf{v})$ に対する相対残差 Θ を式 (A.1) で定義する．ただしベクトルのノルム $||\cdot||$ には 2-ノルムを用いた．

$$\Theta \equiv \frac{||\,A\mathbf{v}-\lambda B\mathbf{v}\,||}{||\,\lambda B\mathbf{v}\,||}\,. \tag{A.1}$$

この Θ の値はベクトル $\mathbf{v}$ の規格化にはよらないし，また共通の非零の値で行列 A と B をスケールしても不変である．またベクトルのノルムとして 2-ノルムを用いたので，幾何学的には，N 次元ユークリッド空間内で 2 つのベクトル $A\mathbf{v}$ と $\lambda B\mathbf{v}$ が挟む角の大きさを ϕ とするとき，不等式 (A.2) が成り立つ．

$$\sin\phi \le \Theta\,. \tag{A.2}$$

複数の近似固有対のベクトルを列としてまとめて並べた行列を V とし，それから行列 A と B の対称性や帯性または疎性を利用しながら行列積 AV と BV を求め，それから複数の相対残差を作る．このように行列積計算の形にすることで，A と B への記憶参照は全部で 1 回ずつになり，複数の相対残差を個別に計算するよりも効率良く計算できる．

A.1.3 直交化付きフィルタの反復による近似固有対の改良

計量 B の正規直交化とフィルタの適用を組み合わせたものを少数（`IT_MAX`）回反復することで，フィルタの伝達特性の形状の悪さを補って，不変部分空間の基底の近似を改善できる [20], [27]．その処理の概要は以下のようになる．

(1) 通過域が $[a,b]$ であるフィルタ $\mathcal{F}$ を用意
(2) 乱数から作成した m 個のベクトルの組を Y とする．
(3) for `IT`=1, `IT_MAX` do
(4)　$X \leftarrow$「Y の切断付き B-正規直交化」
(5)　$Y \leftarrow \mathcal{F}X$
(6) enddo
(7) X と Y とフィルタの特性を考慮して，Y の線形結合で不変部分空間の基底の近似 Z を作成 [3].
(8) 式 (1) の一般固有値問題に対応する Rayleigh-Ritz 法を Z に適用して，得られた Ritz 対を近似固有対とする．

上記処理中の (4)「Y の切断付き B-正規直交化」の処理において Y の実効階数の低下を検出したら，X の持つベクトルの数を減らす（上記処理中の次の (5) で Y の持つベクトルの数も X と同じになる）．

A.1.4 実験に用いた計算機システム

実験例の計算に用いたシステムは，東京大学情報基盤センターの Oakbridge-CX の 1 ノード（CPU は Dual で Intel Xeon 8280（2.7GHz,28cores），共有メモリは 192GiB，倍精度のピーク演算性能は 4.8TFLOPS）である．

プログラムのソースコードの記述には Fortran90 を用いて，並列化のための OpenMP の指示行を適宜追加した．計算に用いた数値と演算は IEEE 754 の倍精度浮動小数点（2 進，64bit）である．コンパイラは intel fortran（version 19.0.5.281）で，コンパイラのオプションとして "`-fast -qopenmp -xCORE-AVX512 -align array64byte`"を指定した．

A.2 実験その 1

以下の各例では有限要素法（FEM）で用いる直方体への要素分割を $(N_1, N_2, N_3) = (40, 50, 60)$ とした．それにより得られる式 (1) の実対称定値一般固有値問題は行列 A と B が帯行列で，行列次数が $N = 120,000$，下帯幅が $w_L = 2,041$ となる．

この一般固有値問題の最小固有値は，変分原理により要素分割によらず常に 3 より大きいので（いまの場合は有効数字 5 桁では 3.0010 である），固有値が区間 $[a, b] = [3, 30]$ にある固有対を近似して求めることにした．固有値がその区間にある固有対の数は 54 である．

フィルタを適用する最初のベクトルの数 m は，$\mu = 2.0$ の場合には $m = 200$，$\mu = 1.5$ の場合には $m = 125$，$\mu = 1.25$ の場合には $m = 100$ とした（伝達関数 $g(t)$ の通過域 $t \in [0, 1]$ と遷移域 $t \in (1, \mu)$ を併せた区間 $t \in [0, \mu)$ に対応する固有値の区間 $\lambda \in [a, b')$ は，$\mu = 2.0$ の場合には $[3, 57)$，$\mu = 1.5$ の場合には $[3, 43.5)$，$\mu = 1.25$ の場合には $[3, 36.75)$ であり，そうしてそれぞれの場合に，区間 $[a, b')$ に固有値がある固有対の数は 163 個，105 個，78 個であり，それぞれの場合に対して m は少し大きくとっている）．

計算で得られた近似固有対で固有値が区間 $[a, b]$ にあるものについて，式 (A.1) で定義した「相対残差」を計算した．

A.2.1 「方式 I」の伝達関数の例

伝達関数 $g(t)$ の形状パラメタ 3 つと次数の組（μ，$g_{\rm p}$，$g_{\rm s}$，n）を指定して，それを満たす「方式 I」の伝達関数を決定した．対応するフィルタの作用は伝達関数から容易に導ける（副節 2.6）．

実験に用いた「方式 I」の 6 通りのフィルタ（I-1，I-2，I-3，I-4，I-5，I-6）について，指定した 4 つのパラメタ μ，$g_{\rm p}$，$g_{\rm s}$，n の値をそれぞれ**表 A·1** に示す（各例において，次数 n 以外の 3 つのパラメタ μ，$g_{\rm p}$，$g_{\rm s}$ を先に指定して，n の値は「方式 I」のフィルタが構成可能な最小のものにしている）．そうして各例について，指定した 4 つのパラメタから決定された式 (8) の σ_k，α_k，$k = 1, 2$ の値をそれぞれ**表 A·2** に示す．

「方式 I」の 6 通りのフィルタについて，正規化座標 t を横軸にとり，伝達関数の大きさ $|g(t)|$ の常用対数を縦軸にとってプロットしたグラフをそれぞれ示す（**図 A·1**，**図 A·2**，**図 A·3**，**図 A·4**，**図 A·5**，**図 A·6**）．

A.2.2 「方式 II」の伝達関数の例

伝達関数 $g(t)$ の形状パラメタ 3 つと次数の組（μ，$g_{\rm p}$，$g_{\rm s}$，n）を指定して，それを満たす「方式 II」の伝達関数を決定した．対応するフィルタの作用は伝達関数から容易に導ける（副節 2.6）．

実験に用いた 6 通りのフィルタ（II-1，II-2，II-3，II-4，II-5，II-6）について，指定した 4 つのパラメタ μ，$g_{\rm p}$，$g_{\rm s}$，n の値をそれぞれ**表 A·5** に示す（各例において，次数 n 以外の 3 つのパラメタ μ，$g_{\rm p}$，$g_{\rm s}$ を先に指定して，n の値は「方式 II」のフィルタが構成可能な最小のものにしている）．そうして各例について，指定した 4 つのパラメタから決定された式 (8) の σ_k，α_k，$k = 1, 2$ の値をそれぞれ**表 A·6** に示す．

「方式 II」の 6 通りのフィルタについて，正規化座標 t を横軸にとり，伝達関数の大きさ $|g(t)|$ の常用対数を縦軸にとってプロットしたグラフをそれぞれ示す（**図 A·13**，**図 A·14**，**図 A·15**，**図 A·16**，**図 A·17**，**図 A·18**）．

A.2.3 各フィルタによる計算結果

「方式 I」の 6 通りのフィルタそれぞれについて，B-正規直交化と組み合わせて 3 回まで反復して求めた近似固有対の固有値を横軸にとり，その相対残差の常用対数を縦軸にとってプロットしたグラフを示す（**図 A·7**，**図 A·8**，**図 A·9**，**図 A·10**，**図 A·11**，**図 A·12**）．

同様に，「方式 II」の 6 通りのフィルタそれぞれについて，B-正規直交化と組み合わせて 3 回まで反復して求めた近似固有対の固有値を横軸にとり，その相対残差の常用対数を縦軸にとってプロットしたグラフを示す（**図 A·19**，**図 A·20**，**図 A·21**，**図 A·22**，**図 A·23**，**図 A·24**）．

フィルタの反復回数に対する近似固有対の相対残差の最大値を「方式 I」のフィルタの各例について**表 A·3** に，同様に「方式 II」のフィルタの各例について**表 A·7** に示す．またフィルタの反復回数に対する対角化までの経過時間を「方式 I」のフィルタの各例について**表 A·4** に，同様に「方式 II」のフィルタの各例について**表 A·8** に示す．どの場合もシフト行列 2 つを分解するのに要した経過時間は約 14 秒であった．反復あたりの経過時間は，フィルタの次数 n が高いほど，ベクトルの数 m が多いほど増える（ただしフィルタを反復すると途中でベクトルの数は一般にはベクトルの組の階数低下により減少する）．

扱った問題の規模は 1 ノードのシステムで計算できる最大のものではなく，示した経過時間についても並列化に対する努力は最善のものではない．またこれら各例では，行列 2 つの分解処理や分解結果 2 つを用いた前進後退代入の処理も，並行して行なうことは可能であるが，実験の計算では順次に行っている．

A.3 実験その 2：レゾルベントが 1 つと 2 つのフィルタの比較

まず比較の便利のために，シフトが実数であるレゾルベントを 1 つ用いたフィルタと 2 つ用いたものとが同じ基準で設定できるようにするため，シフトが実数である単一のレゾルベントを使うフィルタの構成法を従来のものから少しだけ拡張する．そうして，単一のレゾルベントを使う場

合である「単一」と，レゾルベントを 2 つ用いる場合の 2 つの方式である「方式 I」と「方式 II」について，若干の比較実験を行なう．

例題とする式 (1) の実対称定値一般固有値問題は，再び以前と同じ有限要素法から導かれたものであり，FEM の要素分割を $(N1, N2, N3) = (40, 50, 60)$ として，固有値が区間 $[a, b] = [3, 30]$ に含まれる全部で 54 の固有対を求めた．

経過時間を一応示したが，レゾルベントを 2 つ用いる場合には，2 つ分のレゾルベントの構成の準備や 2 つのレゾルベントをベクトルの組に作用させる処理は本来なら並行して計算ができるが，そのようにはせずに 2 つを順次に処理している．また OpenMP によるスレッド並列化も最善の努力のものではない．

A.3.1 実数シフトの単一のレゾルベントで構成されたフィルタの実験

これは，式 (33) で表されるシフトが実数である単一のレゾルベントの Chebyshev 多項式をフィルタに用いた場合の実験の例である．

シフトが実数である単一のレゾルベントを用いたこのフィルタで，指定する μ の値をそれぞれ 2.0，1.5，1.25 とした場合の計算結果を**表 A·9**，**表 A·10**，**表 A·11** に示す．これらすべての場合にパラメタ g_{s} の値は `1E-13` と指定している．表中では，次数 n は 10 から 40 まで 5 刻みで変え，直交化付きフィルタの反復回数を 1 から 3 までとして，近似固有対の相対残差の最大値を各場合について示している．このように 3 つのパラメタ μ，n，g_{s} については値を直接指定し，残りのパラメタ g_{p} の値は 0.5^j の形（j は正の整数）に制限して，そのうちでフィルタを実現可能とする場合の最大値に設定した（表 A·9，表 A·10，表 A·11 の中に実際に用いた g_{p} の値を有効数字 3 桁に丸めたものを載せている）．

そうして，最初の段階においてフィルタに適用するベクトルの数 m は，μ の値を 2.0，1.5，1.25 とした各場合についてそれぞれ 200，125，100 と指定している．

μ の値を 2.0，1.5，1.25 とした各場合について，今回の方式の単一のレゾルベントを用いたフィルタで，近似固有対を求めるために掛かった経過時間をそれぞれ**表 A·21**，**表 A·22**，**表 A·23** に示す．

μ の値を 2.0，1.5，1.25 とした各場合について，単一のレゾルベントを用いたフィルタの適用回数が 1 から 3 まで（IT1，IT2，IT3）について横軸にフィルタの次数 n をとり，縦軸に相対残差の最大値の常用対数の値をとってそれぞれ赤，緑，青の線を用いてプロットしたグラフを**図 A·25**，**図 A·26**，**図 A·27** に示す．

μ の値が 2.0，1.5，1.25 の各場合について，横軸にはフィルタの次数 n をとり，縦軸には単一のレゾルベントで構成されたフィルタを 1 回適用して得られた近似固有対の相対残差の最大値の常用対数をとってプロットしたグラフを**図 A·34** に示す．

A.3.2 実数シフトの 2 つのレゾルベントから構成された「方式 I」と「方式 II」のフィルタの実験

これは実数シフトのレゾルベント 2 つの線形結合の Chebyshev 多項式をフィルタとした場合の実験の例である．

「方式 I」のフィルタで μ の値を 2.0，1.5，1.25 とした各場合の表がそれぞれ**表 A·12**，**表 A·13**，**表 A·14** であり，同様に「方式 II」のフィルタについてはそれぞれ**表 A·15**，**表 A·16**，**表 A·17** である．簡単のためにパラメタ g_{s} の値はどの場合にも `1E-13` とした．各表の中では，次数 n を 10 から 40 まで 5 刻みで変えて，直交化付きフィルタの反復回数を 1 から 3 まで変えて，得られた近似固有対の相対残差の最大値を示している．このように 3 つのパラメタ μ，n，g_{s} のについては値を直接指定をして，残り 1 つのパラメタ g_{p} はその値を 0.5^j（j は正の整数）の形に制限して，そのなかからフィルタを実現可能とする最大のものを選択した．この探索でフィルタが実現不可能な場合には表中で横線を引いて示している（表 A·12，表 A·13，表 A·14，表 A·15，表 A·16，表 A·17 の中に実際に採択された g_{p} の値を有効数字 3 桁に丸めたものを載せている）．

μ の値を 2.0，1.5，1.25 とした各場合について，最初にフィルタに適用するベクトルの数 m はそれぞれ 200，125，100 とした．

「方式 I」で μ の値を 2.0，1.5，1.25 とした各場合の経過時間をそれぞれ**表 A·24**，**表 A·25**，**表 A·26** に示し，同様に「方式 II」についてはそれぞれ**表 A·27**，**表 A·28**，**表 A·29** に示す．フィルタに適用するベクトルの数 m が同じである場合には，「方式 I」も「方式 II」も行列の分解や連立 1 次方程式の組を解く計算の作業量は同じになるので，経過時間もほぼ同じになっている．ただし，フィルタ操作と組み合わせた B-正規直交化により検出されるベクトルの組の実効階数の低下の状況が違うと，反復の 2 回目以降からは作業の対象となるベクトルの数が同じにはならずに経過時間の違いを生む可能性がある．

μ の値を 2.0，1.5，1.25 とした各場合について，レゾルベントを 2 つ用いる「方式 I」のフィルタの適用回数が 1 から 3 まで（IT1，IT2，IT3）について横軸にフィルタの次数 n をとり，縦軸に相対残差の最大値の常用対数の値をとってそれぞれ赤，緑，青の線を用いてプロットしたグラフを**図 A·28**，**図 A·29**，**図 A·30** に示す．$\mu = 2.0$ のときは次数 n が 15 以上の場合には近似固有対の精度の改良はフィルタの反復 2 回目で終了している．そうして $\mu = 1.5$ のときは次数 n が 25 以上の場合には反復 2 回目で終了している．さらに $\mu = 1.25$ のときも反復 3 回目で改良はほぼ終了している．

同様に，μ の値を 2.0，1.5，1.25 とした各場合について，レゾルベントを 2 つ用いる「方式 II」のフィルタの適用回数が 1 から 3 まで（IT1，IT2，IT3）について横軸にフィルタの次数 n をとり，縦軸に相対残差の最大値の常用対数の値をとってそれぞれ赤，緑，青の線を用いてプロットしたグラフを**図 A·31**，**図 A·32**，**図 A·33** に示す．これも $\mu = 2.0$ のときには次数 n が 15 以上の場合には近似固有対の精度の改良はフィルタの反復 2 回目で終了している．そうして $\mu = 1.5$ のときは次数 n が 20 以上の場合には反復 2 回目で終了している．さらに $\mu = 1.25$ のときも反復 3 回目で改良はほぼ終了している．

μ の値を 2.0，1.5，1.25 とした各場合について，横軸にはフィルタの次数 n をとり，縦軸には「方式 I」のフィルタの適用 1 回で得られた近似固有対の相対残差の最大値の常用対数をとってプロットしたグラフを**図 A·35** に示す．同様に，縦軸に「方式 II」のフィルタの適用 1 回で得られた近似固有対の相対残差の最大値の常用対数をとってプロットしたグラフを**図 A·36** に示す．

A.3.3 実験その 2 の結果

μ の値を 2.0，1.5，1.25 とした各場合の g_{p} の値をフィルタの次数 n とフィルタの種類（「単一」、「方式 I」、「方式 II」）について集めたものを，それぞれ**表 A·18**，**表 A·19**，**表 A·20** に掲げる．いまの場合 g_{s} の値はどの場合にも `1E-13` である．通過域におけるフィルタの伝達率の最大最小比は $1/g_{\mathrm{p}}$ であるから，g_{p} の値が大きいほど，固有対の近似精度の一様性の向上が期待できる．これらの 3 つの表から，μ の値とフィルタの次数 n が同じであるときには，「単一」，「方式 I」，「方式 II」の順に g_{p} の値が大きくなっているつまりフィルタの伝達特性を μ，g_{p}，g_{s} の 3 つで特徴付ける場合には，「単一」，「方式 I」，「方式 II」の後のものほど良くなっていることがわかる．

実験その 2 で，フィルタを 1 回適用して得られた近似固有対の相対残差の最大値を比較したグラフ（図 A·34，図 A·35，図 A·36）からは，単一のレゾルベントで構成したフィルタは，2 つのレゾルベントで構成した「方式 I」や「方式 II」のフィルタに比べて，多項式の次数 n を増したときの相対残差の最大値の減少傾向が緩やかである．このことはフィルタの次数 n を増したときに，通過域における伝達率の最小値 g_{p} として選んだ値（「単一」の場合は表 A·9，表 A·10，表 A·11 で「方式 I」の場合は表 A·12，表 A·13，表 A·14 で「方式 II」の場合は表 A·15，表 A·16，表 A·17）が増加している傾向と概ね符合する．どのフィルタの場合についても，μ の値を小さくするほど（伝達関数は遷移域で急峻な変化を強いられてその結果として通過域での伝達関数の最大最小比が大きくなるので），近似固有対の残差の最大値は増加している．固有値が通過域に隣接する遷移域にある不要な固有対が多く存在する場合には，フィルタの種類を固定したときに，パラメタ μ を小さくとると遷移域が狭くなり，より少ない数 m の初期ベクトルをフィルタに適用すればよくなるので，フィルタの中で n 回繰り返してレゾルベントを適用する際のベクトルの数 m に比例する記憶参照や演算量を減らせるという利点があるが，フィルタの特性が悪くなる（今の場合は g_{p} が小さくなる）ので，計算で得られる近似固有対の精度が低下する可能性がある．レゾルベント 2 つで構成されたフィルタの方が 1 つで構成されたものよりも小さい μ の値が指定できる．そうして「方式 2」は「方式 1」に比べて他の 3 つのパラメタを揃えた場合にはより小さい μ の値を指定することができる．

実験その 2 で，フィルタを 1 回だけ適用した場合の近似固有対の相対残差の最大値を比較したグラフ（**図 A·37**，**図 A·38**，**図 A·39**）からは，単一のレゾルベントで構成されたフィルタを用いた場合（赤線）よりも，2 つのレゾルベントから構成された「方式 I」（緑線）と「方式 II」（青線）のフィルタを用いた場合の方が近似固有対の精度が少し高いことがわかる．そうして，「方式 I」と「方式 II」を比較すると，「方式 I」の方が精度が高い．これは「方式 II」では区間 $[3, 30]$ の下端に近い固有値 3.00102667 を持つ固有対が伝達率が低いために精度も低くなり，相対残差の最大値での比較で不利になっている．

A.3.3.1 フィルタ反復との併用の考察

$\mu = 2.0$ の場合（$m = 200$）

- フィルタが「単一」の場合
  表 A·9 と表 A·21 からたとえば次数 $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 `7.4E-06` よりも，次数 $n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 `8.7E-11` の方が精度が 5 桁程度高いことがわかる．ただし経過時間は前者が 56.1 秒に対して後者は（途中に B-正規直交化が 1 回追加されるので）65.2 秒で，少し長くなっている．
  同様に，次数 $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 `3.7E-06` よりも，次数 $n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 `5.6E-13` の方が精度が 7 桁程度高い．そうして経過時間は前者が 73.8 秒に対して後者は 82.9 秒である．
- フィルタが「方式 I」の場合
  表 A·12 と表 A·24 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 `8.6E-07` よりも，$n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 `8.9E-11` の方が精度が 4 桁高い．そうして経過時間は前者が 85.3 秒に対して後者は 95.3 秒である．
  同様に，$n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 `2.3E-07` よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 `1.2E-12` の方が精度が 5 桁程度高い．そうして経過時間は前者が

114.9 秒に対して後者は 124.6 秒である．

- フィルタが「方式 II」の場合
  表 A·15 と表 A·27 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 6.5E-06 よりも，$n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 2.9E-10 の方が精度が 4 桁程度高い．そうして経過時間は前者が 87.0 秒に対して後者は 96.0 秒である．
  同様に $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 2.5E-07 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 3.1E-12 の方が精度が 5 桁程度高い．そうして経過時間は前者が 115.0 秒に対して後者は 125.2 秒である．

$\mu = 1.5$ **の場合**（$m = 125$）

- フィルタが「単一」の場合
  表 A·10 と表 A·22 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 3.2E-04 よりも，$n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 2.3E-07 の方が精度が 3 桁程度高い．そうして経過時間は前者が 43.5 秒に対して後者は 48.0 秒である．
  同様に $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 2.1E-04 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 9.2E-10 の方が精度が 5 桁程度高い．そうして経過時間は前者が 57.4 秒に対して後者は 61.6 秒である．
- フィルタが「方式 I」の場合
  表 A·13 と表 A·25 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 9.2E-05 よりも，$n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 6.7E-08 の方が精度が 3 桁程度高い．そうして経過時間は前者が 68.6 秒に対して後者は 73.1 秒である．
  同様に $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 8.0E-06 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 3.1E-10 の方が精度が 4 桁程度高い．そうして経過時間は前者が 92.6 秒に対して後者は 95.7 秒である．
- フィルタが「方式 II」の場合
  表 A·16 と表 A·28 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 1.2E-04 よりも，$n = 10$ のフィルタを 2 回反復した場合の相対残差の最大値 2.1E-07 の方が精度が 3 桁程度高い．そうして経過時間は前者が 69.4 秒に対して後者は 72.5 秒である．
  同様に $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 1.3E-05 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 5.2E-10 の方が精度が 4 桁程度高い．そうして経過時間は前者が 92.4 秒に対して後者は 97.0 秒である．

$\mu = 1.25$ **の場合**（$m = 100$）

- フィルタが「単一」の場合
  表 A·11 と表 A·23 からたとえば $n = 20$ のフィルタを 1 回適用した場合の相対残差の最大値 1.8E-02 よりも，$n = 10$ のフィルタを 2 回適用した場合の相対残差の最大値 5.2E-05 の方が精度が 2 桁程度高い．そうして経過時間は前者が 36.5 秒に対して後者は 38.9 秒である．
  同様に $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 3.2E-03 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 1.2E-06 の方が精度が 3 桁程度高い．そうして経過時間は前者が 47.3 秒に対して後者は 50.1 秒である．
- フィルタが「方式 I」の場合
  表 A·14 と表 A·26 からたとえば $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 4.3E-04 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 2.1E-07 の方が精度が 3 桁程度高い．そうして経過時間は前者が 76.7 秒に対して後者は 79.4 秒である．
- フィルタが「方式 II」の場合
  表 A·17 と表 A·29 からたとえば $n = 30$ のフィルタを 1 回適用した場合の相対残差の最大値 1.4E-03 よりも，$n = 15$ のフィルタを 2 回適用した場合の相対残差の最大値 2.2E-06 の方が精度が 3 桁程度高い．そうして経過時間は前者が 77.5 秒に対して後者は 79.8 秒である．

この結果から，最初にランダムなベクトルの組から始めて，多項式次数 n が高いフィルタ 1 回適用するよりも，途中に B-正規直交化をはさんで次数が半分のフィルタを 2 回適用する方が，得られる近似固有対の精度の点では相当に有利であることがわかる．ただし，正規直交化の計算時間（いまの例の場合にはそれほど大した割合ではない）が追加になる．並列計算の場合には正規直交化を行うところで同期待ちも必要になる．

表 A·1 実験その 1：「方式 I」の各例の伝達関数に指定した 4 つのパラメタ

Table A·1 EXP1: Specified 4 parameters for type-I transfer functions.

例	μ	$g_{\rm p}$	$g_{\rm s}$	n
I-1	2.0	1E-2	1E-09	25
I-2	2.0	1E-2	1E-10	35
I-3	2.0	1E-3	1E-12	25
I-4	2.0	1E-3	1E-13	32
I-5	2.0	1E-3	1E-14	40
I-6	1.5	1E-4	1E-11	30

表 A·2 実験その 1：「方式 I」の伝達関数の $x(t)$ に用いた σ_k と α_k

Table A·2 EXP1: Values of σ_k and α_k used in $x(t)$ for Type-I transfer function.

例	k	σ_k	α_k
I-1	1	4.09068 41137 85926 9	9.68147 36896 33707 0
	2	2.02528 07667 67491 7	2.37312 19592 31734 7
I-2	1	5.19655 07817 65392 2	15.25918 03018 57066
	2	3.21576 96254 85300 8	5.84346 85487 09282 1
I-3	1	2.22755 26153 98233 9	10.70208 67035 10560
	2	1.59850 75775 766164	5.51114 60688 83539 0
I-4	1	3.32580 23062 73146 3	8.98973 04258 55874 8
	2	1.79146 09244 00880 6	2.60836 57440 39891 1
I-5	1	3.99137 37417 64652 6	11.75250 98719 03449
	2	2.39289 28457 85695 5	4.22408 19519 01427 9
I-6	1	2.69117 50089 59303 0	8.93745 60356 09324 4
	2	1.71861 35211 28330 2	3.64490 72765 50080 1

表 A·3 実験その 1：「方式 I」のフィルタの反復回数と相対残差の最大値

Table A·3 EXP1: Max of relative residual for iterations of type-I filters

例	反復 1 回	反復 2 回	反復 3 回
I-1	1.0E-04	5.5E-12	2.6E-12
I-2	1.2E-05	4.2E-12	4.3E-12
I-3	1.2E-06	3.8E-12	3.9E-12
I-4	9.5E-08	2.5E-12	2.6E-12
I-5	1.2E-08	3.2E-12	3.6E-12
I-6	8.7E-05	4.5E-12	3.8E-12

表 A·4 実験その 1：「方式 I」：フィルタ反復回数と対角化に要した経過時間（秒）

Table A·4 EXP1: Elapsed times for diagonalization (in second) for iterations of type-I filters.

例	n	m	反復 1 回	反復 2 回	反復 3 回
I-1	25	200	100.2	180.5	256.7
I-2	35	200	128.9	237.2	343.8
I-3	25	200	99.8	180.6	257.3
I-4	32	200	121.1	222.1	318.9
I-5	40	200	143.7	248.8	346.5
I-6	30	125	91.5	164.3	236.5

表 A·5 実験その 1：「方式 II」の各例の伝達関数に指定した 4 つのパラメタ

Table A·5 EXP1: Specified 4 parameters for type-II transfer functions.

例	μ	$g_{\rm p}$	$g_{\rm s}$	n
II-1	2.0	1E-2	1E-13	30
II-2	2.0	1E-2	1E-14	35
II-3	2.0	1E-3	1E-13	21
II-4	1.5	1E-4	1E-12	24
II-5	1.5	1E-4	1E-13	28
II-6	1.25	1E-6	1E-13	29

表 A·6 実験その 1：「方式 II」の伝達関数の $x(t)$ に用いた σ_k と α_k

Table A·6 EXP1: Values of σ_k and α_k used in $x(t)$ for Type-II transfer function.

例	k	σ_k	α_k
II-1	1	1.67933 35315 46617 8	12.84712 18363 24346
	2	1.25898 93885 43740 0	8.12041 76097 42180 1
II-2	1	1.92356 13781 91710 9	14.18630 98321 53896
	2	1.45862 38171 49344 4	9.04662 44340 09778 8
II-3	1	1.22291 68196 12936 5	4.32668 10367 40262 2
	2	0.37200 77616 25172 68	0.81235 02518 27033 46
II-4	1	1.23356 16207 65095 2	3.93345 42009 89467 5
	2	0.41603 30166 83183 49	0.84103 96834 36731 41
II-5	1	0.96499 058641 91108 4	12.22386 05471 97841
	2	0.78605 22624 66379 16	9.05045 51521 88670 0
II-6	1	0.97498 17452 4114078	4.55966 85101 81800 2
	2	0.51619 30340 47827 13	1.85327 70031 67030 3

表 A·7 実験その 1：「方式 II」のフィルタの反復回数と相対残差の最大値

Table A·7 EXP1: Max relative residual for iterations of type-II filters

例	反復 1 回	反復 2 回	反復 3 回
II-1	4.8E-07	8.1E-12	9.6E-12
II-2	4.1E-08	9.8E-12	1.0E-11
II-3	1.3E-06	2.2E-12	2.7E-12
II-4	1.1E-04	3.6E-12	4.0E-12
II-5	1.7E-05	1.8E-11	1.9E-11
II-6	2.7E-03	1.8E-10	9.8E-12

表 A·8 実験その 1：「方式 II」：フィルタ反復回数と対角化に要した経過時間（秒）

Table A·8 EXP1: Elapsed times for diagonalization (in second) for iterations of type-II filters.

例	n	m	反復 1 回	反復 2 回	反復 3 回
II-1	30	200	116.0	210.2	301.1
II-2	35	200	129.5	220.8	307.2
II-3	21	200	88.9	159.5	224.4
II-4	24	125	80.3	137.9	195.2
II-5	28	125	88.3	157.0	224.0
II-6	29	100	75.6	131.5	186.8

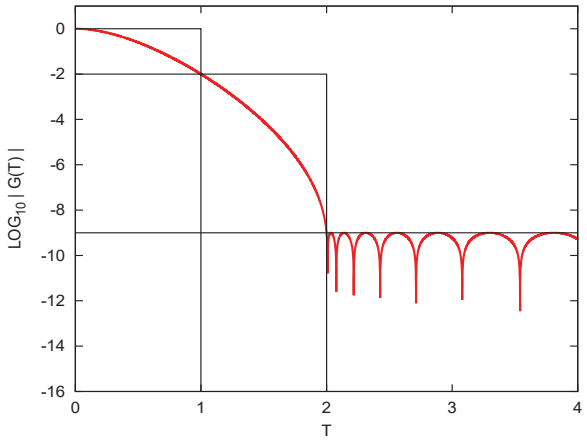

図 A·1 実験その 1：例 I-1：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-2, g_s=1E-09, n=25）

Fig. A·1 EXP1: Example I-1: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-2, g_s=1E-09, n=25).

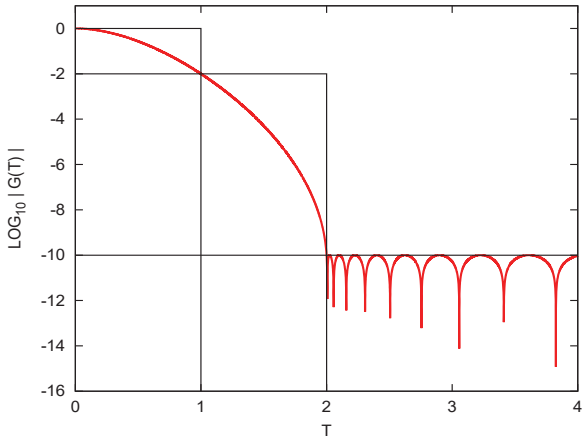

図 A·2 実験その 1：例 I-2：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-2, g_s=1E-10, n=35）

Fig. A·2 EXP1: Example I-2: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-2, g_s=1E-10, n=35).

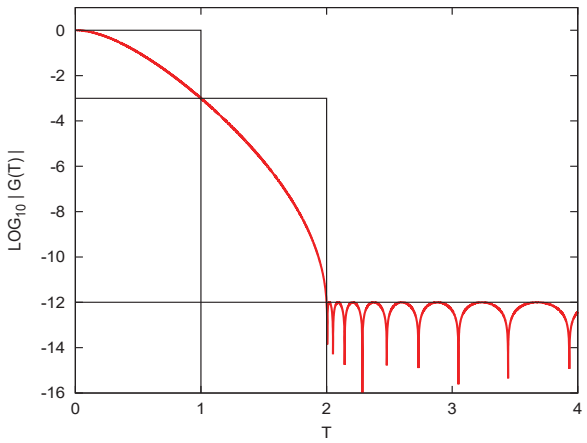

図 A·3 実験その 1：例 I-3：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-3, g_s=1E-12, n=25）

Fig. A·3 EXP1: Example I-3: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-3, g_s=1E-12, n=25).

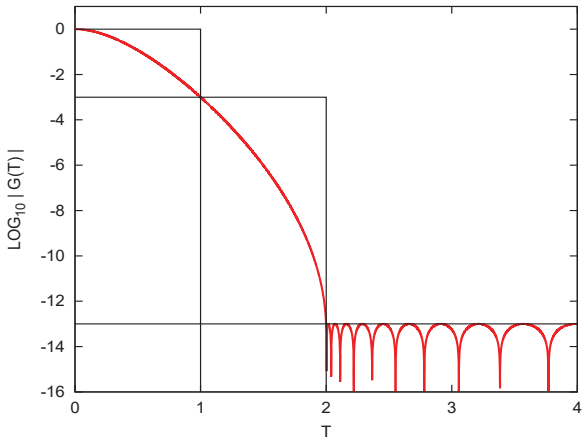

図 A·4 実験その 1：例 I-4：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-3, g_s=1E-13, n=32）

Fig. A·4 EXP1: Example I-4: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-3, g_s=1E-13, n=32).

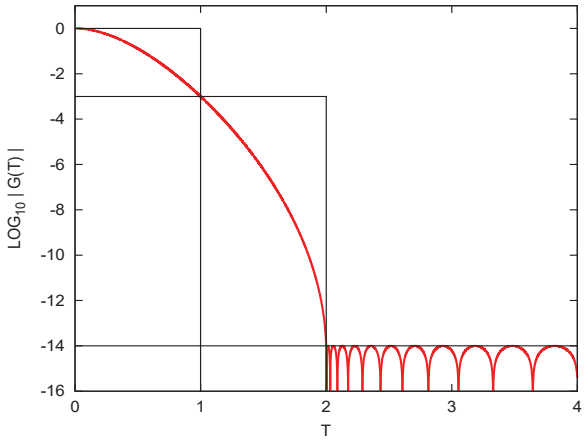

図 A·5 実験その 1：例 I-5：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-3, g_s=1E-14, n=40）

Fig. A·5 EXP1: Example I-5: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-3, g_s=1E-14, n=40).

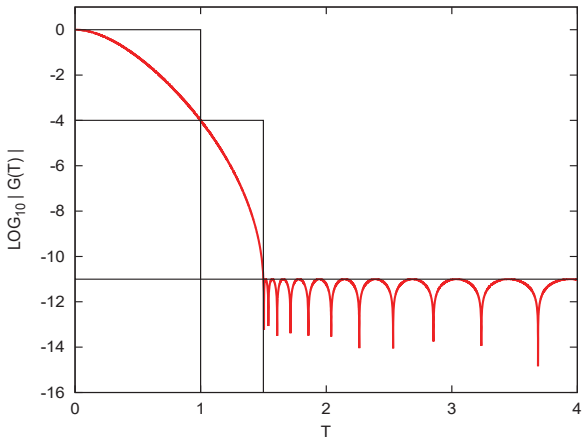

図 A·6 実験その 1：例 I-6：伝達関数の大きさ $|g(t)|$ （μ=1.5, g_p=1E-4, g_s=1E-11, n=30）

Fig. A·6 EXP1: Example I-6: Transfer function magnitude $|g(t)|$ (μ=1.5, g_p=1E-4, g_s=1E-11, n=30).

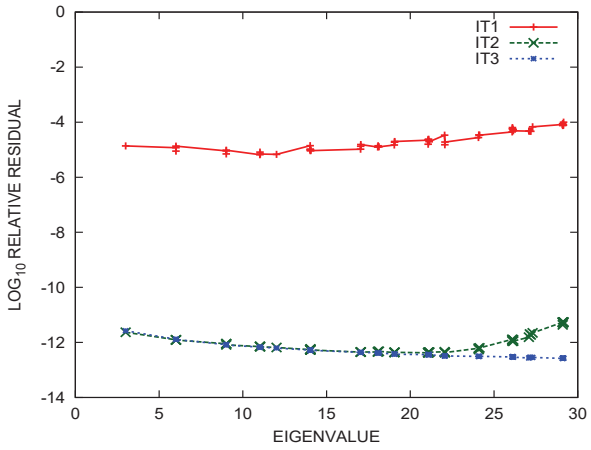

図 A·7 実験その 1：例 I-1：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0, g_{p}=1E-2, g_{s}=1E-09, n=25）

Fig. A·7 EXP1: Example I-1: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-2, g_{s}=1E-09, n=25).

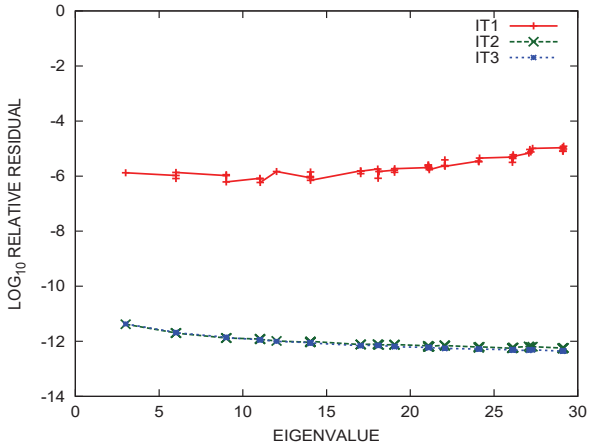

図 A·8 実験その 1：例 I-2：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0, g_{p}=1E-2, g_{s}=1E-10, n=35）

Fig. A·8 EXP1: Example I-2: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-2, g_{s}=1E-10, n=35).

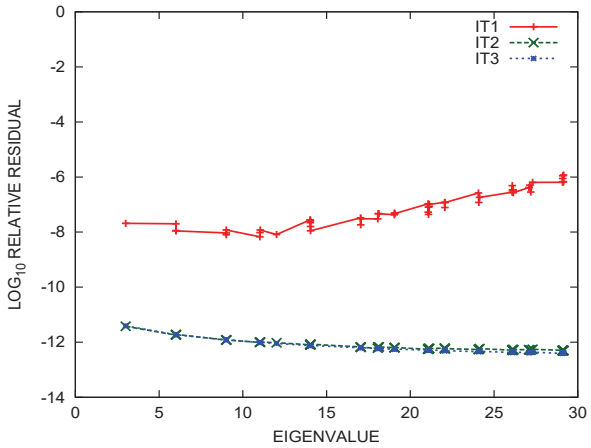

図 A·9 実験その 1：例 I-3：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0, g_{p}=1E-3, g_{s}=1E-12, n=25）

Fig. A·9 EXP1: Example I-3: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-3, g_{s}=1E-12, n=25).

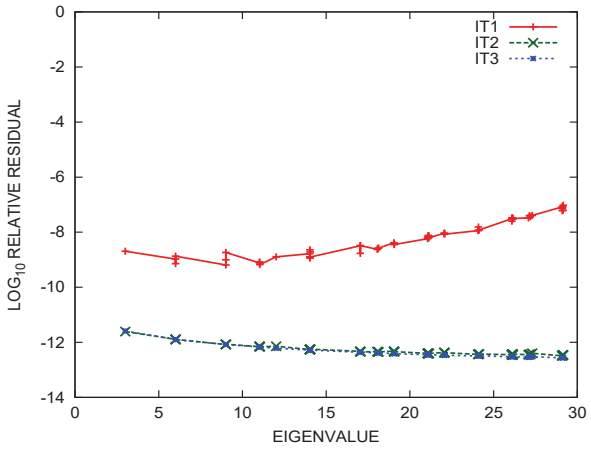

図 A·10 実験その 1：例 I-4：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0, g_{p}=1E-3, g_{s}=1E-13, n=32）

Fig. A·10 EXP1: Example I-4: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-3, g_{s}=1E-13, n=32).

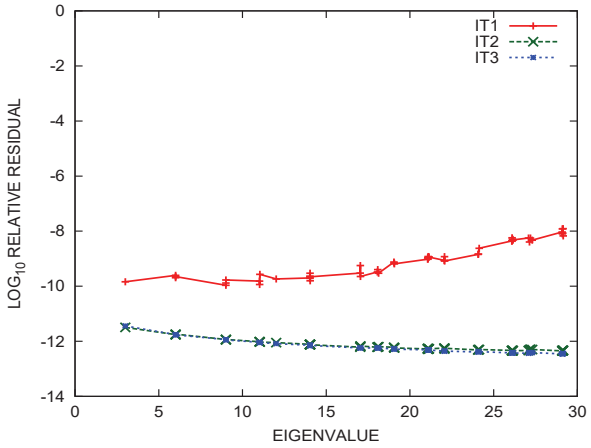

図 A·11 実験その 1：例 I-5：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0, g_{p}=1E-3, g_{s}=1E-14, n=40）

Fig. A·11 EXP1: Example I-5: Relative residuals of eigenpairs of iterations of the filter (μ=2.0, g_{p}=1E-3, g_{s}=1E-14, n=40).

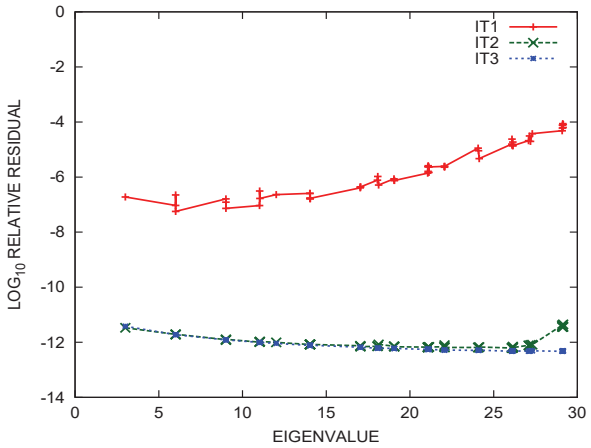

図 A·12 実験その 1：例 I-6：フィルタ反復回数ごとの各固有対の相対残差（μ=1.5, g_{p}=1E-4, g_{s}=1E-11, n=30）

Fig. A·12 EXP1: Example I-6: Relative residuals of eigenpairs for iterations of the filter (μ=1.5, g_{p}=1E-4, g_{s}=1E-11, n=30).

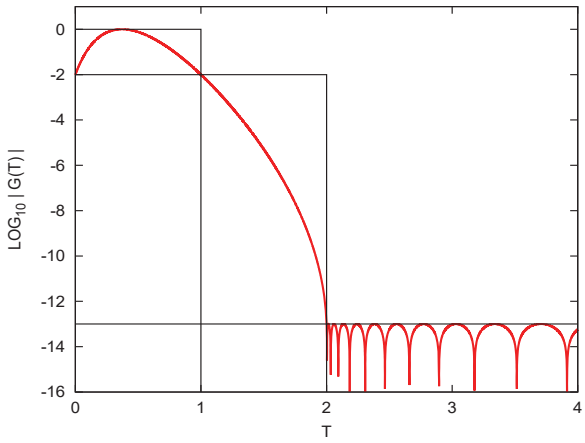

図 A·13 実験その 1：例 II-1：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-2, g_s=1E-13, n=30）

Fig. A·13 EXP1: Example II-1: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-2, g_s=1E-13, n=30).

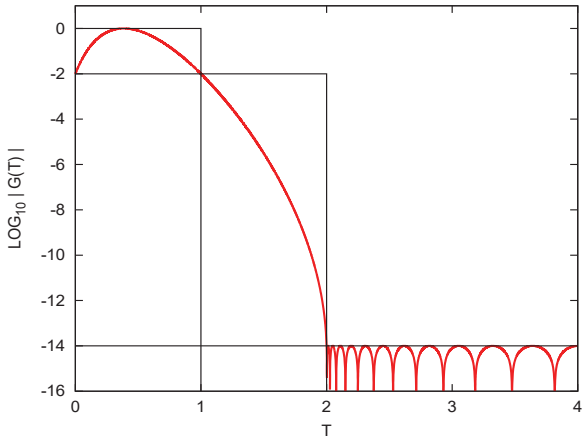

図 A·14 実験その 1：例 II-2：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-2, g_s=1E-14, n=35）

Fig. A·14 EXP1: Example II-2: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-2, g_s=1E-14, n=35).

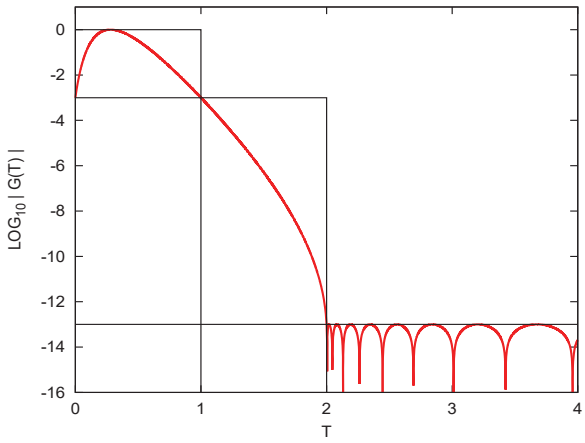

図 A·15 実験その 1：例 II-3：伝達関数の大きさ $|g(t)|$ （μ=2.0, g_p=1E-3, g_s=1E-13, n=21）

Fig. A·15 EXP1: Example II-3: Transfer function magnitude $|g(t)|$ (μ=2.0, g_p=1E-3, g_s=1E-13, n=21).

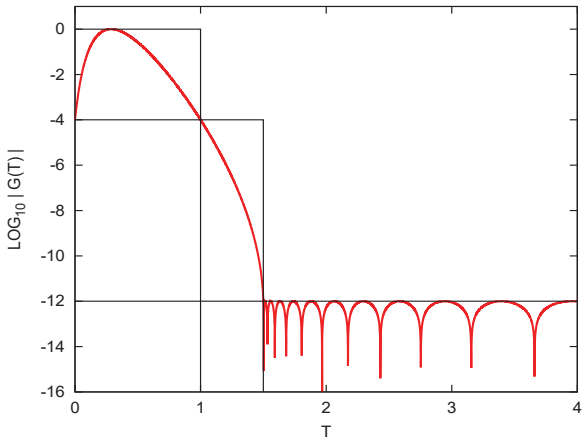

図 A·16 実験その 1：例 II-4：伝達関数の大きさ $|g(t)|$ （μ=1.5, g_p=1E-4, g_s=1E-12, n=24）

Fig. A·16 EXP1: Example II-4: Transfer function magnitude $|g(t)|$ (μ=1.5, g_p=1E-4, g_s=1E-12, n=24).

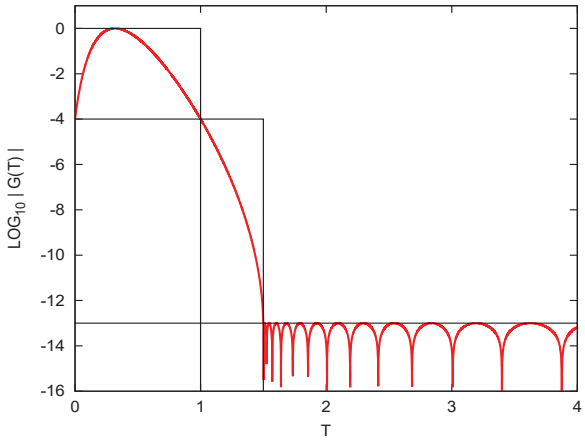

図 A·17 実験その 1：例 II-5：伝達関数の大きさ $|g(t)|$ （μ=1.5, g_p=1E-4, g_s=1E-13, n=28）

Fig. A·17 EXP1: Example II-5: Transfer function magnitude $|g(t)|$ (μ=1.5, g_p=1E-4, g_s=1E-13, n=28).

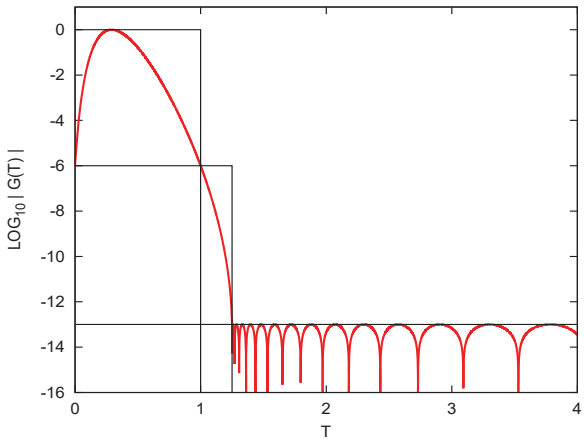

図 A·18 実験その 1：例 II-6：伝達関数の大きさ $|g(t)|$ （μ=1.25, g_p=1E-6, g_s=1E-13, n=29）

Fig. A·18 EXP1: Example II-6: Transfer function magnitude $|g(t)|$ (μ=1.25, g_p=1E-6, g_s=1E-13, n=29).

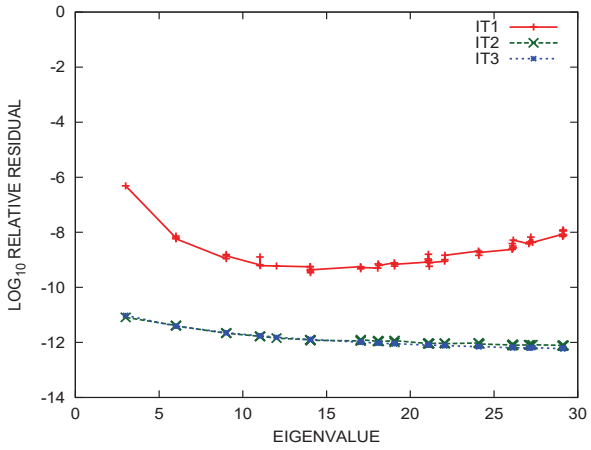

図 A·19 実験その 1：例 II-1：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0，g_{p}=1E-2，g_{s}=1E-13，n=30）

Fig. A·19 EXP1: Example II-1: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-2, g_{s}=1E-13, n=30).

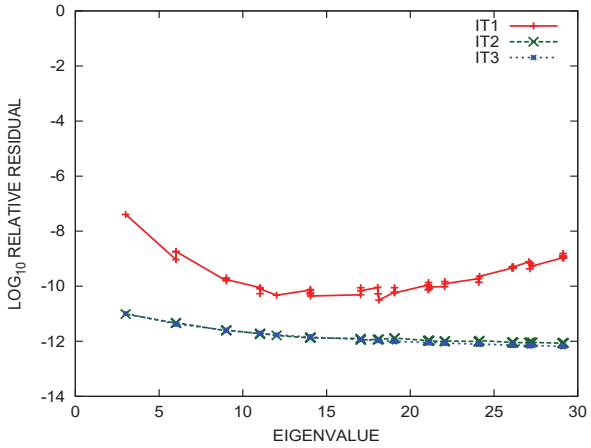

図 A·20 実験その 1：例 II-2：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0，g_{p}=1E-2，g_{s}=1E-14，n=35）

Fig. A·20 EXP1: Example II-2: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-2, g_{s}=1E-14, n=35).

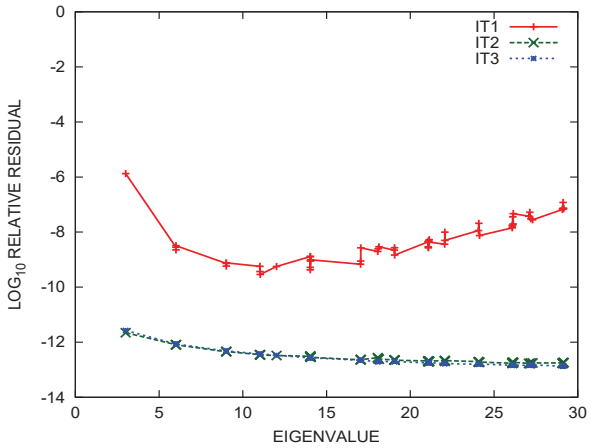

図 A·21 実験その 1：例 II-3：フィルタ反復回数ごとの各固有対の相対残差（μ=2.0，g_{p}=1E-3，g_{s}=1E-13，n=21）

Fig. A·21 EXP1: Example II-3: Relative residuals of eigenpairs for iterations of the filter (μ=2.0, g_{p}=1E-3, g_{s}=1E-13, n=21).

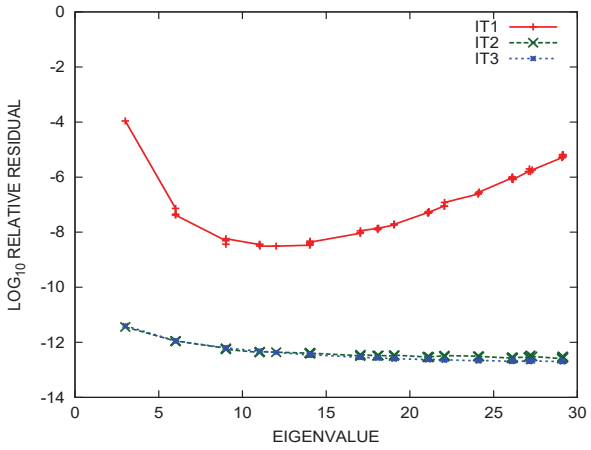

図 A·22 実験その 1：例 II-4：フィルタ反復回数ごとの各固有対の相対残差（μ=1.5，g_{p}=1E-4，g_{s}=1E-12，n=24）

Fig. A·22 EXP1: Example II-4: Relative residuals of eigenpairs for iterations of the filter (μ=1.5, g_{p}=1E-4, g_{s}=1E-12, n=24).

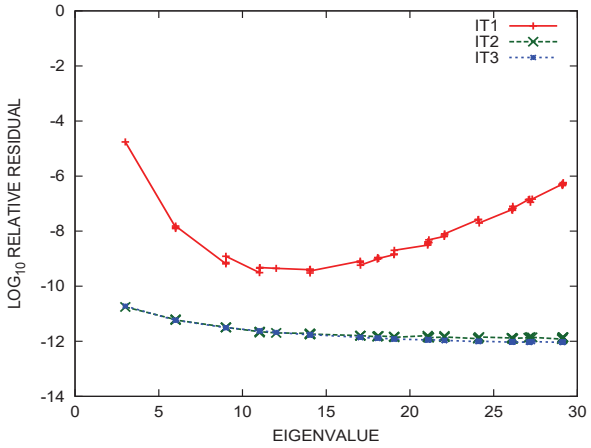

図 A·23 実験その 1：例 II-5：フィルタ反復回数ごとの各固有対の相対残差（μ=1.5，g_{p}=1E-4，g_{s}=1E-13，n=28）

Fig. A·23 EXP1: Example II-5: Relative residuals of eigenpairs for iterations of the filter (μ=1.5, g_{p}=1E-4, g_{s}=1E-13, n=28).

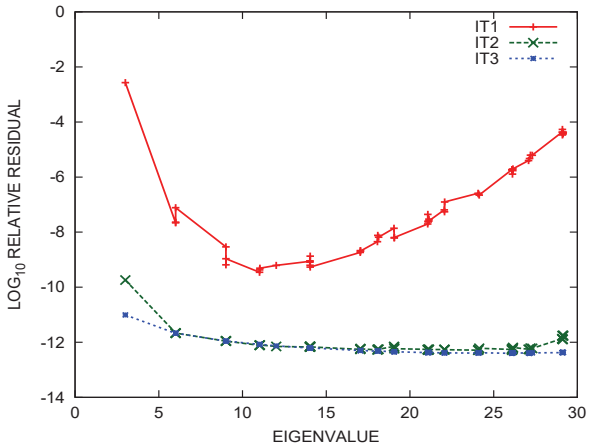

図 A·24 実験その 1：例 II-6：フィルタ反復回数ごとの各固有対の相対残差（μ=1.25，g_{p}=1E-6，g_{s}=1E-13，n=29）

Fig. A·24 EXP1: Example II-6: Relative residuals of eigenpairs for iterations of the filter (μ=1.25, g_{p}=1E-6, g_{s}=1E-13, n=29).

表 A·9 実験その 2：相対残差の最大値（「単一」, $\mu = 2.0$, $m = 200$）
Table A·9 EXP2: Max relative residual ("single", μ=2.0, m=200).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	1.19E-07	5.7E-04	8.7E-11	4.4E-13
15	3.81E-06	2.0E-05	5.6E-13	5.6E-13
20	1.53E-05	7.4E-06	7.1E-13	7.0E-13
25	3.05E-05	5.1E-06	8.2E-13	8.3E-13
30	3.05E-05	3.7E-06	7.0E-13	6.8E-13
35	6.10E-05	2.9E-06	9.4E-13	9.4E-13
40	6.10E-05	2.9E-06	7.6E-13	7.6E-13

表 A·10 実験その 2：相対残差の最大値（「単一」, $\mu = 1.5$, $m = 125$）
Table A·10 EXP2: Max relative residual ("single", μ=1.5, m=125).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	3.73E-09	4.5E-02	2.3E-07	3.7E-12
15	5.96E-08	1.0E-03	9.2E-10	5.4E-13
20	2.38E-07	3.2E-04	4.9E-11	6.3E-13
25	4.77E-07	2.0E-04	1.7E-11	7.3E-13
30	4.77E-07	2.1E-04	1.2E-11	6.2E-13
35	9.54E-07	8.7E-05	4.6E-12	8.8E-13
40	9.54E-07	8.6E-05	6.2E-12	7.5E-13

表 A·11 実験その 2：相対残差の最大値（「単一」, $\mu = 1.25$, $m = 100$）
Table A·11 EXP2: Max relative residual ("single", μ=1.25, m=100).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	2.33E-10	7.4E-02	5.2E-05	1.0E-08
15	1.86E-09	6.2E-02	1.2E-06	4.2E-11
20	3.73E-09	1.8E-02	2.5E-07	3.3E-12
25	7.45E-09	8.9E-03	6.7E-08	6.3E-13
30	1.49E-08	3.2E-03	1.3E-08	8.0E-13
35	1.49E-08	3.7E-03	1.2E-08	6.7E-13
40	1.49E-08	2.9E-03	8.8E-09	6.0E-13

表 A·12 実験その 2：相対残差の最大値（「方式 I」, $\mu = 2.0$, $m = 200$）
Table A·12 EXP2: Max relative residual ("type-I", μ=2.0, m=200).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	2.38E-07	3.3E-04	8.9E-11	4.6E-13
15	1.53E-05	6.0E-06	1.2E-12	1.2E-12
20	1.22E-04	8.6E-07	1.9E-12	2.0E-12
25	2.44E-04	3.9E-07	1.1E-12	1.1E-12
30	4.88E-04	2.3E-07	1.3E-12	1.3E-12
35	9.77E-04	1.1E-07	1.5E-12	1.6E-12
40	1.95E-03	7.3E-08	3.5E-12	3.6E-12

表 A·13 実験その 2：相対残差の最大値（「方式 I」, $\mu = 1.5$, $m = 125$）
Table A·13 EXP2: Max relative residual ("type-I", μ=1.5, m=125).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	7.45E-09	1.6E-02	6.7E-08	7.1E-13
15	1.19E-07	6.5E-04	3.1E-10	6.2E-13
20	9.54E-07	9.2E-05	3.7E-12	8.9E-13
25	3.81E-06	1.7E-05	1.3E-12	1.3E-12
30	7.63E-06	8.0E-06	1.4E-12	1.4E-12
35	1.53E-05	4.6E-06	1.8E-12	1.9E-12
40	1.53E-05	4.7E-06	1.5E-12	1.6E-12

表 A·14 実験その 2：相対残差の最大値（「方式 I」, $\mu = 1.25$, $m = 100$）
Table A·14 EXP2: Max relative residual ("type-I", μ=1.25, m=100).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	---	---	---	---
15	3.73E-09	2.7E-02	2.1E-07	3.7E-12
20	1.49E-08	3.0E-03	1.1E-08	8.4E-13
25	5.96E-08	9.9E-04	8.0E-10	1.5E-12
30	1.19E-07	4.3E-04	1.6E-10	1.8E-12
35	1.19E-07	4.0E-04	1.5E-10	1.4E-12
40	2.38E-07	2.3E-04	3.3E-11	1.8E-12

表 A·15 実験その 2：相対残差の最大値（「方式 II」，$\mu = 2.0$, $m = 200$）

Table A·15 EXP2: Max relative residual ("type-II", μ=2.0, m=200).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	4.77E-07	4.4E-03	2.9E-10	4.7E-12
15	6.10E-05	3.4E-05	3.1E-12	4.6E-12
20	4.88E-04	6.5E-06	1.9E-12	2.0E-12
25	3.91E-03	7.1E-07	4.3E-12	5.0E-12
30	7.81E-03	2.5E-07	2.6E-12	2.8E-12
35	1.56E-02	1.8E-07	3.8E-12	3.7E-12
40	1.56E-02	1.4E-07	2.3E-12	2.3E-12

表 A·16 実験その 2：相対残差の最大値（「方式 II」，$\mu = 1.5$, $m = 125$）

Table A·16 EXP2: Max relative residual ("type-II", μ=1.5, m=125).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	7.45E-09	8.0E-02	2.1E-07	4.5E-12
15	4.77E-07	9.2E-03	5.2E-10	4.2E-12
20	7.63E-06	1.2E-04	4.4E-12	5.2E-12
25	3.05E-05	4.4E-05	3.2E-12	3.3E-12
30	1.22E-04	1.3E-05	4.5E-12	4.6E-12
35	2.44E-04	1.4E-05	3.9E-12	4.0E-12
40	4.88E-04	8.2E-06	5.8E-12	4.9E-12

表 A·17 実験その 2：相対残差の最大値（「方式 II」，$\mu = 1.25$, $m = 100$）

Table A·17 EXP2: Max relative residual ("type-II", μ=1.25, m=100).

n	g_{p}	反復 1 回	反復 2 回	反復 3 回
10	---	---	---	---
15	7.45E-09	7.1E-01	2.2E-06	3.8E-11
20	5.96E-08	5.3E-02	2.8E-08	4.7E-12
25	2.38E-07	8.6E-03	1.9E-09	4.1E-12
30	9.54E-07	1.4E-03	1.3E-10	5.4E-12
35	1.91E-06	2.4E-03	3.1E-11	4.9E-12
40	3.81E-06	5.8E-04	9.0E-12	5.3E-12

表 A·18 実験その 2：次数 n の各種フィルタで採択された g_{p} の値（$\mu = 2.0$ の場合）

Table A·18 EXP2: Selected value of g_{p} for filters of degree n (case $\mu = 2.0$).

n	「単一」	「方式 I」	「方式 II」
10	1.19E-07	2.38E-07	4.77E-07
15	3.81E-06	1.53E-05	6.10E-05
20	1.53E-05	1.22E-04	4.88E-04
25	3.05E-05	2.44E-04	3.91E-03
30	3.05E-05	4.88E-04	7.81E-03
35	6.10E-05	9.77E-04	1.56E-02
40	6.10E-05	1.95E-03	1.56E-02

表 A·19 実験その 2：次数 n の各種フィルタで採択された g_{p} の値（$\mu = 1.5$ の場合）

Table A·19 EXP2: Selected value of g_{p} for filters of degree n (case $\mu = 1.5$).

n	「単一」	「方式 I」	「方式 II」
10	3.73E-09	7.45E-09	7.45E-09
15	5.96E-08	1.19E-07	4.77E-07
20	2.38E-07	9.54E-07	7.63E-06
25	4.77E-07	3.81E-06	3.05E-05
30	4.77E-07	7.63E-06	1.22E-04
35	9.54E-07	1.53E-05	2.44E-04
40	9.54E-07	1.53E-05	4.88E-04

表 A·20 実験その 2：次数 n の各種フィルタで採択された g_{p} の値（$\mu = 1.25$ の場合）

Table A·20 EXP2: Selected value of g_{p} for filters of degree n (case $\mu = 1.25$).

n	「単一」	「方式 I」	「方式 II」
10	2.33E-10	---	---
15	1.86E-09	3.73E-09	7.45E-09
20	3.73E-09	1.49E-08	5.96E-08
25	7.45E-09	5.96E-08	2.38E-07
30	1.49E-08	1.19E-07	9.54E-07
35	1.49E-08	1.19E-07	1.91E-06
40	1.49E-08	2.38E-07	3.81E-06

表 A·21 実験その 2：経過時間（秒）（「単一」, $\mu = 2.0$, $m = 200$）
Table A·21 EXP2: Elapse time in second (”single”, μ=2.0, m=200).

n	反復 1 回	反復 2 回	反復 3 回
10	37.3	65.2	87.0
15	47.4	82.9	114.8
20	56.1	100.4	140.4
25	64.8	118.1	166.1
30	73.8	134.9	191.1
35	82.8	153.6	217.9
40	90.9	170.9	244.4

表 A·22 実験その 2：経過時間（秒）（「単一」, $\mu = 1.5$, $m = 125$）
Table A·22 EXP2: Elapse time in second (”single”, μ=1.5, m=125).

n	反復 1 回	反復 2 回	反復 3 回
10	29.1	48.0	64.2
15	36.1	61.6	84.8
20	43.5	75.1	105.4
25	50.4	89.1	126.2
30	57.4	103.4	147.5
35	65.1	117.1	168.3
40	70.8	130.5	189.8

表 A·23 実験その 2：経過時間（秒）（「単一」, $\mu = 1.25$, $m = 100$）
Table A·23 EXP2: Elapse time in second (”single”, μ=1.25, m=100).

n	反復 1 回	反復 2 回	反復 3 回
10	24.4	38.9	51.6
15	30.5	50.1	68.3
20	36.5	60.9	85.3
25	41.7	73.0	101.8
30	47.3	84.2	118.4
35	52.9	94.3	135.0
40	58.2	106.0	151.7

表 A·24 実験その 2：経過時間（秒）（「方式 I」, $\mu = 2.0$, $m = 200$）
Table A·24 EXP2: Elapse time in second (”type-I”, μ=2.0, m=200).

n	反復 1 回	反復 2 回	反復 3 回
10	56.9	95.3	129.3
15	70.8	124.6	172.1
20	85.3	153.7	214.7
25	101.5	183.3	259.8
30	114.9	211.4	301.9
35	129.4	239.1	344.1
40	144.3	266.9	389.2

表 A·25 実験その 2：経過時間（秒）（「方式 I」, $\mu = 1.5$, $m = 125$）
Table A·25 EXP2: Elapse time in second (”type-I”, μ=1.5, m=125).

n	反復 1 回	反復 2 回	反復 3 回
10	46.3	73.1	99.8
15	57.4	95.7	134.6
20	68.6	119.7	167.8
25	80.7	141.7	202.7
30	92.6	166.1	236.8
35	103.5	188.0	271.6
40	114.6	211.9	306.3

表 A·26 実験その 2：経過時間（秒）（「方式 I」, $\mu = 1.25$, $m = 100$）
Table A·26 EXP2: Elapse time in second (”type-I”, μ=1.25, m=100).

n	反復 1 回	反復 2 回	反復 3 回
10	---	---	---
15	49.0	79.4	108.6
20	59.1	98.1	137.1
25	66.6	116.8	163.8
30	76.7	133.7	192.1
35	85.3	153.2	220.1
40	95.1	171.3	245.8

表 A·27 実験その 2：経過時間 (秒) (「方式 II」, $\mu = 2.0, m = 200$)
Table A·27 EXP2: Elapse time in second ("type-II", μ=2.0, m=200).

n	反復 1 回	反復 2 回	反復 3 回
10	57.0	96.0	129.6
15	72.4	125.2	171.6
20	87.0	152.8	215.2
25	100.5	183.0	258.2
30	115.0	210.8	301.9
35	129.5	238.5	344.0
40	144.3	268.3	387.6

表 A·28 実験その 2：経過時間 (秒) (「方式 II」, $\mu = 1.5, m = 125$)
Table A·28 EXP2: Elapse time in second ("type-II", μ=1.5, m=125).

n	反復 1 回	反復 2 回	反復 3 回
10	45.7	72.5	98.0
15	59.0	97.0	134.8
20	69.4	120.5	168.6
25	81.5	142.6	202.5
30	92.4	165.5	237.4
35	104.7	188.6	271.2
40	115.6	211.7	305.5

表 A·29 実験その 2：経過時間 (秒) (「方式 II」, $\mu = 1.25, m = 100$)
Table A·29 EXP2: Elapse time in second ("type-II", μ=1.25, m=100).

n	反復 1 回	反復 2 回	反復 3 回
10	---	---	---
15	49.4	79.8	110.3
20	58.6	98.8	138.0
25	68.1	116.6	165.6
30	77.5	134.9	191.9
35	86.4	153.8	222.0
40	96.5	173.1	246.8

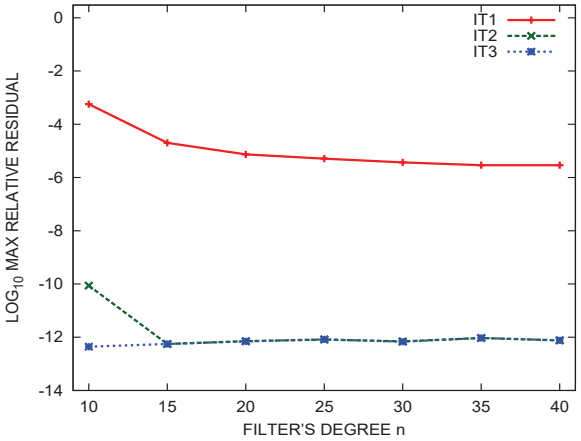

図 A·25 実験その 2：フィルタの次数と相対残差の最大値 (「単一」, $\mu = 2.0,\ m = 200$)
Fig. A·25 EXP2: Filter's degree vs. max of relative residuals ("single", $\mu = 2.0$, $m = 200$).

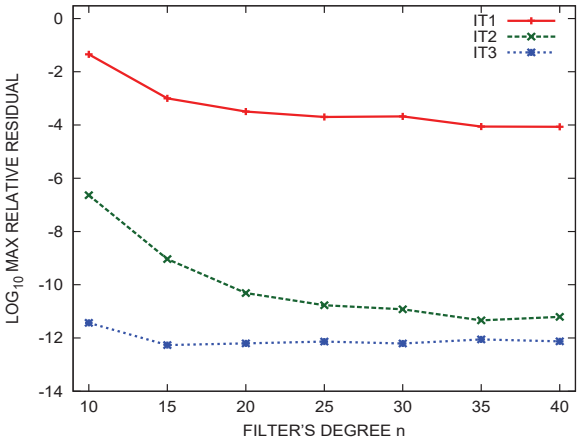

図 A·26 実験その 2：フィルタの次数と相対残差の最大値 (「単一」, $\mu = 1.5,\ m = 125$)
Fig. A·26 EXP2: Filter's degree vs. max of relative residuals ("single", $\mu = 1.5$, $m = 125$).

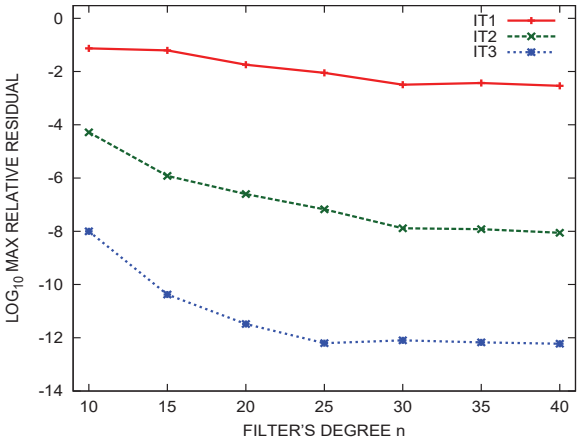

図 A·27 実験その 2：フィルタの次数と相対残差の最大値 (「単一」, $\mu = 1.25,\ m = 100$)
Fig. A·27 EXP2: Filter's degree vs. max of relative residuals ("single", $\mu = 1.25$, $m = 100$).

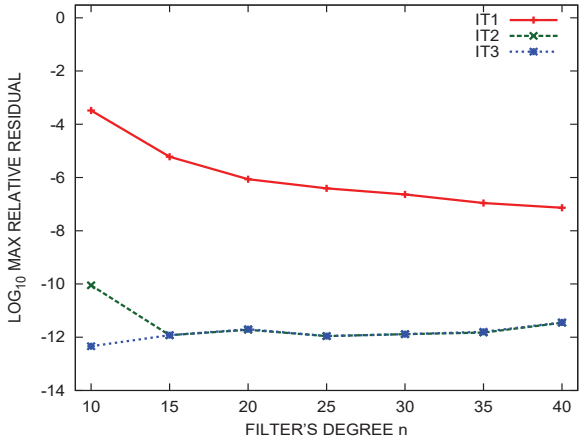

図 A·28 実験その 2：フィルタの次数と相対残差の最大値（「方式 I」，$\mu = 2.0$，$m = 200$）

Fig. A·28 EXP2: Filter's degree vs. max of relative residuals ("type-I", $\mu = 2.0$, $m = 200$).

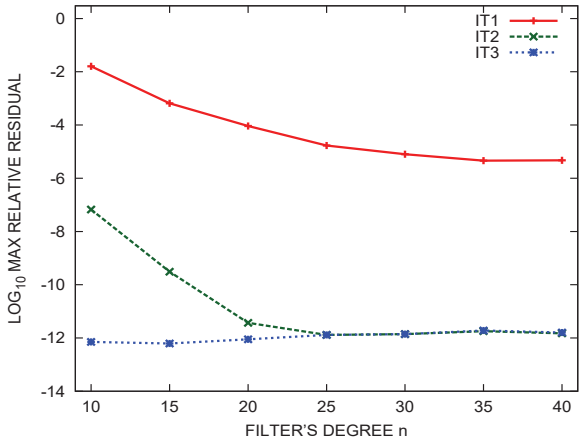

図 A·29 実験その 2：フィルタの次数と相対残差の最大値（「方式 I」，$\mu = 1.5$，$m = 125$）

Fig. A·29 EXP2: Filter's degree vs. max of relative residuals ("type-I", $\mu = 1.5$, $m = 125$).

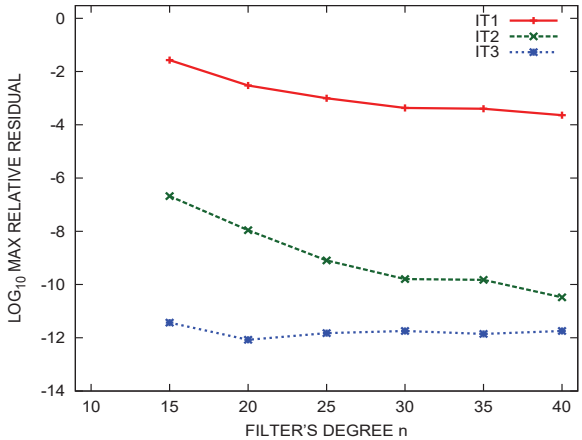

図 A·30 実験その 2：フィルタの次数と相対残差の最大値（「方式 I」，$\mu = 1.25$，$m = 100$）

Fig. A·30 EXP2: Filter's degree vs. max of relative residuals ("type-I", $\mu = 1.25$, $m = 100$).

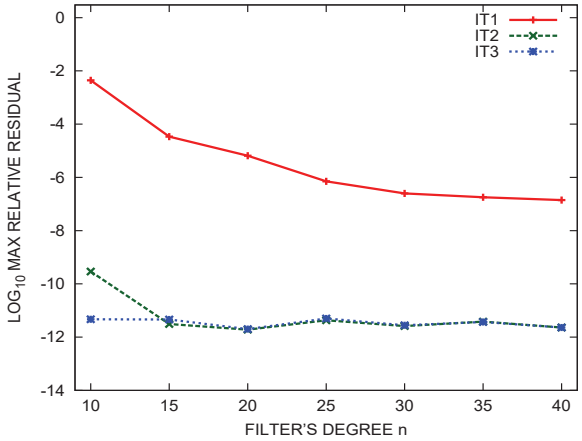

図 A·31 実験その 2：フィルタの次数と相対残差の最大値（「方式 II」，$\mu = 2.0$，$m = 200$）

Fig. A·31 EXP2: Filter's degree vs. max of relative residuals ("type-II", $\mu = 2.0$, $m = 200$).

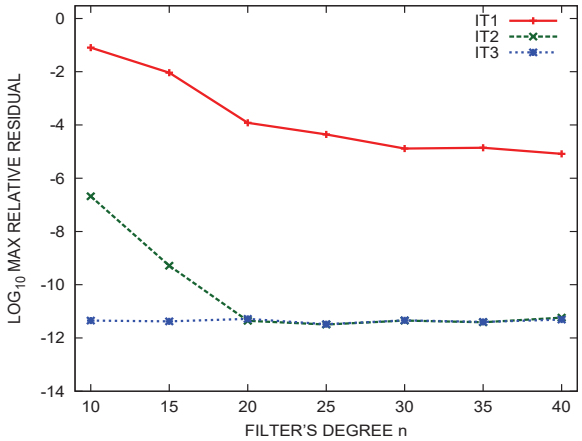

図 A·32 実験その 2：フィルタの次数と相対残差の最大値（「方式 II」，$\mu = 1.5$，$m = 125$）

Fig. A·32 EXP2: Filter's degree vs. max of relative residuals ("type-II", $\mu = 1.5$, $m = 125$).

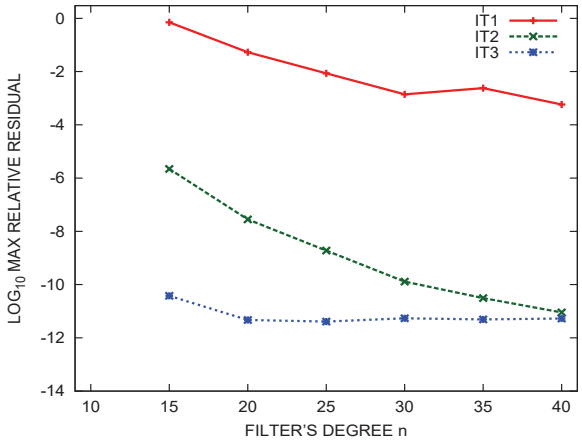

図 A·33 実験その 2：フィルタの次数と相対残差の最大値（「方式 II」，$\mu = 1.25$，$m = 100$）

Fig. A·33 EXP2: Filter's degree vs. max of relative residuals ("type-II", $\mu = 1.25$, $m = 100$).

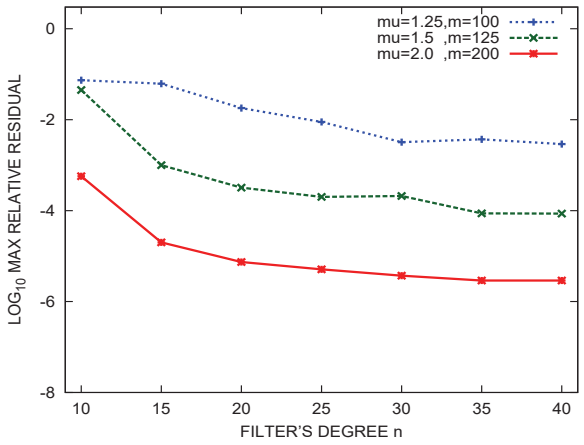

図 A·34 実験その 2：フィルタ適用 1 回目の相対残差の最大値（「単一」）

Fig. A·34 EXP2: Max relative residual of 1st filter application ("single").

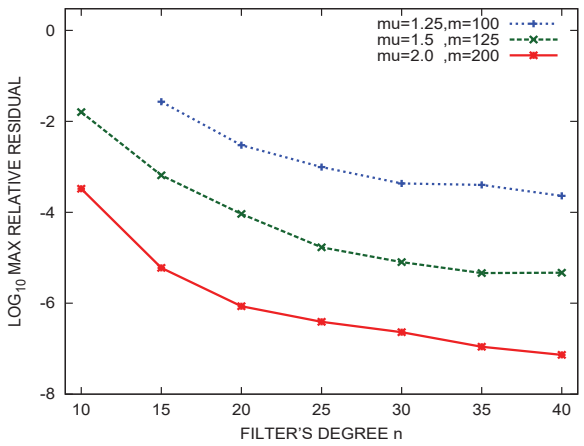

図 A·35 実験その 2：フィルタ適用 1 回目の相対残差の最大値（「方式 I」）

Fig. A·35 EXP2: Max relative residual of 1st filter application ("type-I").

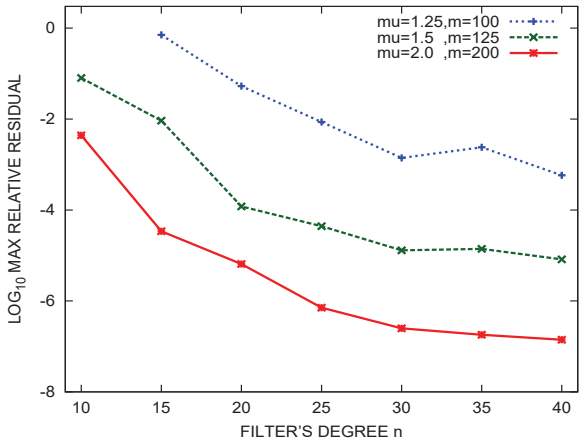

図 A·36 実験その 2：フィルタ適用 1 回目の相対残差の最大値（「方式 II」）

Fig. A·36 EXP2: Max relative residual of 1st filter application ("type-II").

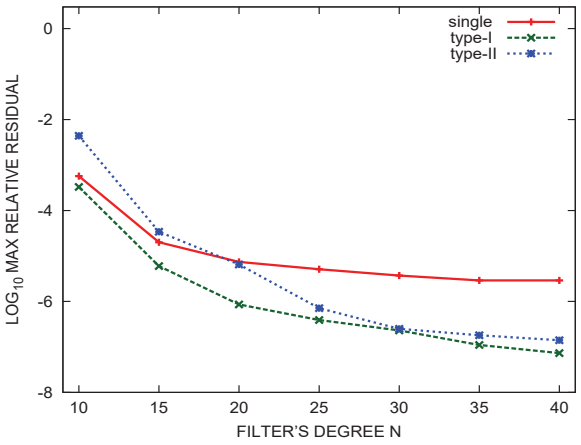

図 A·37 実験その 2：フィルタ適用 1 回目の相対残差の最大値（$\mu = 2.0,\ m = 200$）

Fig. A·37 EXP2: Max relative residual of 1st filter application (μ=2.0, m=200).

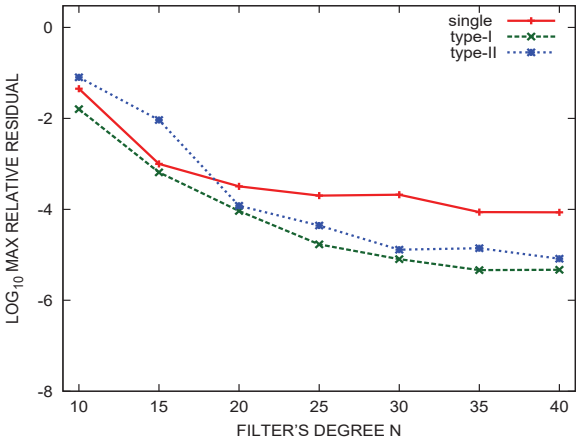

図 A·38 実験その 2：フィルタ適用 1 回目の相対残差の最大値（$\mu = 1.5,\ m = 125$）

Fig. A·38 EXP2: Max relative residual of 1st filter application (μ=1.5, m=125).

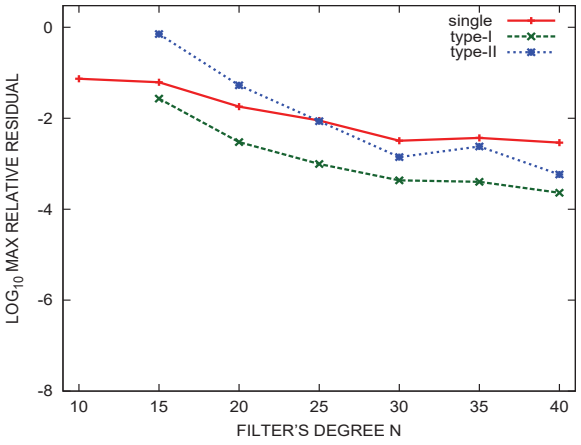

図 A·39 実験その 2：フィルタ適用 1 回目の相対残差の最大値（$\mu = 1.25,\ m = 100$）

Fig. A·39 EXP2: Max relative residual of 1st filter application (μ=1.25, m=100).

A.4 計算式の導出

本文中（副節 §2.1 と副節 §2.2）では省略をしたが，「方式 I」と「方式 II」についてそれぞれ，伝達関数を決定するための 4 つの実数値 σ_1，σ_2，α_1，α_2 の算出方法の導出過程をここに記す．

A.4.1 「方式 I」の場合

まず連立式 (13) の第 4 番目の式から式 (A.3) の最初の等式が得られるが，その値は t によらない定数なのでそれをここでは C とおく．

$$\frac{\alpha_1}{\sigma_1^2} = \frac{\alpha_2}{\sigma_2^2} = C\,. \tag{A.3}$$

すると，2 つの極の係数はそれぞれ式 (A.4) で表せるので，それを用いて α_1 と α_2 を消去できる．

$$\begin{cases} \alpha_1 = C\sigma_1^2, \\ \alpha_2 = C\sigma_2^2\,. \end{cases} \tag{A.4}$$

連立式 (13) の第 1 番目の式と式 (A.4) から，定数 C の値の逆数は式 (A.5) で表せる．

$$\frac{1}{C} = \frac{\sigma_1^2}{\mu+\sigma_1} - \frac{\sigma_2^2}{\mu+\sigma_2} = (\sigma_1-\sigma_2)\times\frac{\mu(\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(\mu+\sigma_1)(\mu+\sigma_2)}\,. \tag{A.5}$$

連立式 (13) の第 3 番目の式と式 (A.4) と式 (A.5) を併せると，式 (A.6) が得られる．

$$x_H = C\times(\sigma_1-\sigma_2) = \frac{(\mu+\sigma_1)(\mu+\sigma_2)}{\mu(\sigma_1+\sigma_2)+\sigma_1\sigma_2}\,. \tag{A.6}$$

この式 (A.6) から（構成が可能な場合には）$\sigma_1 > \sigma_2$ であり，x_H が正であることから $C > 0$ であること，そのことと式 (A.4) により $\alpha_1 > \alpha_2 > 0$ であることもわかる．

連立式 (13) の第 2 番目の式と式 (A.4) と式 (A.5) と (A.6) を併せると，式 (A.7) が得られる．

$$\begin{aligned} x_L &= C\times\Big(\frac{\sigma_1^2}{1+\sigma_1} - \frac{\sigma_2^2}{1+\sigma_2}\Big) \\ &= \frac{(\mu+\sigma_1)(\mu+\sigma_2)}{\mu(\sigma_1+\sigma_2)+\sigma_1\sigma_2}\times\frac{(\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(1+\sigma_1)(1+\sigma_2)} \\ &= x_H\times\frac{(\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(1+\sigma_1)(1+\sigma_2)}\,. \end{aligned} \tag{A.7}$$

式 (8) に式 (A.4) と式 (A.6) を併せると，式 (A.8) が得られる．

$$\begin{aligned} x(t) &= \frac{(\mu+\sigma_1)(\mu+\sigma_2)}{\mu(\sigma_1+\sigma_2)+\sigma_1\sigma_2}\times\frac{t(\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(t+\sigma_1)(t+\sigma_2)} \\ &= x_H\times\frac{t(\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(t+\sigma_1)(t+\sigma_2)}\,. \end{aligned} \tag{A.8}$$

すると与えられた 3 つの値 μ，x_H，x_L の組から以下の手順で σ_1 と σ_2 の値が求められる．まず (A.6) と (A.7) から以下の関係 (A.9) が得られる．

$$\begin{cases} \dfrac{1}{x_H} &= \dfrac{(\mu+\sigma_1)(\mu+\sigma_2)-\mu^2}{(\mu+\sigma_1)(\mu+\sigma_2)} = 1 - \dfrac{\mu^2}{(\mu+\sigma_1)(\mu+\sigma_2)}\,, \\ \dfrac{x_L}{x_H} &= \dfrac{(1+\sigma_1)(1+\sigma_2)-1}{(1+\sigma_1)(1+\sigma_2)} = 1 - \dfrac{1}{(1+\sigma_1)(1+\sigma_2)}\,. \end{cases} \tag{A.9}$$

この式 (A.9) を式 (A.10) の置き換えを用いて書き直すと式 (A.11) が得られる．

$$\begin{cases} p \equiv \dfrac{\mu^2 x_H}{x_H-1}\,, \\ q \equiv \dfrac{x_H}{x_H-x_L}\,. \end{cases} \tag{A.10}$$

$$\begin{cases} (\mu+\sigma_1)(\mu+\sigma_2) = p\,, \\ (1+\sigma_1)(1+\sigma_2) = q\,. \end{cases} \tag{A.11}$$

さらに式 (A.11) を少し変形して，式 (A.12) を得る．

$$\begin{cases} \sigma_1\sigma_2 + \mu(\sigma_1+\sigma_2) = p-\mu^2\,, \\ \sigma_1\sigma_2 + (\sigma_1+\sigma_2) = q-1\,. \end{cases} \tag{A.12}$$

式 (A.12) は σ_1 と σ_2 の基本対称式である $S_1 \equiv \sigma_1+\sigma_2$ と $S_2 \equiv \sigma_1\sigma_2$ についての連立 1 次方程式であり，それを解いて式 (A.13) が得られる．

$$\begin{cases} S_1 = \dfrac{p-q}{\mu-1} - (\mu+1)\,, \\ S_2 = \mu + \dfrac{\mu\,q-p}{\mu-1}\,. \end{cases} \tag{A.13}$$

すると 2 次方程式 (A.14) の 2 根が相異なる正の実数である場合に限り，それらは σ_1 と σ_2（$\sigma_1 > \sigma_2 > 0$）である．

$$w^2 - S_1 w + S_2 = 0\,. \tag{A.14}$$

この 2 次方程式 (A.14) の 2 根が相異なる正の実数であるための必要十分条件は，$S_1 > 0$ かつ $S_2 > 0$ かつ $D \equiv S_1^2 - 4S_2 > 0$ である．

こうして相異なる正の実数 σ_1 と σ_2 が求まれば，式 (A.6) と式 (A.4) から得られる式 (A.15) を順に計算することで，2 つの極の係数 α_1 と α_2 がそれぞれ求まる．

$$\begin{cases} C \leftarrow \dfrac{x_H}{\sigma_1-\sigma_2}\,, \\ \alpha_1 \leftarrow C\sigma_1^2\,, \\ \alpha_2 \leftarrow C\sigma_2^2\,. \end{cases} \tag{A.15}$$

以上が式 (8) の $x(t)$ を決定する手順である．

A.4.2 「方式 II」の場合

まず連立方程式 (16) の第 1 番目と第 2 番目の式をあわせると式 (A.16) の最初の等式が得られる．その等式の値は t にはよらない定数なので，それを C とおく．

$$\frac{\alpha_1}{\sigma_1(1+\sigma_1)} = \frac{\alpha_2}{\sigma_2(1+\sigma_2)} = C\,. \tag{A.16}$$

すると極の係数はそれぞれ式 (A.17) で表される．

$$\begin{cases} \alpha_1 = C\sigma_1(1+\sigma_1), \\ \alpha_2 = C\sigma_2(1+\sigma_2). \end{cases} \tag{A.17}$$

連立式 (16) の第 3 番目の式から式 (A.17) を用いて α_1 と α_2 を消去すれば，式 (A.18) が得られる．

$$1 = \frac{\alpha_1}{\mu+\sigma_1} - \frac{\alpha_2}{\mu+\sigma_2} = C\left\{\frac{\sigma_1(1+\sigma_1)}{\mu+\sigma_1} - \frac{\sigma_2(1+\sigma_2)}{\mu+\sigma_2}\right\}. \tag{A.18}$$

すると C の逆数の値は式 (A.19) で表される．

$$\begin{aligned} \frac{1}{C} &= \frac{\sigma_1(1+\sigma_1)}{\mu+\sigma_1} - \frac{\sigma_2(1+\sigma_2)}{\mu+\sigma_2} \\ &= (\sigma_1-\sigma_2) \times \frac{\mu(1+\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(\mu+\sigma_1)(\mu+\sigma_2)}\,. \end{aligned} \tag{A.19}$$

そうして x_L の値は式 (A.20) で表される．

$$\begin{aligned} x_L &= \frac{\alpha_1}{\sigma_1} - \frac{\alpha_2}{\sigma_2} \\ &= C \times (\sigma_1-\sigma_2) \\ &= \frac{(\mu+\sigma_1)(\mu+\sigma_2)}{\mu(1+\sigma_1+\sigma_2)+\sigma_1\sigma_2}\,. \end{aligned} \tag{A.20}$$

なお，この式 (A.20) から $C > 0$ であることがわかる．そのことと式 (A.17) をあわせると，$\alpha_1 > \alpha_2 > 0$ であることもわかる．

つぎに極大点の位置 t_p についての条件は連立式 (16) の第 5 番目の式から導かれる式 (A.21) である．

$$0 = \frac{\alpha_1}{(t_p+\sigma_1)^2} - \frac{\alpha_2}{(t_p+\sigma_2)^2} = C\left\{\frac{\sigma_1(1+\sigma_1)}{(t_p+\sigma_1)^2} - \frac{\sigma_2(1+\sigma_2)}{(t_p+\sigma_2)^2}\right\}. \tag{A.21}$$

そうして $t_p > 0$, $\sigma_1 > \sigma_2 > 0$ であることを用いて式 (A.21) の平方根を開くと，式 (A.22) の最初の等式が得られる．その等式の値は t によらない定数であるのでそれを Γ とおいた．

$$\frac{\sqrt{\sigma_1(1+\sigma_1)}}{t_p+\sigma_1} = \frac{\sqrt{\sigma_2(1+\sigma_2)}}{t_p+\sigma_2} = \Gamma\,. \tag{A.22}$$

すると $\sigma_1 \neq \sigma_2$ であるから，式 (A.23) が得られる．

$$\begin{aligned} \Gamma &= \frac{\sqrt{\sigma_1(1+\sigma_1)} - \sqrt{\sigma_2(1+\sigma_2)}}{\sigma_1-\sigma_2} \\ &= \frac{1+\sigma_1+\sigma_2}{\sqrt{\sigma_1(1+\sigma_1)} + \sqrt{\sigma_2(1+\sigma_2)}}\,. \end{aligned} \tag{A.23}$$

そうして，式 (A.22) から式 (A.24) が得られる．

$$\begin{cases} t_p+\sigma_1 = \dfrac{1}{\Gamma} \times \sqrt{\sigma_1(1+\sigma_1)}, \\ t_p+\sigma_2 = \dfrac{1}{\Gamma} \times \sqrt{\sigma_2(1+\sigma_2)}\,. \end{cases} \tag{A.24}$$

式 (A.24) の中の 2 つの式を連立させると，t_p を表す式 (A.25) が得られる．

$$\begin{aligned} t_p &= \frac{1}{\Gamma} \times \frac{\sigma_1\sqrt{\sigma_2(1+\sigma_2)} - \sigma_2\sqrt{\sigma_1(1+\sigma_1)}}{\sigma_1-\sigma_2} \\ &= \frac{1}{\Gamma} \times \frac{\sigma_1\sigma_2}{\sigma_1\sqrt{\sigma_2(1+\sigma_2)} + \sigma_2\sqrt{\sigma_1(1+\sigma_1)}}\,. \end{aligned} \tag{A.25}$$

よって，式 (A.24)，式 (A.16)，式 (A.23)，式 (A.20) を用いて，式 (16) の第 4 番目である x_H を表す式を書き換えると，式 (A.26) が得られる．

$$\begin{aligned} x_H &= \frac{\alpha_1}{t_p+\sigma_1} - \frac{\alpha_2}{t_p+\sigma_2} \\ &= \Gamma\left\{\frac{\alpha_1}{\sqrt{\sigma_1(1+\sigma_1)}} - \frac{\alpha_2}{\sqrt{\sigma_2(1+\sigma_2)}}\right\} \\ &= C\,\Gamma\left\{\sqrt{\sigma_1(1+\sigma_1)} - \sqrt{\sigma_2(1+\sigma_2)}\right\} \\ &= C\,\Gamma^2 \times (\sigma_1-\sigma_2) \\ &= \Gamma^2 x_L\,. \end{aligned} \tag{A.26}$$

すると式 (A.26)，式 (A.23)，および (A.20) を用いて，式 (A.27) が導かれる．

$$\begin{cases} \dfrac{1}{\Gamma} = \sqrt{\dfrac{x_L}{x_H}} = \dfrac{\sqrt{\sigma_1(1+\sigma_1)} + \sqrt{\sigma_2(1+\sigma_2)}}{1+\sigma_1+\sigma_2}, \\ \dfrac{1}{x_L} = \dfrac{\mu(1+\sigma_1+\sigma_2)+\sigma_1\sigma_2}{(\mu+\sigma_1)(\mu+\sigma_2)} = 1 - \dfrac{\mu(\mu-1)}{(\mu+\sigma_1)(\mu+\sigma_2)}\,. \end{cases} \tag{A.27}$$

この式 (A.27) から σ_1 と σ_2 についての連立方程式 (A.28) が得られるので，それを解いて σ_1 と σ_2 を求めればよい (ただし $\sigma_1 > \sigma_2 > 0$ である)．

$$\begin{cases} \dfrac{\sqrt{\sigma_1(1+\sigma_1)} + \sqrt{\sigma_2(1+\sigma_2)}}{1+\sigma_1+\sigma_2} = \sqrt{\dfrac{x_L}{x_H}}, \\ (\mu+\sigma_1)(\mu+\sigma_2) = \mu(\mu-1) \times \dfrac{x_L}{x_L-1}\,. \end{cases} \tag{A.28}$$

そこでいま (A.28) の上側の式に含まれる平方根をはずすために，式 (A.29) で表される変数の置換を行う．

$$\begin{cases} \sigma_1 \equiv \dfrac{z_1^2}{1-z_1^2}, \\ \sigma_2 \equiv \dfrac{z_2^2}{1-z_2^2}\,. \end{cases} \tag{A.29}$$

ただし $0 < z_1 < 1$，$0 < z_2 < 1$ で，さらに $\sigma_1 > \sigma_2$ であるから，$1 > z_1 > z_2 > 0$ である．

するとこの置換により (A.28) の上側の式の左辺は，式 (A.30) の右辺に書き換えられる．

$$\frac{\sqrt{\sigma_1(1+\sigma_1)} + \sqrt{\sigma_2(1+\sigma_2)}}{1+\sigma_1+\sigma_2} = \frac{z_1+z_2}{1+z_1z_2}\,. \tag{A.30}$$

よって (A.28) の上側の式は，式 (A.31) に書き換えられる．

$$z_1 + z_2 = (1 + z_1 z_2)\sqrt{\frac{x_L}{x_H}}\,. \tag{A.31}$$

(さらに $t_p = \dfrac{z_1 z_2}{1 + z_1 z_2}$ であることもわかる)．

いま S_1 と S_2 をそれぞれ式 (A.32) で表される z_1 と z_2 の基本対称式とする．

$$\begin{cases} S_1 \equiv z_1 + z_2\,, \\ S_2 \equiv z_1 z_2\,. \end{cases} \tag{A.32}$$

そうすると関係式 (A.31) は式 (A.33) に書き換えられる．

$$S_1 = (1 + S_2)\sqrt{\frac{x_L}{x_H}}\,. \tag{A.33}$$

さらに (A.28) の下側の式について，σ_1 と σ_2 を z_1 と z_2 を用いて書き換えて，式 (A.34) が得られる．

$$(z_1^2 - \kappa)(z_2^2 - \kappa) = \nu\kappa\,(z_1^2 - 1)(z_2^2 - 1)\,. \tag{A.34}$$

ただしここで導入した記号 κ と ν は，式 (A.35) で表されるものである．

$$\begin{cases} \kappa \equiv \dfrac{\mu}{\mu - 1}\,, \\ \nu \equiv \dfrac{x_L}{x_L - 1}\,. \end{cases} \tag{A.35}$$

式 (A.34) を整理することで式 (A.36) が得られる．

$$\eta_0 (z_1 z_2)^2 + \eta_1 (z_1^2 + z_2^2) + \eta_2 = 0\,. \tag{A.36}$$

ただしここで導入した 3 つの係数 η_0，η_1，η_2 はそれぞれ，式 (A.37) により与えられる．

$$\begin{cases} \eta_0 \equiv 1 - \nu\kappa, \\ \eta_1 \equiv (\nu - 1)\kappa, \\ \eta_2 \equiv (\kappa - \nu)\kappa\,. \end{cases} \tag{A.37}$$

式 (A.36) を z_1 と z_2 の基本対称式 (A.32) を用いて書き換えて，式 (A.38) が得られる．

$$\eta_0 {S_2}^2 + \eta_1 ({S_1}^2 - 2S_2) + \eta_2 = 0\,. \tag{A.38}$$

式 (A.33) の関係を用いて，式 (A.38) から S_1 を消去すると，S_2 についての 2 次方程式 (A.39) が得られる．

$$\zeta_0 {S_2}^2 + \zeta_1 S_2 + \zeta_2 = 0\,. \tag{A.39}$$

ただしこの方程式の各係数は式 (A.40) により与えられる．

$$\begin{cases} \zeta_0 \equiv \eta_0 + \dfrac{x_L}{x_H} \times \eta_1, \\ \zeta_1 \equiv 2\,(\,\dfrac{x_L}{x_H} - 1\,)\,\eta_1, \\ \zeta_2 \equiv \eta_2 + \dfrac{x_L}{x_H} \times \eta_1\,. \end{cases} \tag{A.40}$$

式 (A.39) の係数 ζ_0 と ζ_1 は共に負である．

実際 $\mu > 1$，$x_L > 1$ であることから，$\kappa \equiv \dfrac{\mu}{\mu - 1} > 1$，$\nu \equiv \dfrac{x_L}{x_L - 1} > 1$ であるから，$\nu\kappa > 1$ であるので，$\zeta_0 \equiv 1 - \nu\kappa < 0$ である．

また $\nu - 1 > 0$ であることから $\eta_1 \equiv (\nu - 1)\kappa > 0$ であり，さらに $x_H > x_L > 1$ より $\dfrac{x_L}{x_H} < 1$ であるから，$\dfrac{x_L}{x_H} - 1 < 0$ なので $\zeta_1 \equiv 2(\dfrac{x_L}{x_H} - 1)\eta_1 < 0$ である．

すると S_2 についての 2 次方程式 (A.39) が正根を持つためには ζ_2 が正であることが必要である．そうして ζ_2 が正であるときには，2 次方程式 (A.39) の判別式 $D_1 \equiv \zeta_1^2 - 4\zeta_0\zeta_2$ は正で実根は 2 つあるが，根と係数の関係から正根は単一である．

いま S_2 についての 2 次方程式 (A.39) が区間 $(0, 1)$ に入る実根を持つとする（そうでなければ σ_1 と σ_2 には適切な解は無い）．そのような S_2 が存在するとき，式 (A.33) を用いて S_2 から S_1 を作る．

そうして，2 次方程式 $w^2 - S_1 w + S_2 = 0$ の相異なる 2 つの実根がどちらも区間 $(0, 1)$ にあるとき，それらを z_1 と z_2（$1 > z_1 > z_2 > 0$）とする（そのような 2 つの根 z_1 と z_2 が無ければ，適切な σ_1 と σ_2 も存在しない）．そうして式 (A.29) の関係を用いて，z_1 と z_2 の値から σ_1 と σ_2 の値を求める．式 (A.41) を順に計算することで 2 つの極の係数 α_1 と α_2 が求まる．

$$\begin{cases} C \leftarrow \dfrac{x_L}{\sigma_1 - \sigma_2}\,, \\ \alpha_1 \leftarrow C\sigma_1(1 + \sigma_1)\,, \\ \alpha_2 \leftarrow C\sigma_2(1 + \sigma_2)\,. \end{cases} \tag{A.41}$$

以上の手順により，式 (8) の $x(t)$ が決定される．

参考文献

[1] 村上弘：レゾルベントの線形結合によるフィルタ対角化法，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.49, No.SIG 2(ACS21), pp.66–87 (2008).
[2] 村上弘：固有値が指定された区間内にある固有対を解くための対称固有値問題用のフィルタの設計，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.3, No.3(ACS31), pp.1–21 (2010).
[3] 村上弘：対称一般固有値問題のフィルタ作用素を用いた不変部分空間の近似構成，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.4, No.4 (ACS35), pp.1–14 (2011).
[4] 村上弘：レゾルベントを用いたフィルタによる固有値問題の解法について，**情報処理学会研究報告**，Vol.2012-HPC-133, No.22, pp.1–8 (2012).
[5] 村上弘：実対称定値一般固有値問題の最小側固有値を持つ固有対に対する実数シフトのレゾルベントを組み合わせたフィルタによる解法，**先進的計算基盤システムシンポジウム論文集** *2012*，pp.81–82 (2012).
[6] 村上弘：レゾルベントの線形結合をフィルタに用いたエルミート定値一般固有値問題のフィルタ対角化法，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.7, No.1 (ACS45), pp.57–72 (2014).
[7] 村上弘：レゾルベントの多項式をフィルタとして用いる対角化法について，**情報処理学会研究報告**，Vol.2014-HPC-146, No.13, pp.1–4 (2014).
[8] 村上弘：実対称定値一般固有値問題に対するレゾルベントの多項式によるフィルタの構成法の検討，**情報処理学会研究報告**，Vol.2014-HPC-147, No.2, pp.1–10 (2014).
[9] 村上弘：実数シフトのレゾルベントを組み合わせたフィルタによる実対称定値一般固有値問題の下端付近の固有値を持つ固有対の解法，*HPCS2015* **シンポジウム論文集**，Vol.2015, pp.38–51 (2015).
[10] Anthony P. Austin and Lloyd N. Trefethen: "Computing Eigenvalues of Real Symmetric Matrices with Rational Filters in Real Arithmetic", SIAM J. Sci. Comput., Vol.37, No.3, pp.A1365-1387 (2015).
[11] 村上弘：一つのレゾルベントから構成されたフィルタを用いた実対称定値一般固有値問題に対するフィルタ対角化法の実験，**情報処理学会研究報告**，Vol.2015-HPC-149, No.7, pp.1–16 (2015).
[12] 村上弘：実対称定値一般固有値問題の最小側固有対を解くための実数シフトのレゾルベントの多項式によるフィルタの簡易な設計法，**情報処理学会研究報告集**，Vol.2016-HPC-155, No.44, pp.1–27 (2016).
[13] 村上弘：レゾルベントの多項式によるフィルタを用いた実対称定値一般固有値問題の解法，**情報処理学会研究報告集**，Vol.2016-HPC-157, No.4, pp.1–15 (2016).
[14] 村上弘：チェビシェフ展開形で表わされたレゾルベントの多項式によるフィルタの伝達特性の調整，**数理解析研究所講究録**，No.2019, pp.96–112 (2017).
[15] 村上弘：実対称定値一般固有値問題を解くためのレゾルベントの多項式型フィルタの設計について，**情報処理学会研究報告**，Vol.2017-HPC-158, No.7, pp.1–10 (2017).
[16] 村上弘：実対称定値一般固有値問題を解くための少数のレゾルベントの多項式を用いたフィルタの設計法，**情報処理学会研究報告**，Vol.2017-HPC-159, No.4, pp.1–13 (2017).
[17] 村上弘：少数のレゾルベントから構成されたフィルタを用いた実対称定値一般固有値問題の解法，**情報処理学会研究報告**，Vol.2017-HPC-160, No.32, pp.1–32 (2017).
[18] 村上弘：少数のレゾルベントで構成された多項式型フィルタによる対称定値一般固有値問題の解法，**情報処理学会研究報告**，Vol.2017-HPC-161, No.7, pp.1–13 (2017).
[19] 村上弘：少数のレゾルベントから構成されたフィルタを用いた対称定値一般固有値問題の解法，**情報処理学会研究報告**，Vol.2017-HPC-162, No.21, pp.1–34 (2017).
[20] 村上弘：少数のレゾルベントの多項式型フィルタを用いた一般固有値問題の解法，**情報処理学会研究報告**, Vol.2018-HPC-165, No.15, pp.1–21 (2018).
[21] 村上弘：フィルタにレゾルベントの線形結合の多項式を用いた複素エルミート定値一般固有値問題の解法，**情報処理学会研究報告**，Vol.2018-HPC-166, No.10, pp.1–17 (2018).
[22] 村上弘：フィルタ対角化法による近似固有対の精度の改良について，**情報処理学会研究報告**，Vol.2018-HPC-167, No.29, pp.1–31 (2018).
[23] 村上弘：単一のレゾルベントのチェビシェフ多項式による実対称定値一般固有値問題の解法用の簡易型フィルタ，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.12, No.2 (ACS64), pp.1–26 (2019).
[24] 村上弘：フィルタ対角化法による固有値問題の近似対の改良，**情報処理学会研究報告**，Vol.2019-HPC-168, No.18, pp.1–36 (2019).
[25] 村上弘：直交化付きフィルタ適用による固有値問題の近似対の反復改良について，**情報処理学会研究報告**，Vol.2019-HPC-169, No.1, pp.1–31 (2019).
[26] Hiroshi Murakami: Filters consist of a few resolvents to solve real symmetric-definite generalized eigenproblems, *JJIAM*, Vol.36, No.2, pp.579–618 (2019).
[27] 村上弘：フィルタの反復適用による実対称定値一般固有値問題の近似対の改良，**情報処理学会論文誌：コンピューティングシステム** (ACS)，Vol.12, No.3(ACS65), pp.14–33 (2019).
[28] 村上弘：少数のレゾルベントの線形結合の多項式をフィルタとして用いた実対称定値一般固有値問題の解法，**情報処理学会研究報告**，Vol.2019-HPC-171, No.7, pp.1–45 (2019).
[29] 村上弘：少数のレゾルベントで構成されたフィルタを用いた実対称定値一般固有値問題の解法，**情報処理学会論文誌：コンピューティングシステム** (ACS67)，Vol.13, No.1, pp.1–27 (2020).

A.5 フィルタを実数シフトの単一のレゾルベントから構成する設計法の追補

既に副節 2.7 でフィルタを単一の実数シフトのレゾルベントの Chebyshev 多項式とする場合の伝達関数の設計について記述したが，ここでは 4 つのパラメアのうち 3 つだけを指定した場合にフィルタの性質が最も望ましくなるように決める方法も含めて，より詳しく扱うことにする．

伝達関数 $g(t)$ は前と同じ式 (31) で表されるとする．そうして伝達関数は条件 $g(0)=1$，$g(1)=g_p$，$g(\mu)=g_s$ を満たすとする．ただし $1<\mu$，$1>g_p>g_s>0$ である．

式 (A.42) で y_H と y_L を定義する．

$$\begin{cases} y_H \equiv \cosh\left(\dfrac{1}{n}\cosh^{-1}\dfrac{1}{g_s}\right), \\ y_L \equiv \cosh\left(\dfrac{1}{n}\cosh^{-1}\dfrac{g_p}{g_s}\right). \end{cases} \tag{A.42}$$

すると 1 次の有理関数 $y(t)$ は式 (A.43) を満たす．

$$\begin{cases} t=0 \quad : \quad \dfrac{\alpha}{\sigma}+\beta=y_H, \\ t=1 \quad : \quad \dfrac{\alpha}{\sigma+1}+\beta=y_L, \\ t=\mu \quad : \quad \dfrac{\alpha}{\sigma+\mu}+\beta=1. \end{cases} \tag{A.43}$$

ただし，$1>g_p>g_s>0$ であることから $y_H>y_L>1$ であり，$t\geq 0$ で $g(t)$ が極を持たないためには $\sigma>0$ であり，さらに $t\in[0,\mu)$ で $g(t)>0$ が n の偶奇に依らずに成り立つように α を正とする．すると $y(t)$ は正の実軸上で単調減少関数であるから，$t\in[\mu,\infty)$ で $|g(t)|\leq g_s$ を常に満たすように β の値を $[-1,1)$ の範囲に制限する．

A.5.1 4 つのパラメタ n，μ，g_s，g_p を指定した場合の設計

式 (A.42) を用いると，この場合には n，μ，y_H，y_L を指定したことになる．

いま $\mu>1$ かつ $y_H>y_L>1$ であるときに，連立方程式 (A.43) を解けば式 (A.44) が得られる．

$$\begin{cases} \sigma=1\Big/\left(\dfrac{\mu-1}{\mu}\times\dfrac{y_H-1}{y_L-1}-1\right), \\ \alpha=(y_H-y_L)\Big/\left(\dfrac{1}{\sigma}-\dfrac{1}{\sigma+1}\right), \\ \beta=1-\dfrac{\alpha}{\sigma+\mu}. \end{cases} \tag{A.44}$$

伝達関数が構成可能になるためには，σ が正でかつ $-1\leq\beta<1$ であることが必要である．

A.5.2 3 つの値 n と μ と比 g_p/g_s を指定して，g_s を最小にする設計

式 (A.42) を用いると，この場合には n，μ，y_L を指定して，y_H を最小に選ぶことになる．

連立方程式 (A.43) から 2 つの式の組 (A.45) が得られる．

$$\begin{cases} y_H-1=\alpha\times(\dfrac{1}{\sigma}-\dfrac{1}{\sigma+\mu}), \\ y_L-1=\alpha\times(\dfrac{1}{\sigma+1}-\dfrac{1}{\sigma+\mu}). \end{cases} \tag{A.45}$$

それらの両辺の比をとることにより，式 (A.46) が得られる．

$$\frac{y_H-1}{y_L-1}=\frac{\mu}{\mu-1}\times\left(1+\frac{1}{\sigma}\right). \tag{A.46}$$

すると今の場合には y_L も μ も値が固定されているので，y_H を最小にするには σ を最大にすればよいことがわかる．連立方程式 (A.43) から以下の式の組 (A.47) が導かれる．

$$\begin{cases} (y_H-1)=\alpha\left(\dfrac{1}{\sigma}-\dfrac{1}{\sigma+\mu}\right), \\ y_H-\beta=\dfrac{\alpha}{\sigma}. \end{cases} \tag{A.47}$$

それらから式変形により，以下の式が得られる．

$$\frac{\sigma+1}{\mu-1}=\frac{1-\beta}{y_L-1}. \tag{A.48}$$

すると，σ が最大になるのは β が最小の値をとるとき，すなわち $\beta=-1$ のときであることがわかる．そうしてそのとき式 (A.49) が得られる．

$$\begin{cases} \sigma=\dfrac{2(\mu-1)}{y_L-1}-1, \\ \alpha=(y_L-1)\Big/\left(\dfrac{1}{\sigma+1}-\dfrac{1}{\sigma+\mu}\right), \\ y_H=\dfrac{\alpha}{\sigma}-1. \end{cases} \tag{A.49}$$

ただし $\sigma>0$ が満たされていることが必要であり，そうであれば $y(t)=\alpha/(t+\sigma)-1$ が決まる．

A.5.3 3 つのパラメタ n と μ と g_s を指定して，g_p を最大にする設計

式 (A.42) を用いると，この場合には n と μ と y_H だけを指定して，y_L を最大に選ぶことになる．

連立方程式 (A.43) から式 (A.46) が導かれる．すると今の場合には μ と y_H の値は固定されているので，y_L を最大にするには σ を最大にすればよいことがわかる．

連立方程式 (A.43) から導かれる 2 つの式の組 (A.47) の式の両辺の比をとると，以下の式が得られる．

$$\frac{\sigma+\mu}{\mu}=\frac{y_H-\beta}{y_H-1}. \tag{A.50}$$

今の場合には μ と y_H の値は固定されているから，上式から σ の値を最大にするためには β の値をその最小値 -1 にとればよいことがわかる．すると式 (A.51) が得られる．

$$\begin{cases} \sigma=\dfrac{2\mu}{y_H-1}, \\ \alpha=2(\sigma+\mu)=2\mu\times\dfrac{y_H+1}{y_H-1}, \\ y_L=\dfrac{\alpha}{\sigma+1}-1. \end{cases} \tag{A.51}$$

これにより $y(t)=\alpha/(t+\sigma)-1$ が決まる．

A.5.4 3 つのパラメタ n と g_s と g_p を指定して，μ を最小にする設計

式 (A.42) を用いると，この場合には y_H と y_L を指定して，μ を最小に選ぶことになる．

式 (A.46) を少し変形すると式 (A.52) が得られる．

$$\left(1+\frac{1}{\sigma}\right) = \frac{y_H-1}{y_L-1} \times \left(1-\frac{1}{\mu}\right). \tag{A.52}$$

この式から，今の場合には y_H も y_L も値が固定されているので，μ が最小になるのは σ が最大になるときである．そこで β を表す式 (A.53) を変形する．

$$\beta = \left(\frac{y_L}{\sigma} - \frac{y_H}{\sigma+1}\right) \Big/ \left(\frac{1}{\sigma} - \frac{1}{\sigma+1}\right), \tag{A.53}$$

それにより関係式 (A.54) が得られる．

$$\sigma = \frac{y_L-\beta}{y_H-y_L}. \tag{A.54}$$

この式から，σ は $\beta=-1$ のときに最大値をとり，その値は式 (A.55) で与えられる．

$$\sigma = \frac{y_L+1}{y_H-y_L}. \tag{A.55}$$

そのときの α と μ の値は式 (A.56) で与えられる．

$$\begin{cases} \alpha = 1\Big/\left(\dfrac{1}{y_L+1} - \dfrac{1}{y_H+1}\right) = \dfrac{(y_H+1)(y_L+1)}{y_H-y_L}, \\ \mu = \dfrac{(y_L+1)(y_H-1)}{2(y_H-y_L)}. \end{cases} \tag{A.56}$$

不等式 (A.57) により，この μ の最小値は 1 よりも必ず大きい．

$$\mu - 1 = \frac{(y_L-1)(y_H+1)}{2(y_H-y_L)} > 0. \tag{A.57}$$

そうして $y(t) = \alpha/(t+\sigma) - 1$ と決まる．

A.5.5 この節のまとめ

A.5.5.1 4 つのパラメタ n，μ，g_s，g_p を指定した場合の設計

まず式 (A.42) を用いて y_H と y_L を求めてから，以下を計算する．

$$\begin{cases} D \leftarrow \mu(y_H-y_L) - (y_H-1), \\ \beta \leftarrow \{\mu(y_H-y_L) - y_L(y_H-1)\}/D, \\ \sigma \leftarrow \mu(y_L-1)/D, \\ \alpha \leftarrow \mu(y_L-1)/D \times (\mu-1)(y_H-1)(y_H-y_L)/D. \end{cases}$$

この場合は条件 $D>0$ と $-1 \le \beta < 1$ の両方が成立するときにだけ適切な実現が可能である．

A.5.5.2 3 つの量 n と μ と比 g_p/g_s を指定して，g_s を最小にする設計

まず式 (A.42) を用いて y_H の値を求めてから，以下を計算する．

$$\begin{cases} \beta \leftarrow -1, \\ \sigma \leftarrow \dfrac{(2\mu-1)-y_L}{y_L-1}, \\ \alpha \leftarrow \dfrac{2(\mu-1)(y_L+1)}{y_L-1}, \\ y_H \leftarrow \dfrac{(2\mu-1)y_L-1}{(2\mu-1)-y_L}, \\ g_s \leftarrow 1/\cosh\{n\cosh^{-1}(y_H)\}. \end{cases}$$

この場合は条件 $\sigma>0$ が成立するときにだけ適切な実現が可能である．

A.5.5.3 n と μ と g_s を指定して，g_p を最大にする設計

まず式 (A.42) を用いて y_H の値を求めてから，以下を計算する．

$$\begin{cases} \beta \leftarrow -1, \\ \sigma \leftarrow \dfrac{2\mu}{y_H-1}, \\ \alpha \leftarrow \dfrac{2\mu}{y_H-1} \times (y_H+1), \\ y_L \leftarrow \dfrac{(2\mu-1)y_H+1}{(2\mu-1)+y_H}, \\ g_p \leftarrow g_s\cosh\{n\cosh^{-1}(y_L)\}. \end{cases}$$

この場合は常に適切な実現が可能である．

A.5.5.4 3 つのパラメタ n と g_s と g_p を指定して，μ を最小にする設計

まず式 (A.42) を用いて y_H と y_L の値を求めてから，以下を計算する．

$$\begin{cases} \beta \leftarrow -1, \\ \sigma \leftarrow \dfrac{y_L+1}{y_H-y_L}, \\ \alpha \leftarrow \dfrac{y_L+1}{y_H-y_L} \times (y_H+1), \\ \mu \leftarrow \dfrac{y_L+1}{y_H-y_L} \times \dfrac{y_H-1}{2}. \end{cases}$$

この場合は常に適切な実現が可能である．

A.6 フィルタを実数シフトのレゾルベント 2 つで構成する別の設計手法

実数シフトのレゾルベント 2 つと恒等作用素の線形結合の Chebyshev 多項式をフィルタとする場合の伝達関数の設計については，以前の節 2 で既に示したが，この節では，有理関数補間の方法を用いた手法で若干の拡張を試みる．

いま伝達関数 $g(t)$ は式 (A.58) の形で表されるとする．ここで σ_1，σ_2 は相異なる正の実数で，α_1，α_2，β は実数である．以前の節 2 の場合は $\beta=-1$ に限定したことになる．

$$\begin{cases} g(t)=g_s T_n(y(t)), \\ y(t)=\dfrac{\alpha_1}{t+\sigma_1}-\dfrac{\alpha_2}{t+\sigma_2}+\beta,. \end{cases} \tag{A.58}$$

伝達関数がこの形をしていれば，それに対応するフィルタはシフトが最小固有値よりも小さい実数とする 2 つのレゾルベントの作用と恒等作用素の線形結合の n 次 Chebyshev 多項式として実現できることがわかる．

いま，$\mu>1$ で，伝達関数の 2 つの閾値である正の実数 g_s，g_p は $0<g_s<g_p<1$ を満たしているとする．関数 $g(t)$ は $[0,1]$ を通過域，$(1,\mu)$ を遷移域，$[\mu,\infty)$ を阻止域として，関数値は通過域では区間 $[g_p,1]$ に含まれ，遷移域では (g_s,g_p) に含まれ，阻止域では $[-g_s,g_s]$ に含まれるものとする．そうして，伝達関数が通過域の中で最大値 1 をとる点を z とする．すると $g(t)$ についての条件 $g(z)=1$，$g'(z)=0$，$g(1)=g_p$，$g(\mu)=g_s$ が得られる．

いま y_H と y_L を前節と同じ式 (A.42) で定義しておく．すると，関数 $y(t)$ については以下の条件を課すことになる．

$$\begin{cases} t=z & : \quad y(z)=y_H,\ y'(z)=0, \\ t=1 & : \quad y(1)=y_L, \\ t=\mu & : \quad y(\mu)=1\,. \end{cases} \tag{A.59}$$

さらに $t=\infty$ で $y(\infty)=\beta$ と置く

関数 $y(t)$ のとる値の範囲は通過域では $[y_L,y_H]$ であり，遷移域では $(1,y_L)$ であり，阻止域では $[-1,1]$ である．以降で $y(0)$ の値を指定していない場合も，不等式 $y_L\le y(0)$ は満たす必要がある．

いま 3 点 (z,y_H)，$(1,y_L)$，$(\mu,1)$ を通り，$t=z$ では微分の値が零である高々 3 次の有理補間関数を $R_1(t)$ とする．有理関数 $R_1(t)$ の重心形式による表現は式 (A.60) になる．そうして導入された 3 つの未定パラメタ d_z，c_z，c_1 を決定する．ただし，d_z と c_1 はどちらも零ではないと仮定しておく．するとその仮定の下では $R_1(t)$ が式 (A.59) の 4 条件を最初から満たすことを容易に確認できる．

$$R_1(t)=\frac{\left\{\dfrac{d_z}{(t-z)^2}+\dfrac{c_z}{t-z}\right\}y_H+\dfrac{c_1}{t-1}\,y_L+\dfrac{1}{t-\mu}}{\left\{\dfrac{d_z}{(t-z)^2}+\dfrac{c_z}{t-z}\right\}+\dfrac{c_1}{t-1}+\dfrac{1}{t-\mu}}\,. \tag{A.60}$$

この $R_1(t)$ を t の多項式 $N_1(t)$ と $D_1(t)$ の比の形として $R_1(t)=N_1(t))/D_1(t)$ と書けば，それらの多項式は高々 3 次になる．それらの多項式の 3 次項の係数を両方とも零にする条件を求めると，パラメタ c_z と c_1 についての 2 連立線形方程式になり，それを解くと式 (A.61) が得られる．

$$c_z=\frac{y_L-1}{y_H-y_L},\ c_1=-\frac{y_H-1}{y_H-y_L}\,. \tag{A.61}$$

それを用いて $N_1(t)$ と $D_1(t)$ から c_z と c_1 を含まない多項式 $N_2(t)$ と $D_2(t)$ が得られ，それに対応する t の 2 次の有理関数は $R_2(t)\equiv N_2(t)/D_2(t)$ である．多項式 $N_2(t)$ と $D_2(t)$ はパラメタ d_z を含むが，無限遠における条件 $R_2(\infty)=\beta$，つまり $N_2(t)$ と $D_2(t)$ の 2 次の項の係数の比が β である条件を課せば，それも d_t についての 1 次方程式に帰着して，それを解くと式 (A.62) が得られる．

$$d_z=\frac{y_H\{\beta-z+\mu(1-\beta)\}-y_L\{y_H(1-z)+z\beta-1+\mu(1-\beta)\}-\beta(1-z)}{(y_H-y_L)(\beta-y_H)}\,. \tag{A.62}$$

これを $R_2(t)$ と $D_2(t)$ の d_t に代入して多項式 $N_3(t)$ と $D_3(t)$ を作ると，$R_3(t)\equiv N_3(t)/D_3(t)$ は 4 つの条件 (A.59) の他に，$R_3(t)=\beta$ という条件も満たす t の有理関数になる．

こうして多項式 $N_3(t)$ と $D_3(t)$ が含む 5 つの実数の量 (μ，y_H，y_L，β，z) を指定すると，実数係数の 2 次多項式 $N(t)$ と $D(t)$ が四則演算だけで構成できる．さらに $y(0)\ge y_L$，つまり $N(t)$ と $D(t)$ の定数項の比の値が y_L 以上である必要がある．あとは分母の 2 次多項式 $D(t)$ の零点 2 つが負の実数であるときに限って，$y(t)\leftarrow N(t)/D(t)$ とする．

A.6.1 6 つのパラメタ n，μ，g_s，g_p，β，z を与えて $y(t)$ を求める場合

ここでは上記の記述に従って，パラメタとして n，μ，g_s，g_p，β，z の 6 つを与え，それらを満たせる伝達関数が構成可能な場合にはそれを具体的に求める計算法を示す．前提として，$n>1$，$\mu>1$，$0<g_s<g_p<1$，$\beta\in[-1,1)$，$z\in[0,1)$ であるとする．

まず y_H と y_L を式 (A.42) の右辺を計算して求める．すると $1<y_L<y_H$ である．

式 (A.63) で与えられる 2 次の有理関数 $R(t)$ が，5 つの条件 $R(z)=y_H$，$R'(z)=0$，$R(1)=y_L$，$R(\mu)=1$，$R(\infty)=\beta$ を満たすとする．ここで z は区間 $[0,1]$ 内の $R(t)$ の最大点の t 座標である．

そのとき，式 (A.63) の中の 6 つの係数 p_2，p_1，p_0，q_2，q_1，q_0 のそれぞれは（それら 6 つに共通の定数を乗じる違いを無視すると）5 つのパラメタ y_H，y_L，β，μ，z を用いて式 (A.64) で与えられることが計算で示せる．

$$R(t)\equiv\frac{N(t)}{D(t)}=\frac{p_2t^2+p_1t+p_0}{q_2t^2+q_1t+q_0}\,. \tag{A.63}$$

$$
\begin{cases}
p_2 = \beta(y_H-1)(y_H-y_L)(\mu-1), \\
p_1 = y_H^2(\mu-z)^2 - y_H y_L\,\big[\, y_H(1-z)^2 \\
\quad +(\mu-1)\{\,2(1-z)+(\mu-1)\,\}\,\big] - \beta\times \\
\quad \big[\, y_L\big\{ -y_H\{(1-z)^2+(\mu^2-1)\}+2(\mu-1)z\big\} \\
\quad +y_H\big\{y_H(\mu^2-1)+\{(\mu-z)^2-(\mu^2-1)\}\big\}\,\big], \\
p_0 = -y_H^2(\mu-z)^2 \\
\quad +y_H y_L\,\{\,y_H(1-z)^2\mu+(\mu-z^2)(\mu-1)\,\} \\
\quad +\beta\big[\, y_H^2(\mu-1)\mu + y_H\{\,(\mu-z)^2-(\mu-1)\mu\,\} \\
\quad +y_L\big\{ -y_H\{(\mu-1)z^2+(\mu-z)^2\}+(\mu-1)z^2\big\}\,\big], \\
q_2 = (y_H-1)(y_H-y_L)(\mu-1), \\
q_1 = -y_H^2\times 2(\mu-1)z + y_H\{\,(1-z)^2+(\mu^2-1)\,\} \\
\quad -y_L\big[\, y_H\{\,(\mu-z)^2-(\mu^2-1)\,\}+(\mu^2-1)\,\big] \\
\quad +\beta\big[-y_H\big\{(\mu-1)+2(1-z)\big\}(\mu-1) \\
\quad +y_L(\mu-z)^2-(1-z)^2\,\big], \\
q_0 = y_H\big[\, y_H(\mu-1)z^2-\{(1-z)^2+(\mu-1)\}\mu\,\big] \\
\quad +y_L\big[\, y_H\{(\mu-z)^2-(\mu-1)\mu\}+(\mu-1)\mu\,\big] \\
\quad +\beta\big\{y_H(\mu-z^2)(\mu-1)-y_L(\mu-z)^2+\mu(z-1)^2\big\}.
\end{cases}
\tag{A.64}
$$

以前の節 2 の場合と違って，伝達関数の原点での値は指定していないが，条件 $g_p \le g(0)$ に対応するものとして，式 (A.64) で求めた p_0 と q_0 について不等式による条件 $R(0)=p_0/q_0 \ge y_L$ を満たすことは必要になる．

求めた式 (A.63) の分母の 2 次式 $D(t)$ が 2 つの負の実根 $-\sigma_1$，$-\sigma_2$ を持つためには式 (A.64) で求めた q_2，q_1，q_0 についての 3 つの不等式 $q_0>0$，$q_1>0$，$q_1^2-4q_2q_0>0$ がすべて成立することが必要十分条件である．そうなっていれば，式 (A.63) の実数の範囲での部分分数分解を行うことで式 (A.65) を得る．

$$
y(t) = \frac{\alpha_1}{t+\sigma_1} - \frac{\alpha_2}{t+\sigma_2} + \beta\,. \tag{A.65}
$$

パラメタ 6 つの組み合わせをうまく選んで，部分分数分解の式が実数の範囲で求まれば，シフトが実数であるレゾルベント 2 つを用いた簡易型のフィルタの伝達関数が構成できる．簡易型のフィルタの伝達関数は $g(t)=g_sT_n(y(t))$ であり，$y(t)$ の部分分数分解に対応して，シフトが（最小固有値未満の）実数であるレゾルベント 2 つの作用と恒等作用素の線形結合の n 次チェビシェフ多項式としてフィルタを実現できる．パラメタとして $z=0$ にとれば，伝達関数は原点で平坦で最大値をとる副節 2.1 の「方式-I」のフィルタの（$\beta=-1$ 以外を許すので）拡張を与える．

伝達関数を構成した例 1

この節で紹介した方法で伝達関数を構成してみた例をまず 1 つ示す（**図 A·40**）．用いた 6 つのパラメターは $n=35$，$\mu=1.75$，$g_s=$1E-10，$g_p=$1E-2，$\beta=-1$，$z=0.3$ である．

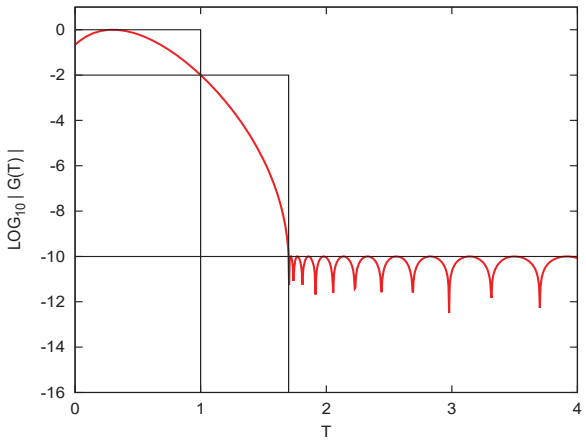

図 A·40　フィルタの伝達関数の対数プロット（例 1）

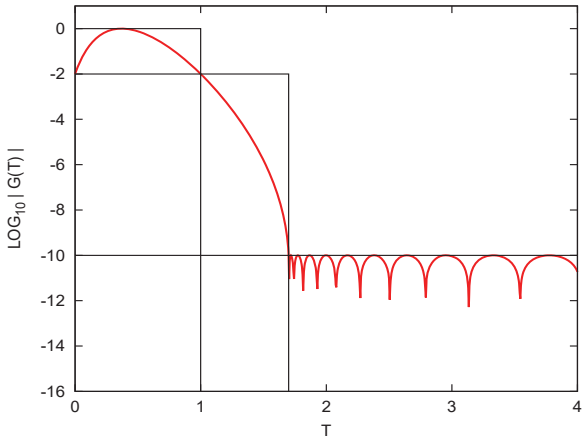

図 A·41　フィルタの伝達関数の対数プロット（例 2）

A.6.1.1　伝達関数を構成した例 2

この節で紹介した伝達関数を構成する方法において，有理関数 $R_2(t)$ に課した $g(\infty)=\beta$ に対応する条件を「方式 II」の $g(0)=g_p$ に対応する $R_2(0)=y_L$ に変更して $R(t)$ の係数を与える新たな数式を導ける（具体的な式の表示は省略）．これにより「方式 II」を少しだけ一般化できる．この方法で求めた伝達関数の例を示す．使用したパラメタは $n=35$，$\mu=1.7$，$g_s=$1E-10，$g_p=$1E-2，$z=0.37$ である（$y(t)$ の 2 つの極はどちらも負の実数で，$|y(\infty)|<1$ も満たしている）．有理関数 $y(t)$ の部分分数分解は $y(t)=\alpha_1/(t+\sigma_1)-\alpha_2/(t+\sigma_2)+\beta$ で $\alpha_1=10.22757865254875$，$\alpha_2=0.8193254233380222$，$\sigma_1=3.694115047029982$，$\sigma_2=0.7802911937373652$，$\beta=-0.5657280193766834$ となった．

A.6.1.2　3 次以上の有理関数 $y(t)$ の構成方法について

今回の本報告の有理関数補間法に基づく伝達関数の構成法と同様の方針により，$y(x)$ を 3 次の有理関数として，たとえば $0\le s<z<1$ で条件 $g(s)=g_p$，$g'(s)=0$，$g(z)=1$，$g'(z)=0$，$g(1)=g_p$，$g(\mu)=g_s$ をすべて満たすものを数式で決定できる．ただし，$y(t)$ の極 3 つがすべて負の実数になる適切なパラメタの組の例を現時点ではまだ見出せていない．