

チェビシェフ多項式と極がすべて実数である低次有理関数の合成により実対称定値一般固有値問題の少数の下側固有対を解くためのフィルタの伝達関数を構成する方法

村上 弘^{1,a)}

受付日 2022年6月10日, 採録日 2022年9月14日

概要: 行列の実対称定値一般固有値問題に対して, 固有値が下端付近にある比較的少数の固有対の近似をフィルタを用いて求める. 本論文ではフィルタは少数のレゾルベントの線型結合の作用の Chebyshev 多項式とする. するとフィルタの伝達関数は Chebyshev 多項式と低次の有理関数の合成であり, その有理関数は重複する極を持たない無限遠で有界なものになる. その有理関数の極をすべて最小固有値未満の実数にできれば, フィルタで用いるレゾルベントのシフトもすべて最小固有値未満の実数になり, レゾルベントの作用を与える連立 1 次方程式の係数行列はすべて実対称正定値になる. そこで本論文では, Chebyshev 多項式と低次の有理関数で極がすべて最小固有値未満の実数であるものの合成により, 特性がある程度良いフィルタの伝達関数をうまく構成する方法を検討する.

キーワード: 有理関数, フィルタ, 固有値問題, レゾルベント, 伝達関数, 実数シフト, Remez 法

A Method to Construct the Transfer Function of the Filter for the Solution of a Small Number of Lower Eigenpairs of a Real Symmetric-definite Generalized Eigenproblem by the Composition of a Chebyshev Polynomial and a Lower-Degree Rational Function Whose Poles Are All Reals

HIROSHI MURAKAMI^{1,a)}

Received: June 10, 2022, Accepted: September 14, 2022

Abstract: For a real symmetric-definite generalized eigenproblem of matrices, we use a filter to approximate a small number of eigenpairs whose eigenvalues are near the lower-end. In this paper, we assume the filter is a composition of a linear combination of a few resolvents and a Chebyshev polynomial. The transfer function which corresponds to the filter is a composition of a rational function of a lower degree and the Chebyshev polynomial, the rational function is finite at the infinity and has no poles duplicated. If all poles of the rational function are made real numbers below the minimum eigenvalue, then all shifts of resolvents are also real numbers below the minimum eigenvalue, and all coefficient matrices of systems of linear equations to give actions of resolvents are real symmetric positive-definite. Thus in this paper we studied the method to construct the filter's transfer function whose properties are moderately well as a composition of a rational function of the lower degree and a Chebyshev polynomial, and all poles of the rational function are below the minimum eigenvalue.

Keywords: rational function, filter, eigenproblem, resolvent, transfer function, real shift, remez method

¹ 東京都立大学数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan
University, Hachioji, Tokyo 192-0397, Japan

a) mrkhrsh@tmu.ac.jp

1. はじめに

行列の実対称定値一般固有値問題に対して, フィルタを利

用して固有値が下端付近にある少数の固有対の近似を求めることにする。我々はすでに文献 [4] において Chebyshev 多項式と単一のレゾルベントの線形結合の合成をフィルタとして用いる方法について研究した。本論文ではフィルタとして Chebyshev 多項式と少数のレゾルベントの線形結合の合成を用いることにする。そのようなフィルタの伝達関数は、Chebyshev 多項式と重複する極を持たない無限遠で有界な低次の有理関数を合成したものになる。そのような有理関数の極がすべて最小固有値よりも小さい実数であれば、すべてのレゾルベントのシフトも最小固有値よりも小さい実数になり、レゾルベントの作用を与える連立 1 次方程式の係数行列は実対称正定値になる。そこで本論文では、Chebyshev 多項式と極がすべて最小固有値よりも小さい実数である低次の有理関数を合成したものをフィルタの伝達関数として採用する。そうしてそのような伝達関数がある程度良い特性を持つようによく構成する方法について検討する。

なお、文献 [1] と [3] はフィルタとしてある程度の個数のレゾルベントの線形結合を用いる方法で、その方法では各レゾルベントのシフトは複素数であるが、本論文では Chebyshev 多項式とレゾルベントの線形結合を合成したものをフィルタにすることで、必要なレゾルベントの個数をごく少ない数にすることにより、レゾルベントの作用を実現する連立 1 次方程式を直接法を用いて解く場合の必要な計算資源の量を抑えることや、さらに下端付近の固有値を持つ固有対を求めることに特化することでレゾルベントのシフトをすべて実数にして、計算全体を実数の演算と記憶だけを用いてできるようにすることを狙ったものである。

フィルタを Chebyshev 多項式と複数のレゾルベントの線形結合を合成したものとする場合に、レゾルベントのシフトを複素数の範囲で選ぶ場合については、有理関数の最良近似理論に対応してレゾルベントの線形結合をバターワース型、チェビシェフ型、逆チェビシェフ型、楕円型として、それらに対応する (B-合成, C-合成, I-合成, E-合成) と呼ぶ 4 種類が構成可能である。それら 4 種類のうちで最初の 3 種類については論文 [5] において導出を示したが、その後の論文 [7] ではさらに 4 番目の種類である楕円型 (E-合成) についても追加で導出を具体的に示して、それを用いた実験の例を示した。これらの合成 (B, C, I, E) によるフィルタは、レゾルベントのシフトがすべて虚数になる場合には、中間固有値の固有対を求めるのに使用できる。レゾルベントの個数は少なくできるが、シフトが虚数のレゾルベントの作用を実現するためには複素数を用いた計算が必要になる。シフトを複素数の範囲で選べるのであれば、実現できるフィルタの特性は実数の範囲に制限している本論文の場合に比べて良いものになる。さらにフィルタのパラメタを指定したときにそれを実現する各レゾルベントのシフトと線形結合係数は、解析的な表式に数値を入

れて計算すれば求められる。

論文 [8] では、シフトが実数である単一のレゾルベントの Chebyshev 多項式、あるいはシフトが実数である 2 つのレゾルベントの線形結合の Chebyshev 多項式としてフィルタを構成して、それにより近似固有対を求める実験を試みた例である。レゾルベントの個数が 2 つの場合のフィルタの設計に現れる連立代数方程式は四則と開平だけを用いる代数的操作で解いたが、本論文の Remez 法類似の反復解法を用いた方法に比べて、解を計算するための式の導出の過程が長く見通しが悪い。また本論文の場合に比べて若干特別な状況の解に限定して求めていることがある。そのかわり、指定するパラメタの条件を満たすフィルタが存在するかどうかをよりはっきりと判断できる。

論文 [6] では、下端固有値の固有対を求めるためのフィルタとして Chebyshev 多項式をシフトが実数の単一のレゾルベントの作用と合成したものを、あるいは中間固有値の固有対を求めるためのフィルタとして Chebyshev 多項式をシフトが虚数の単一のレゾルベントの作用の実部と合成したものをそれぞれ用いて、それらをベクトルの組の再直交化とあわせて反復適用することにより必要な固有ベクトルをベクトルの組が含有する比率を高めることで近似固有対の精度改良を行う方法を記述し、その計算例を IEEE 754 の倍精度浮動小数点数 (binary64) を用いて示した。さらに論文 [9] では、同様の計算を IEEE 754 の単精度浮動小数点数 (binary32) を用いて行った例を示した。

2. 固有値問題の解法に用いるフィルタと有理関数

2.1 フィルタの構成

与えられた実対称定値一般固有値問題 $\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}$ (行列 \mathbf{A} と \mathbf{B} は実対称で、 \mathbf{B} は正定値) に対して、固有値が下端付近にある比較的少数の固有対の近似を構成する手段として、フィルタと呼ぶ線形作用素を用いることにする (フィルタは必要な固有ベクトルの張られる不変部分空間への射影演算子の近似である)。

本論文で扱うフィルタ \mathcal{F} は、Chebyshev 多項式 $T_n(x)$ と線形作用素 \mathcal{Y} の合成であるとする。ここで \mathcal{Y} は異なるシフト ρ_i を持つ少数 K 個のレゾルベント $\mathcal{R}(\rho_i)$ と恒等作用素 \mathcal{I} の線形結合である (式 (1))。ここで γ_i や β は線形結合の係数であり、いまの一般固有値問題に対してはシフト ρ のレゾルベント $\mathcal{R}(\rho)$ は $(\mathbf{A} - \rho\mathbf{B})^{-1}\mathbf{B}$ として定義される。

$$\begin{cases} \mathcal{F} &= g_s T_n(\mathcal{Y}), \\ \mathcal{Y} &= \sum_{i=1}^K \gamma_i \mathcal{R}(\rho_i) + \beta \mathcal{I}. \end{cases} \quad (1)$$

そのとき一般固有値問題の任意の固有対 (λ, \mathbf{v}) に対しては $\mathcal{F}\mathbf{v} = f(\lambda)\mathbf{v}$ が成り立つ。ここで $f(\lambda)$ はフィルタ \mathcal{F} の伝

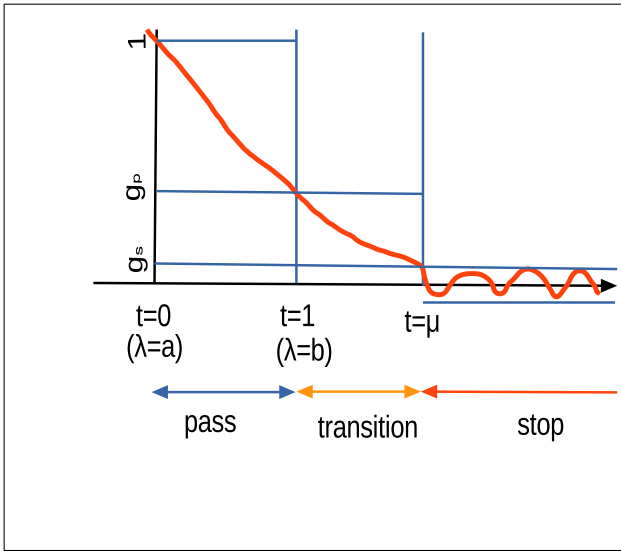


図 1 伝達関数 $g(t)$ の値の範囲の概形の例

Fig. 1 Rough sketch of the range of the transfer function $g(t)$ for each interval.

達関数であり、式 (2) で表されるように Chebyshev 多項式と有理関数 $\hat{y}(\lambda)$ の合成になる。

$$\begin{cases} f(\lambda) = g_s T_n(\hat{y}(\lambda)), \\ \hat{y}(\lambda) = \sum_{i=1}^K \frac{\gamma_i}{\lambda - \rho_i} + \beta. \end{cases} \quad (2)$$

便宜のために、固有値の座標 λ の下端付近の区間 $[a, b]$ と正規化座標と呼ぶ t の単位区間 $[0, 1]$ を線形変換 $\lambda = a + (b - a)t$ で対応させる。そうして引数の λ を正規化座標 t へと変更した伝達関数を $g(t) \equiv f(\lambda)$ とし、同様に $y(t) \equiv \hat{y}(\lambda)$ とすれば式 (3) の形になる。

$$\begin{cases} g(t) = g_s T_n(y(t)), \\ y(t) = \sum_{i=1}^K \frac{\alpha_i}{t - \tau_i} + \beta. \end{cases} \quad (3)$$

上記の過程を逆にたどることにより、式 (3) の形で表される関数 $g(t)$ に対しては、それを伝達関数として持つ式 (1) の形で表される線形作用素であるフィルタ \mathcal{F} を構成できる。

2.2 伝達関数 $g(t)$ の値の範囲に課す条件

正規化座標 t で表された伝達関数 $g(t)$ は $t \geq 0$ で連続であるとし、それを特徴づける最も主要な 3 つの形状パラメータを μ, g_s, g_p とする。ただし $1 < \mu, 0 < g_s \ll g_p < 1$ であるとする。そうして通過域、遷移域、阻止域と呼ばれる t のそれぞれの区間において条件 (4) を満たすものとする。すると $g(t)$ の連続性から $g(1) = g_p, g(\mu) = g_s$ となる必要がある (図 1)。

$$\begin{cases} \text{通過域 } [0, 1] & \text{では } g_p \leq g(t) \leq 1, \\ \text{遷移域 } (1, \mu) & \text{では } g_s < g(t) < g_p, \\ \text{阻止域 } [\mu, \infty) & \text{では } -g_s \leq g(t) \leq g_s. \end{cases} \quad (4)$$

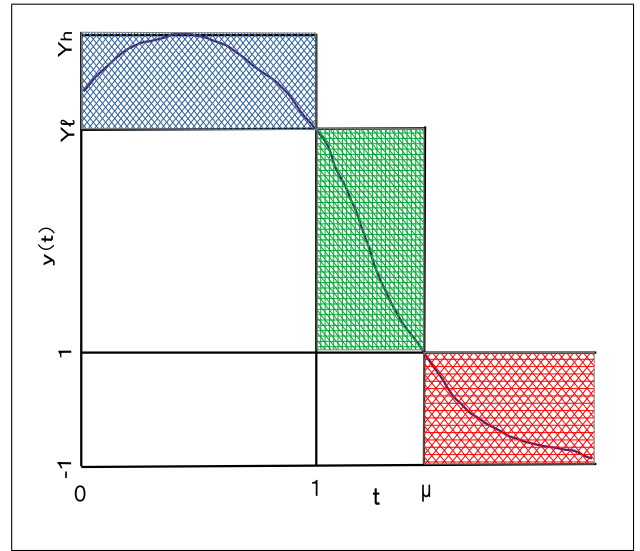


図 2 有理関数 $y(t)$ の値の範囲の概形

Fig. 2 Rough sketch of the range of the rational function $y(t)$ for each intervals.

有理関数 $y(t)$ もまた $t \geq 0$ で連続であるとする。そうして関係 $g = g_s T_n(y)$ に基づいて閾値 y_h と y_l を式 (5) で定義する。

$$\begin{cases} y_h \equiv \cosh\left(\frac{1}{n} \cosh^{-1} \frac{1}{g_s}\right), \\ y_l \equiv \cosh\left(\frac{1}{n} \cosh^{-1} \frac{g_p}{g_s}\right). \end{cases} \quad (5)$$

すると g の閾値と y の閾値の対応は式 (6) となる。

$$\begin{cases} g = 1 & \Rightarrow y = y_h, \\ g = g_p & \Rightarrow y = y_l, \\ g = g_s & \Rightarrow |y| = 1. \end{cases} \quad (6)$$

そうして通過域、遷移域、阻止域のそれぞれにおける有理関数 $y(t)$ の値の範囲についての条件は式 (7) になる。

$$\begin{cases} \text{通過域 } [0, 1] & \text{では } y_l \leq y(t) \leq y_h, \\ \text{遷移域 } (1, \mu) & \text{では } 1 < y(t) < y_l, \\ \text{阻止域 } [\mu, \infty) & \text{では } -1 \leq y(t) \leq 1. \end{cases} \quad (7)$$

有理関数 $y(t)$ の値の範囲についての概形を図 2 に示す。

2.3 レゾルベント $\mathcal{R}(\rho)$ のシフト ρ を実数にする利点

レゾルベントのシフトが実数である場合には、レゾルベントを与えられたベクトルに作用させる計算は実数の演算だけで行えて、さらに必要な記憶についても実数の値の格納だけで済む。連立 1 次方程式を解くなどの線形計算では乗算と加算の演算は通常は 1 対 1 の割合になる。すると複素数とその実部と虚部を実数の組で表して複素数の乗算と加減算を通常の方法により実現する場合には、実数の乗算と加減算の手間を等しいものとして扱う場合には複素数による線形計算の手間は実数による線形計算の手間の 4 倍に

なる。レゾルベントの作用を与える連立1次方程式の係数はシフト行列 $C(\rho) \equiv A - \rho B$ であり、それを行列分解を用いて解く場合にはシフト ρ が実数でかつ最小固有値未満のときは行列 $C(\rho)$ は実対称正定値になるので、ピボット選択をせずに実対称行列用の LDL^T 分解を用いて連立1次方程式を数値安定に解くことができる。シフト ρ が虚数であれば $C(\rho)$ は複素対称なので、複素対称行列用の LDL^T 分解を用いて解くことができる。

今回扱う対象には含めなかったが、一般固有値問題の係数行列の組が実対称定値の場合よりも一般的な複素 Hermite 定値の場合には、シフト ρ が実数でかつ最小固有値未満であればシフト行列 $C(\rho)$ は複素 Hermite 正定値になるので、ピボット選択をせずに複素 Hermite 行列用の LDL^H 分解を用いて連立1次方程式を数値安定に解くことができる。しかし ρ が虚数の場合には $C(\rho)$ は行列分解において利用可能な対称性がないので、たとえば複素 LU 分解を用いて（数値安定性が問題になればさらにピボット行交換付きで）解くことになり、演算量や記憶の量の面でかなり不利になる。

2.4 レゾルベント $\mathcal{R}(\rho)$ のシフト ρ を実数にする欠点

フィルタに用いるレゾルベントのシフトを複素数から実数に制限すると、それは選択範囲を狭めたことになり、実現可能なフィルタの特性は悪くなる。

シフトを実数に制限したことによるフィルタ特性の悪化を補うには、フィルタの構成に用いるレゾルベントの数を増す、用いる Chebyshev 多項式の次数を上げる、フィルタで濾過するベクトルの数を増す、フィルタを反復する回数を増す、ことが必要となる可能性が出てくる。すると、1つのレゾルベントの作用に対応する連立1次方程式を解く手間はシフトが虚数の場合よりもシフトを実数とする場合の方が少なくても、近似固有対のある要求された精度で求める場合の手間が少ないとは断定できない。そのため固有対に対する要求精度が高い場合には特に、シフトを実数に制限する場合としない場合とで性能の比較は簡単ではない。なお今回の論文では扱わないが、「中間固有値」の固有対を求める場合には、レゾルベントのシフトは実軸を避けてすべて虚数にする。

3. 実数シフトのレゾルベントを少数用いたフィルタの構成

今回の論文では、シフトを実数に制限したレゾルベントを少数個用いて特性がある程度良いフィルタをうまく構成することを目指した。

フィルタの特性を代表する3つの形状パラメタ μ , g_s , g_p については以下のことがいえる。パラメタ μ の値は1よりも大きい、 μ が1に近ければそれだけ遮断特性が急峻になるのでフィルタで濾過するベクトルの数は必要な固有

ベクトルの数に近くできる。 μ が1から離れるほど固有値が遷移域にある不要な固有ベクトルが混入してくる数が増すのでそれに見合って濾過するベクトルの数を多くする必要が出てくる。パラメタ g_s の値は微小で正であるが、零に近いほど阻止域に固有値を持つ固有ベクトルを阻止する能力が高いため近似固有対の精度が良くなる。パラメタ g_p の値はつねに1未満であるが、 g_p が1に近いほど通過域に固有値を持つ必要な固有ベクトル相互の伝達率の均一さが良くなるので、求めた近似固有対の間の精度のばらつきが減少することが期待できる。また、フィルタの構成に用いる Chebyshev 多項式の次数 n については、ベクトルの組にフィルタを1回適用する処理の内側ではベクトルの組にレゾルベントの線形結合を n 回（順次に）繰り返して適用するので、 n が大きいとそれだけフィルタを適用する際の計算量と経過時間は増加する。3つの形状パラメタ μ , g_s , g_p と次数 n の合計4つは、どれかを良いものにするとは劣化するトレードオフの関係にある。

なお、以下にあげる各例に対する解である有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ については、それらのグラフを図に描くのは次に述べる方式を採用している。

(1) 変数 t を非負領域 $t \geq 0$ の全体で動かして、グラフの横軸には $s = t/(t+1)$ の値を、縦軸には有理関数 $y(t)$ の値をとって「黒線」でプロットする。これにより有理関数 $y(t)$ の非負領域全体での値の振舞いを確認できる。

なお非負領域は区間 $s \in [0, 1]$ に対応し、通過域 $t \in [0, 1]$ は必ず区間 $s \in [0, 0.5]$ に対応する。またたとえば $\mu = 1.5$ である場合は、遷移域 $t \in (1, \mu)$ は区間 $s \in (0.5, 0.6)$ に、阻止域 $t \in [\mu, \infty)$ は区間 $s \in [0.6, 1]$ にそれぞれ対応する。

(2) グラフの横軸には t の値を、縦軸には伝達関数の大きさ $|g(t)|$ の対数をとって「赤線」でプロットする。ただし $t \in [0, 3]$ の範囲だけを示す。

なお本論文で、実際にフィルタの伝達関数を構成した各例の計算においては、浮動小数点の数値とその演算には数値丸め誤差の影響をなるべく減らすために、IEEE 754 規格の四倍精度 (binary128) を用いており、それにより得られた数値を倍精度 (binary64) 程度に丸めたものを示している。またその際に現れる小規模の連立1次方程式はピボット行交換を行う Gauss の消去法を用いて解いている。

3.1 レゾルベントを1つ用いた構成

レゾルベント1つと恒等演算子の線形結合である \mathcal{Y} に対する伝達関数は式 (8) で与えられる。ここで α , τ , β はすべて実数であるとし、極 τ は負とする。

$$y(t) = \frac{\alpha}{t - \tau} + \beta. \quad (8)$$

有理関数 $y(t)$ の値の範囲に対する制約条件は式 (7) であ

表 1 レゾルベントを1つ用いたフィルタとパラメタの組の例
Table 1 Parameters of filters consist of a single resolvent.

フィルタ	μ	g_s	g_p	n
Ex1-1	2.0	1E-12	1E-4	27
Ex1-2	1.75	1E-15	1E-6	28
Ex1-3	1.5	1E-15	1E-7	33
Ex1-4	1.5	1E-16	1E-8	25
Ex1-5	1.3	1E-15	1E-8	109

り, それから得られる必要条件として $y(t)$ が1次の有理関数であり非負領域で連続かつ単調減少になることも考慮して) 式 (9) をとる.

$$\begin{cases} t=0 & \text{では} & \frac{\alpha}{0-\tau} + \beta = y_h, \\ t=1 & \text{では} & \frac{\alpha}{1-\tau} + \beta = y_\ell, \\ t=\mu & \text{では} & \frac{\alpha}{\mu-\tau} + \beta = 1, \\ t=\infty & \text{では} & \beta \geq -1. \end{cases} \quad (9)$$

必要条件である式 (9) の3つの等式を解くと, 式 (10) を上から順に計算することで β, τ, α の値が得られる.

$$\begin{cases} \beta = \frac{\gamma - y_\ell}{\gamma - 1}, \text{ ただし } \gamma \equiv \frac{y_h - y_\ell}{y_h - 1} \times \mu, \\ \tau = -\frac{y_\ell - \beta}{y_h - y_\ell}, \\ \alpha = -(y_h - \beta)\tau. \end{cases} \quad (10)$$

ただし不等式 $\beta \geq -1$ から, いま $\mu_{\min} \equiv \frac{(y_\ell + 1)(y_h - 1)}{2(y_h - y_\ell)}$ とおくと, 条件 $\mu \geq \mu_{\min}$ を満たすことが必要である.

実数シフトのレゾルベントを1つ用いて構成できたフィルタの5通りの例について, 指定したパラメタ μ, g_s, g_p, n の値を表 1 に示す. これらの例は3つの形状のパラメタ μ, g_s, g_p を先に指定して, それから条件 $\beta \geq -1$ を満たす最小の次数 n を選んでいる.

注記: 以前の論文 [4] では, 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式は $\mathcal{F} = g_s T_n(2\gamma R(\rho) - I)$ として, それに対応する伝達関数には $g(t) = g_s T_n(y)$, $y = 2x - 1, x = \frac{\mu + \sigma}{t + \sigma}$ という式を採用していた. それは本論文で採用している式 (8) において β の値を -1 に固定する場合に相当する. 実際には $y(t)$ が式 (7) の条件をすべて満たせば β は -1 以外の値をとってもよい (ただし $y(\infty) = \beta$ なので, 少なくとも $-1 \leq \beta \leq 1$ であることが必要である).

3.1.1 フィルタ Ex1-1

パラメタの組は $\mu = 2.0, g_s = 1E-12, g_p = 1E-4, n = 27$ である. それを満たす解として式 (11) が得られた.

$$\begin{cases} \beta = -0.96358454650751472, \\ \tau = -6.5170892693114966, \\ \alpha = 16.724024870445035. \end{cases} \quad (11)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそ

れぞれ図 3 と図 4 に示す.

このフィルタ Ex1-1 は μ の値が 2.0 で, 遷移域の相対幅は広くて遮断特性は緩慢である. g_s の値は 1E-12 であるから阻止域での阻止能力は他のものに比べて悪いが, g_p の値は 1E-4 であるから通過域での伝達率の均一性は他のものよりも良い.

3.1.2 フィルタ Ex1-2

パラメタの組は $\mu = 1.75, g_s = 1E-15, g_p = 1E-6, n = 28$ である. それを満たす解として式 (12) が得られた.

$$\begin{cases} \beta = -0.99245324139048424, \\ \tau = -3.8666639819616197, \\ \alpha = 11.190940356660613. \end{cases} \quad (12)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 6 と図 7 に示す.

このフィルタ Ex1-2 は μ の値が 1.75 で, 遷移域の相対幅は少し広く遮断特性は緩慢である. g_s の値は 1E-15 であり阻止域での阻止能力は良いが, g_p の値は 1E-6 でありフィルタ Ex1-1 に比べて通過域での伝達率の均一性は悪い.

3.1.3 フィルタ Ex1-3

パラメタの組は $\mu = 1.5, g_s = 1E-15, g_p = 1E-7, n = 33$ である. それを満たす解として式 (13) が得られた.

$$\begin{cases} \beta = -0.98699306841269848, \\ \tau = -4.7599830140151067, \\ \alpha = 12.438542857229249. \end{cases} \quad (13)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 9 と図 10 に示す.

このフィルタ Ex1-3 は μ の値が 1.5 で, 遷移域の相対幅はほどほどであり遮断特性もほどほどである. g_s の値は 1E-15 であり阻止域での阻止能力は良いが, g_p の値は 1E-7 であり, フィルタ Ex1-1 や Ex1-2 に比べて通過域での伝達率の均一性は悪い.

3.1.4 フィルタ Ex1-4

パラメタの組は $\mu = 1.5, g_s = 1E-16, g_p = 1E-8, n = 25$ である. それを満たす解として式 (14) が得られた.

$$\begin{cases} \beta = -0.91180782525544148, \\ \tau = -2.1158443350135284, \\ \alpha = 6.9127994945844218. \end{cases} \quad (14)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 12 と図 13 に示す.

このフィルタ Ex1-4 も μ の値は 1.5 で, 遷移域の相対幅はほどほどであり遮断特性もほどほどである. g_s の値は 1E-16 であり阻止域での阻止能力は良いが, g_p の値は 1E-8 なのでフィルタ Ex1-1, Ex1-2, Ex1-3 に比べて通過域での伝達率の均一性は悪い. ただし Chebyshev の多項式の次数 n は 25 で, ほかに比べて低めになっている.

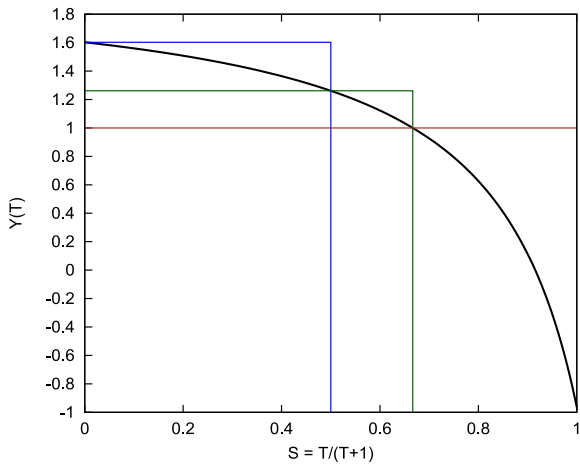


図 3 Ex1-1: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 2.0, g_s = 1E-12, g_p = 1E-4, n = 27)$

Fig. 3 Ex1-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

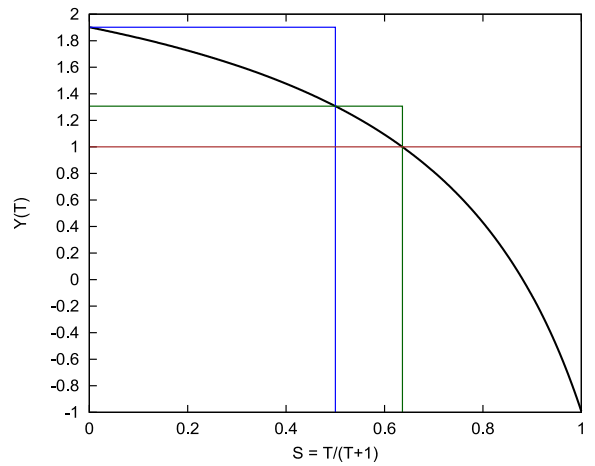


図 6 Ex1-2: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.75, g_s = 1E-15, g_p = 1E-6, n = 28)$

Fig. 6 Ex1-2: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

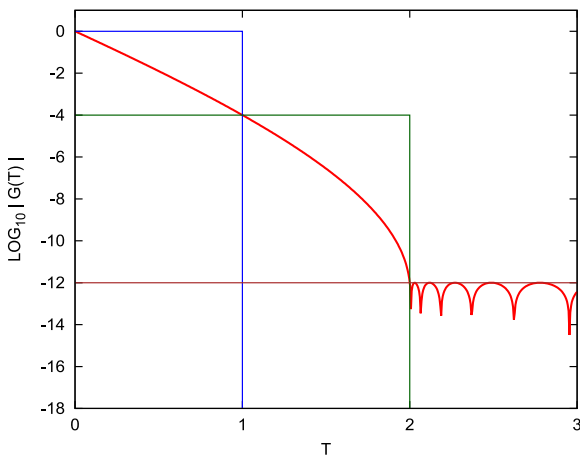


図 4 Ex1-1: 伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 2.0, g_s = 1E-12, g_p = 1E-4, n = 27)$

Fig. 4 Ex1-1: Log plot of the magnitude of $g(t)$.

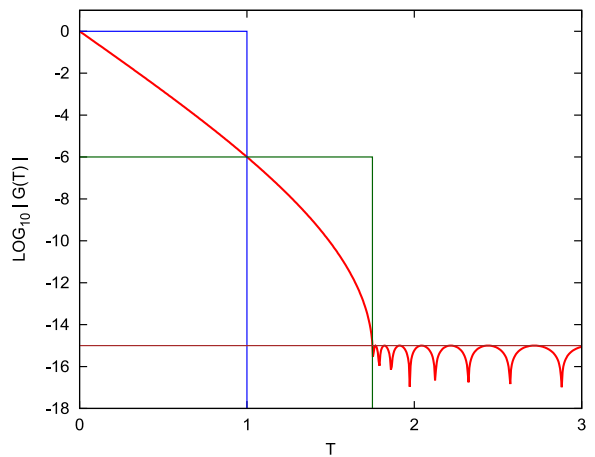


図 7 Ex1-2: 伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.75, g_s = 1E-15, g_p = 1E-6, n = 28)$

Fig. 7 Ex1-2: Log plot of the magnitude of $g(t)$.

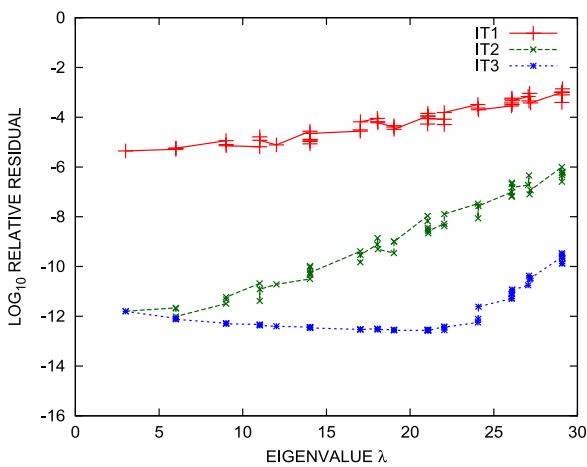


図 5 Ex1-1: 近似固有対の相対残差の大きさの対数
 $(\mu = 2.0, g_s = 1E-12, g_p = 1E-4, n = 27)$

Fig. 5 Ex1-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

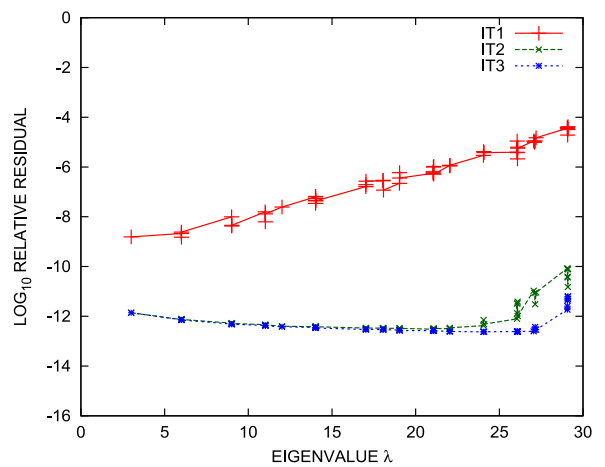


図 8 Ex1-2: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.75, g_s = 1E-15, g_p = 1E-6, n = 28)$

Fig. 8 Ex1-2: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

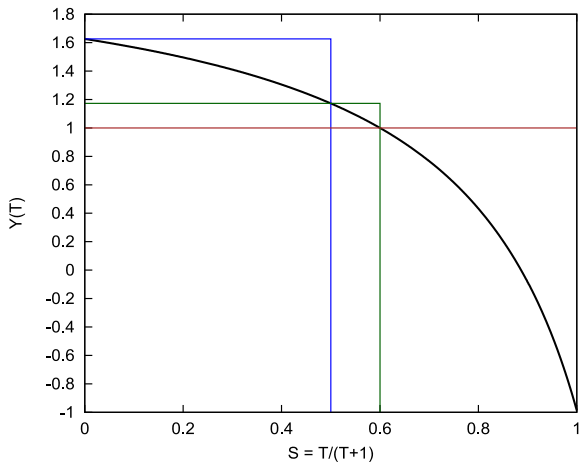


図 9 Ex1-3: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-7, n = 33)$

Fig. 9 Ex1-3: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

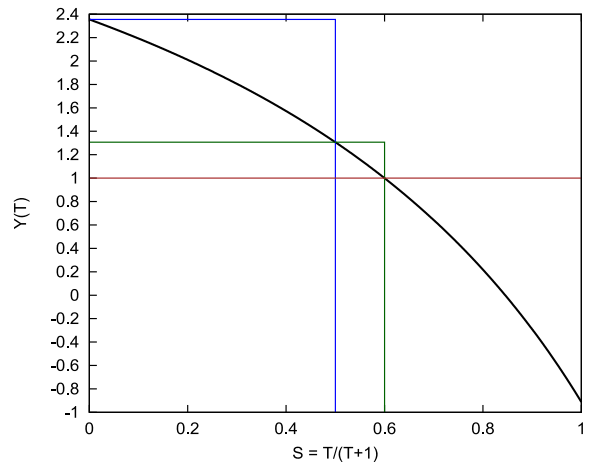


図 12 Ex1-4: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-16, g_p = 1E-8, n = 25)$

Fig. 12 Ex1-4: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

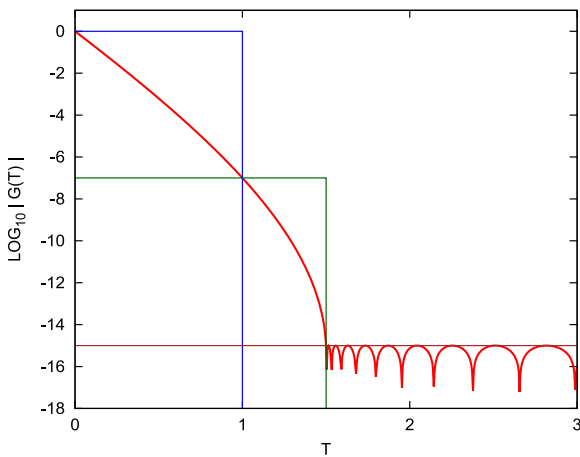


図 10 Ex1-3: 伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-7, n = 33)$

Fig. 10 Ex1-3: Log plot of the magnitude of $g(t)$.

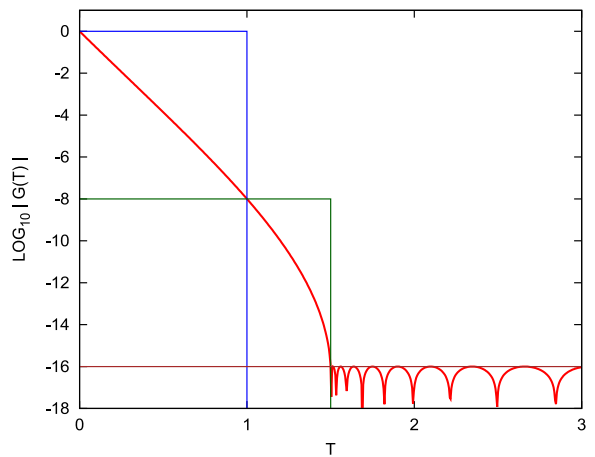


図 13 Ex1-4: 伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-16, g_p = 1E-8, n = 25)$

Fig. 13 Ex1-4: Log plot of the magnitude of $g(t)$.

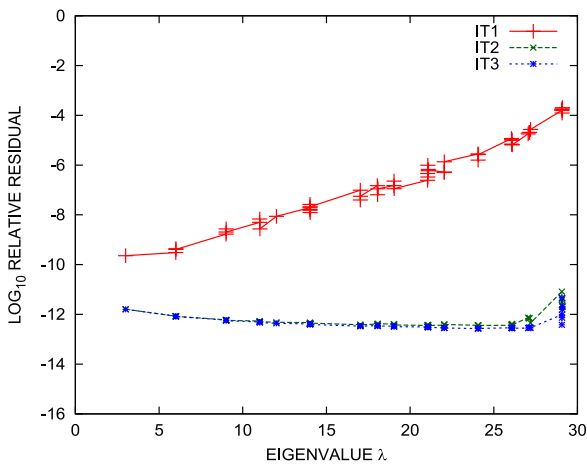


図 11 Ex1-3: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-7, n = 33)$

Fig. 11 Ex1-3: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

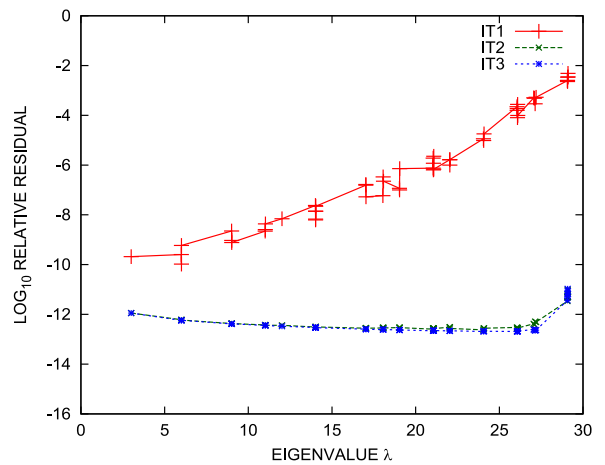


図 14 Ex1-4: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-16, g_p = 1E-8, n = 25)$

Fig. 14 Ex1-4: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

3.1.5 フィルタ Ex1-5

パラメタの組は $\mu = 1.3, g_s = 1E-15, g_p = 1E-8, n = 109$ である. それを満たす解として式 (15) が得られた.

$$\begin{cases} \beta = -0.98752042609237490, \\ \tau = -49.032877479946448, \\ \alpha = 100.03762209539847. \end{cases} \quad (15)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 15 と図 16 に示す.

このフィルタ Ex1-5 は μ の値が 1.3 で小さいので遷移域の相対幅は少し狭く, 遮断特性も他よりは急になっている. g_s の値は $1E-15$ であり阻止域での阻止能力は良いが, g_p の値は $1E-8$ であり通過域での伝達率の均一性はフィルタ Ex1-4 と同じで Ex1-1, Ex1-2, Ex1-3 よりも悪い. この場合は Chebyshev の多項式の次数 n が 109 であり, ほかに場合に比べてかなり大きくなっている. これは μ の値を小さくしたためである.

3.2 レゾルベントを 2 つ用いた構成の例

レゾルベント 2 つと恒等演算子の線形結合 \mathcal{Y} に対する伝達関数を式 (16) とする. ここで $\alpha_1, \alpha_2, \tau_1, \tau_2, \beta$ はすべて実数であり, 極 τ_1 と τ_2 は負である.

$$y(t) = \frac{\alpha_1}{t - \tau_1} + \frac{\alpha_2}{t - \tau_2} + \beta. \quad (16)$$

有理関数 $y(t)$ への制約条件は式 (7) である. フィルタの設計パラメタとして μ, g_s, g_p, n の値を指定する. まず閾値 y_ℓ と y_h を式 (5) から計算する. そうして通過域 $t \in [0, 1]$ で $y(t)$ が最大値 y_h をとる点の座標を $t = z$ とする. $z \neq 0$ ならば $y(t)$ は以下の 4 つの条件 (17) を満たす必要がある.

$$\begin{cases} y(z) = y_h, \quad y'(z) = 0, \\ y(1) = y_\ell, \\ y(\mu) = 1. \end{cases} \quad (17)$$

また $|y(\infty)| \leq 1$ つまり $|\beta| \leq 1$ も必要である. そうして, 与えられた β の値に対して, 4 つの等式 (18) を解くことで 4 つの変数 $\alpha_1, \alpha_2, \tau_1, \tau_2$ の実数値を求めることになる.

$$\begin{cases} \frac{\alpha_1}{z - \tau_1} + \frac{\alpha_2}{z - \tau_2} + \beta = y_h, \\ -\frac{\alpha_1}{(z - \tau_1)^2} - \frac{\alpha_2}{(z - \tau_2)^2} = 0, \\ \frac{\alpha_1}{1 - \tau_1} + \frac{\alpha_2}{1 - \tau_2} + \beta = y_\ell, \\ \frac{\alpha_1}{\mu - \tau_1} + \frac{\alpha_2}{\mu - \tau_2} + \beta = 1. \end{cases} \quad (18)$$

この連立代数方程式 (18) は初等代数的な操作で解くことができる [8]. そうして解いて得られた $y(t)$ が実数の範囲での部分分数分解を持ち, 通過域, 遷移域, 阻止域のそれぞれにおいて関数値が条件 (7) を実際に満たしているかを調べればよいが, 記述が長くなるのでここでは省略する.

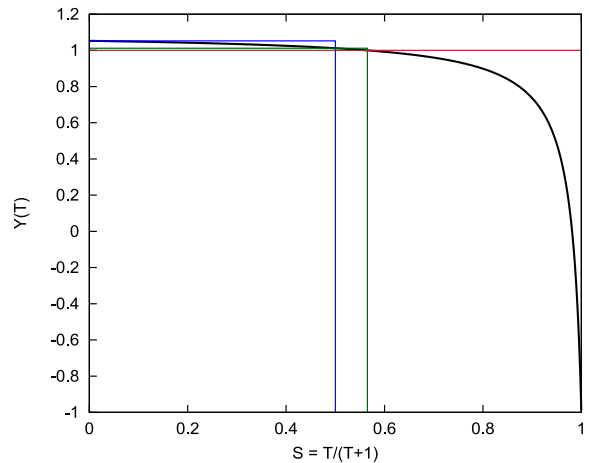


図 15 Ex1-5: 有理関数 $y(t)$ の値 (横軸 $s = t/(t + 1)$) ($\mu = 1.3, g_s = 1E-15, g_p = 1E-8, n = 109$)

Fig. 15 Ex1-5: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t + 1)$).

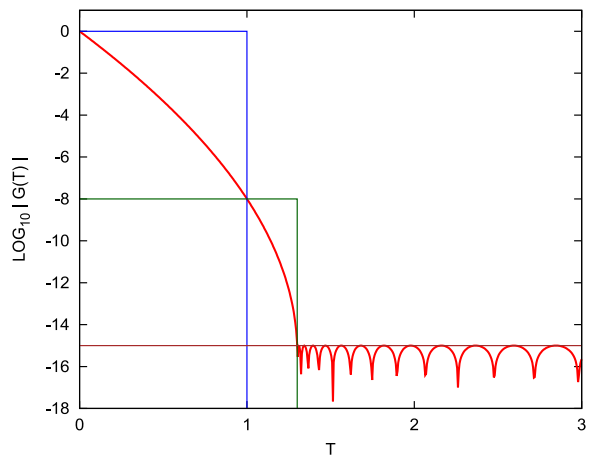


図 16 Ex1-5: 伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.3, g_s = 1E-15, g_p = 1E-8, n = 109$)

Fig. 16 Ex1-5: Log plot of the magnitude of $g(t)$.

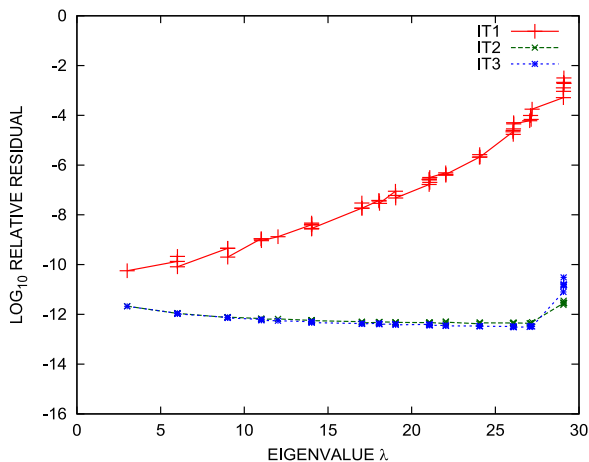


図 17 Ex1-5: 近似固有対の相対残差の大きさの対数 ($\mu = 1.3, g_s = 1E-15, g_p = 1E-8, n = 109$)

Fig. 17 Ex1-5: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

指定するパラメタ μ, g_s, g_p, n の値の組によっては実数解のない場合や、あるいは実数であっても非負の極が現れたり、あるいは実数解で2つの極が負の実数になっていたとしても、 $y(t)$ の値の範囲が通過域や阻止域における制限を満たしていない可能性がある。

また代数的な解法の場合には、解を実数に制限することは容易ではないので、以下では反復法を用いて実数の範囲で連立代数方程式を解くことにする。

通過域における最大点の座標 z は、文献 [2] の Remez 法と同様の反復の手法により解くことにする。まずフィルタのパラメタとして指定された g_s, g_ℓ, n の値から式 (5) により、 $y(t)$ に対する閾値である y_ℓ, y_h の値をそれぞれ求める。そうしてまず極の位置である τ_1, τ_2 の値を仮定して、さらに z の初期値も与える。

z が極大点であるための必要条件 $y'(z) = 0$ を式 (17) から除いた残りの3つの必要条件から、連立1次方程式 (19) が得られる。するといま仮定している τ_1, τ_2, z の値から連立1次方程式の係数や右辺の数値が定まるので、それを解いて $\alpha_1, \alpha_2, \beta$ の数値が求まり、式 (16) の $y(t)$ に含まれる定数の数値がすべて決まる。

$$\begin{bmatrix} 1 & 1 & 1 \\ z - \tau_1 & z - \tau_2 & 1 \\ \frac{1}{1 - \tau_1} & \frac{1}{1 - \tau_2} & 1 \\ \frac{1}{\mu - \tau_1} & \frac{1}{\mu - \tau_2} & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta \end{bmatrix} = \begin{bmatrix} y_h \\ y_\ell \\ 1 \end{bmatrix}. \quad (19)$$

このようにして得られた有理関数 $y(t)$ の通過域における最大点の座標がちょうど z に一致していて、しかも通過域で最小値が y_ℓ より小さくならず、さらに阻止域では $|y(t)|$ が1を超えなければ条件を満たす $y(t)$ の構成に成功したことになる。しかし $y(t)$ の最大点の座標が z に一致しなければ、Remez 法と同様の反復により、 z に近い最大点の座標を求めてそれを次回の z の値に設定して再び連立1次方程式 (19) を解く処理に戻って繰り返す (図 18)。

この Remez 法と同様の反復で z の値が収束したときに、通過域での $y(t)$ の最小値が y_ℓ よりも小さかったり、あるいは阻止域 $[\mu, \infty)$ の中で $|y(t)|$ が1を超えることがあれば、 $y(t)$ は必要な制約を満たしておらず、構成は失敗である。失敗した場合には、最初に仮定した2つの極の位置 τ_1, τ_2 を変更して再び探索を行う。今回はこの2つの極の位置の探索は人の手で試行錯誤により行った。

実数シフトのレゾルベントを2つ用いて実現できたフィルタの4通りの構成例について、指定したパラメタ μ, g_s, g_p, n の値を表 2 に示す (これらの例はまだ改良できる可能性がある)。これらの例ではすべて $\mu = 1.5$ としている。

3.2.1 フィルタ Ex2-1

パラメタの組は $\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 23$ である。それを満たす解として式 (20) が得られた。

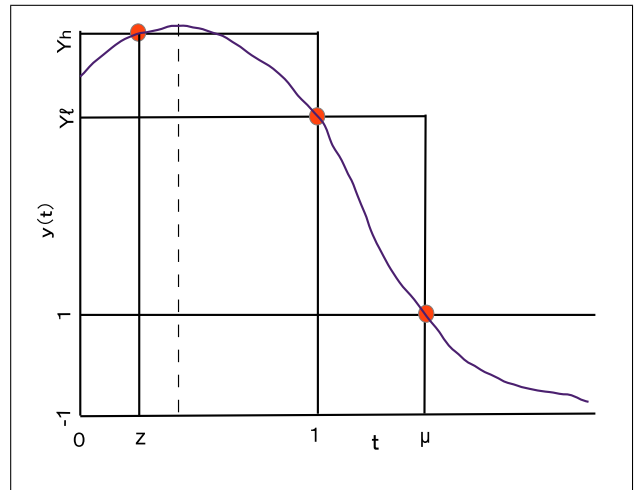


図 18 Remez 法 の 概 念 (仮 定 し た 極 大 点 の 座 標 z を 修 正 す る)
 Fig. 18 Concept of Remez method (The abscissa z of the assumed maximum point will be corrected).

表 2 レゾルベントを2つ用いたフィルタとパラメタの組の例

Table 2 Parameters of filters consist of 2 resolvents.

フィルタ	μ	g_s	g_p	n
Ex2-1	1.5	1E-10	1E-3	23
Ex2-2	1.5	1E-12	1E-3	38
Ex2-3	1.5	1E-15	1E-5	38
Ex2-4	1.5	1.1E-15	1E-4	40

表 3 Ex2-1: Remez 法と同様の反復による z の値の収束状況

Table 3 Ex2-1: Convergence of z by Remez-like iteration.

反復回数	z
0	0.25
1	0.31679715742666303
2	0.32638897662218369
3	0.32657165874206839
4	0.32657172404697121
5	0.32657172404697121

$$\begin{cases} \tau_1 = -0.76, \tau_2 = -1.2, \\ z = 0.32657172404697121, \\ \beta = -0.99230416024046638, \\ \alpha_1 = -6.9037344873220844, \\ \alpha_2 = 13.627045620157944. \end{cases} \quad (20)$$

Remez 法と同様の反復の収束の状況を、 z の初期値を 0.25 とした場合について表 3 に示す (反復において1回前の値と一致する部分の有効数字の桁を下線で示す)。

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 19 と図 20 に示す。

フィルタ Ex2-1 は Chebyshev 多項式の次数 n が 23 と小さい。 g_s の値は 1E-10 とかなり大きく阻止域における遮断特性は悪い。 g_p の値は 1E-3 で2次の有理関数を用いた4つの例の中では Ex2-2 と並んで最も大きく通過域での伝達率の均一性は4つの中では Ex2-2 と並んで最も良い。

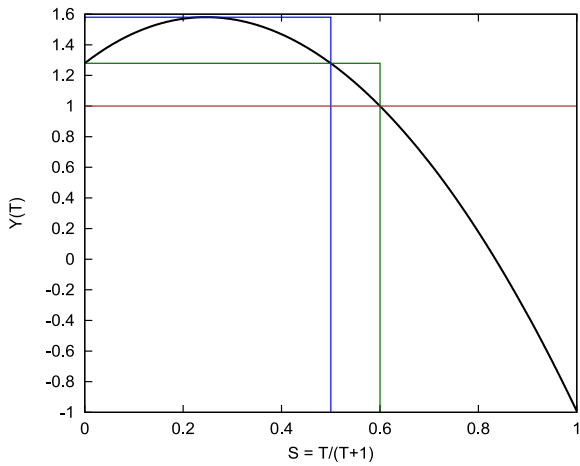


図 19 Ex2-1: 有理関数 $y(t)$ の値 (横軸 $s = t/(t + 1)$)
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 23)$

Fig. 19 Ex2-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t + 1)$).

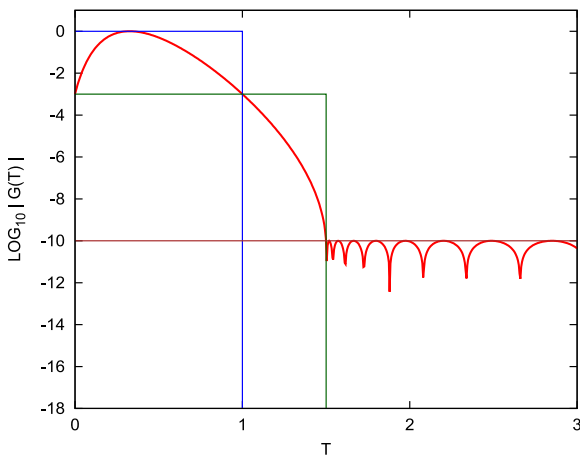


図 20 Ex2-1: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 23)$

Fig. 20 Ex2-1: Log plot of the magnitude of $g(t)$.

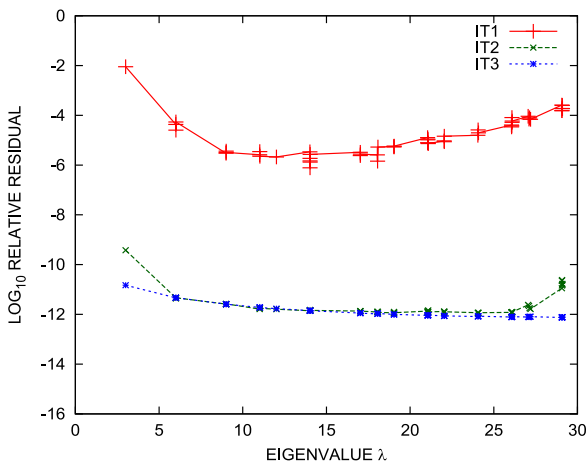


図 21 Ex2-1: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 23)$

Fig. 21 Ex2-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

3.2.2 フィルタ Ex2-2

パラメタの組は $\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 38$ である. それを満たす解として式 (21) が得られた.

$$\begin{cases} \tau_1 = -1.25, \tau_2 = -2.03, \\ z = 0.37507\ 63215\ 83946\ 99, \\ \beta = -0.94058\ 93038\ 64582\ 17, \\ \alpha_1 = -7.55520\ 41413\ 37089\ 6, \\ \alpha_2 = 16.54841\ 50131\ 58312. \end{cases} \quad (21)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 22 と図 23 に示す.

フィルタ Ex2-1 に比べてフィルタ Ex2-2 では g_s の値を $1E-12$ に下げることにより阻止域における遮断性を高めた. g_p の値は Ex2-1 と同じである. g_s の値を下げるためには Chebyshev 多項式の次数 n を 23 から 38 に増やす必要があった.

3.2.3 フィルタ Ex2-3

パラメタの組は $\mu = 1.5, g_s = 1E-15, g_p = 1E-5, n = 38$ である. それを満たす解として式 (22) が得られた.

$$\begin{cases} \tau_1 = -0.8, \tau_2 = -2.5, \\ z = 0.17911\ 72407\ 28373\ 87, \\ \beta = -0.97197\ 32868\ 06298\ 33, \\ \alpha_1 = -1.37229\ 04037\ 26818\ 0, \\ \alpha_2 = 10.27448\ 51537\ 06616. \end{cases} \quad (22)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 25 と図 26 に示す.

フィルタ Ex2-2 に比べてフィルタ Ex2-3 では g_s の値をさらに $1E-15$ に下げて阻止域における遮断性を高めた. g_p の値は Ex2-2 よりも 2 桁小さい $1E-5$ にしたので, 通過域における伝達率の均一性はかなり悪くなっている. Ex2-2 に比べて Ex2-3 では g_s の値を 3 桁小さくしたが, 同時に g_p の値を 2 桁小さくしたことにより, Chebyshev 多項式の次数 n は Ex2-2 と同じく 38 にすることができた.

3.2.4 フィルタ Ex2-4

パラメタの組は $\mu = 1.5, g_s = 1.1E-15, g_p = 1E-4, n = 40$ である. それを満たす解として式 (23) が得られた (パラメタ g_s の値が $1E-15$ ではなくて $1.1E-15$ になっているのは, そうしなければ条件を満たす解が得られなかったからである).

$$\begin{cases} \tau_1 = -1.25, \tau_2 = -1.28, \\ z = 0.35778\ 64394\ 87426\ 83, \\ \beta = -0.99864\ 90114\ 17620\ 76, \\ \alpha_1 = -207.65212\ 64969\ 0007, \\ \alpha_2 = 215.47366\ 66740\ 6178. \end{cases} \quad (23)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 28 と図 29 に示す.

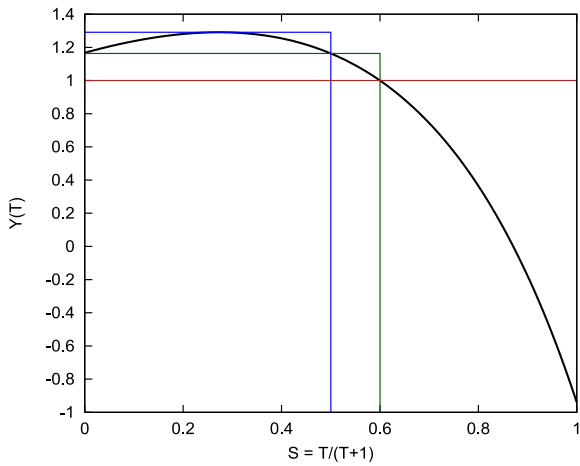


図 22 Ex2-2: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 38)$

Fig. 22 Ex2-2: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

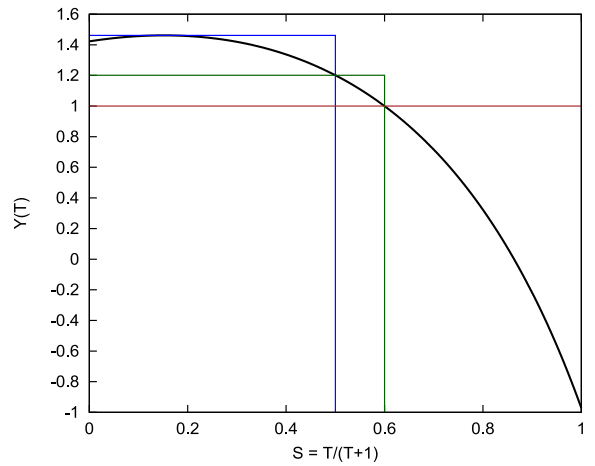


図 25 Ex2-3: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-5, n = 38)$

Fig. 25 Ex2-3: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

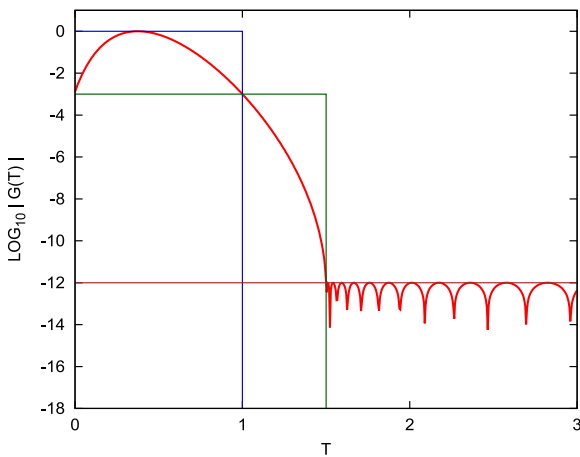


図 23 Ex2-2: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 38)$

Fig. 23 Ex2-2: Log plot of the magnitude of $g(t)$.

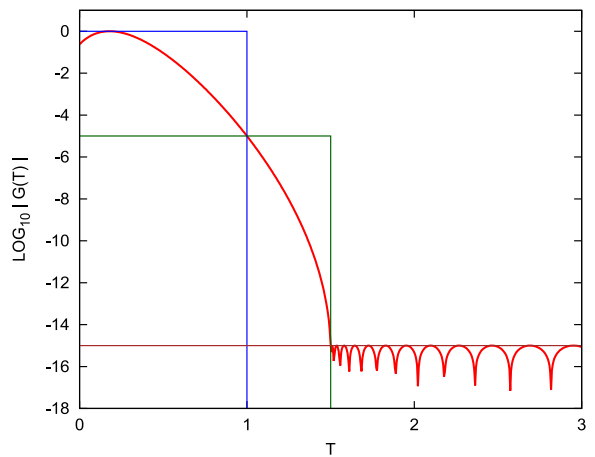


図 26 Ex2-3: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-5, n = 38)$

Fig. 26 Ex2-3: Log plot of the magnitude of $g(t)$.

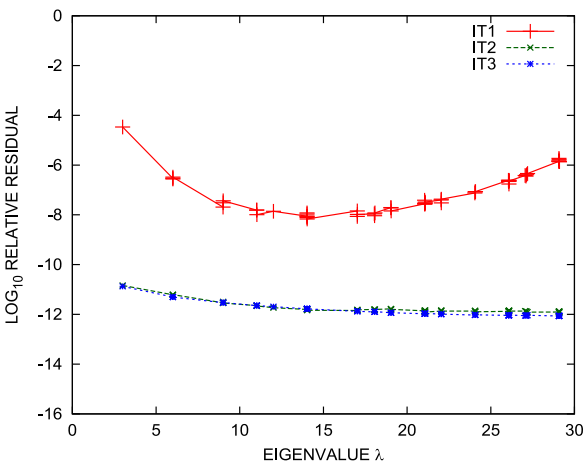


図 24 Ex2-2: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 38)$

Fig. 24 Ex2-2: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

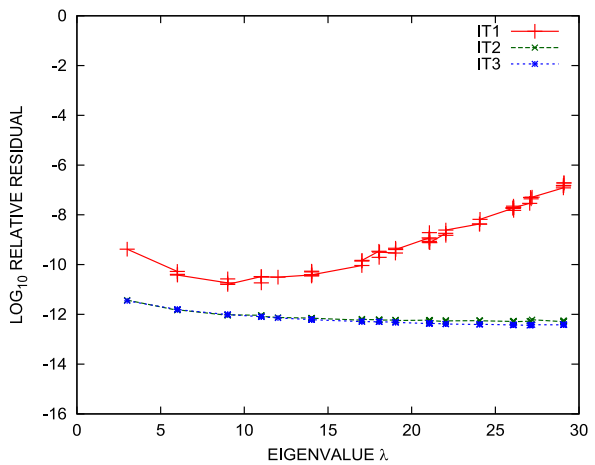


図 27 Ex2-3: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-15, g_p = 1E-5, n = 38)$

Fig. 27 Ex2-3: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

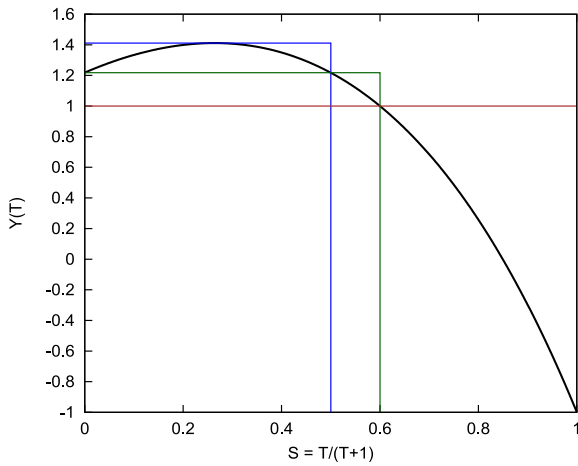


図 28 Ex2-4: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1.1E-15, g_p = 1E-4, n = 40)$

Fig. 28 Ex2-4: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

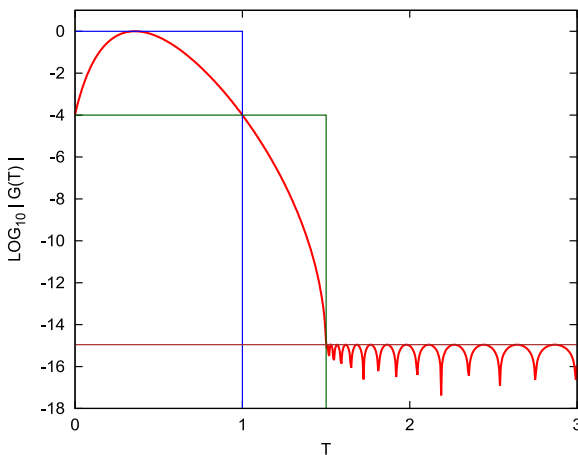


図 29 Ex2-4: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1.1E-15, g_p = 1E-4, n = 40)$

Fig. 29 Ex2-4: Log plot of the magnitude of $g(t)$.

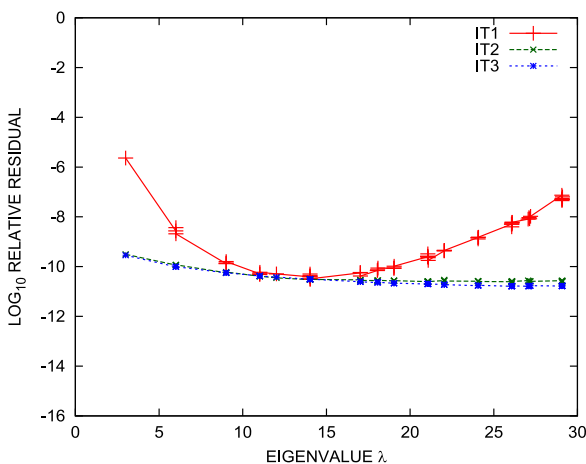


図 30 Ex2-4: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1.1E-15, g_p = 1E-4, n = 40)$

Fig. 30 Ex2-4: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

フィルタ Ex2-3 に比べてフィルタ Ex2-4 では g_s の値はほとんど同じであるが, g_p の値を 1桁大きく $1E-4$ にして通過域における伝達率の均一性を良くしようとした例である. しかし 2つの極 τ_1 と τ_2 の位置がかなり接近していること, それにともなって逆符号である各極の係数 α_1 と α_2 の値の絶対値が大きくなったことから, 2次の有理関数を用いた他のフィルタに比べてベクトルの組に 2つのレゾルベントを作用させた結果の線形結合を作る際に必要な固有ベクトルに対する数値相殺の程度が強まり, それにより丸め誤差が拡大されて計算結果の相対精度が低下する傾向が高くなると考えられる.

3.3 レゾルベントを 3つ用いた構成の例 (方式 I)

レゾルベント 3つと恒等演算子の線形結合に対する伝達関数を式 (24) とする. ここで $\alpha_j, \tau_j, j = 1, 2, 3$ と β はすべて実数であり, 3つの極 τ_j はすべて負であるとする.

$$y(t) = \sum_{j=1}^3 \frac{\alpha_j}{t - \tau_j} + \beta. \quad (24)$$

有理関数 $y(t)$ についての制約条件は式 (7) になる. フィルタの設計パラメータとして μ, g_s, g_p, n の値を指定する (ただし $1 < \mu, 0 < g_s \ll g_p < 1$ である). $y(t)$ の閾値 y_ℓ と y_h は式 (5) から計算で求める.

この方式 I では通過域 $[0, 1]$ における $y(t)$ の最大点の座標を $t = z$ とする. すると $y(t)$ は式 (25) の 4つの条件を満たす必要がある.

$$\begin{cases} y(z) = y_h, \\ t \in [0, 1] \text{ での } y(t) \text{ の最大点が } z, \\ y(1) = y_\ell, \\ y(\mu) = 1. \end{cases} \quad (25)$$

ここでは β の値をうまく仮定して指定することにする. $\beta = y(\infty)$ であるから $|\beta| \leq 1$ である必要がある.

仮定した 3つの極の位置 τ_1, τ_2, τ_3 と仮定した β の値から連立 1次方程式 (26) の係数と右辺の数値が決まり, それを解くことで $\alpha_1, \alpha_2, \alpha_3$ の数値がそれぞれ求まり, それにより式 (24) の $y(t)$ の定数の数値がすべて決まる.

$$\begin{bmatrix} 1/(z - \tau_1) & 1/(z - \tau_2) & 1/(z - \tau_3) \\ 1/(1 - \tau_1) & 1/(1 - \tau_2) & 1/(1 - \tau_3) \\ 1/(\mu - \tau_1) & 1/(\mu - \tau_2) & 1/(\mu - \tau_3) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} y_h - \beta \\ y_\ell - \beta \\ 1 - \beta \end{bmatrix} \quad (26)$$

こうして得られた $y(t)$ の通過域 $[0, 1]$ における最大点の座標が z に一致して, 通過域での最小値が y_ℓ でありさらに阻止域 $[\mu, \infty)$ で $|y(t)|$ の値が 1 を超えることがなければ条件を満たす $y(t)$ の構成に成功したことになる. しかし通過域における最大点の座標が z でない場合には, 真の最大点の座標を求めてそれを次回の z に用いて繰り返しを

表 4 レゾルベントを 3 つ用いたフィルタとパラメタの例 (方式 I)

Table 4 Parameters of filters which consist of 3 resolvents (method I).

フィルタ	μ	g_s	g_p	n
Ex3-I-1	1.5	1E-10	1E-3	20
Ex3-I-2	1.5	1E-10	1E-2	30
Ex3-I-3	1.5	1E-12	1E-3	30
Ex3-I-4	1.5	1E-12	5E-3	40
Ex3-I-5	1.5	1.2E-12	1E-3	20
Ex3-I-6	1.5	2E-14	1E-3	30

表 5 Ex3-I-1: Remez 法と同様の反復による z の値の収束状況

Table 5 Ex3-I-1: Convergence of z by Remez-like iteration.

反復回数	z
0	0.25
1	0.17915257662534714
2	0.18592556801810861
3	0.18599261632282287
4	0.18599262285418808
5	0.18599262285418808

行って収束させる (Remez 法と同様の反復). 収束が得られたときに通過域における $y(t)$ の最小値が y_ℓ に比べて小さかったり, あるいは阻止域で $|y(t)|$ の値が 1 を超える点があれば $y(t)$ の構成は失敗であり, その場合には極の位置 τ_1, τ_2, τ_3 あるいは β の値を変更して探索をさらに行う. 今回はこの探索を人手による試行錯誤で行っている.

指定したフィルタのパラメタ μ, g_s, g_p, n の値によっては, 条件を満たす解が存在しない可能性がある. ただし適切な解が見つからない場合, それが探索が不十分であるからなのか, それとも条件を満たす解が本当に存在しないからなのかを現状では区別できていない.

実数シフトのレゾルベントを 3 つ用いて構成できたフィルタの 6 通りの例について, 指定したパラメタ μ, g_s, g_p, n の値を表 4 に示す (これらはまだ改良できる可能性がある). この 6 つの例すべてで μ の値は共通の 1.5 にした.

3.3.1 フィルタ Ex3-I-1

パラメタの組は $\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 20$ である. それを満たす解として式 (27) が得られた.

$$\begin{cases} \beta = 1.0, \tau_1 = -1.5, \tau_2 = -5.0, \tau_3 = -7.5, \\ z = 0.18599262285418808, \\ \alpha_1 = -10.543472067071244, \\ \alpha_2 = 230.91814774839765, \\ \alpha_3 = -288.10240375810610. \end{cases} \quad (27)$$

Remez 法と同様の反復による収束の状況を z の初期値を 0.25 とした場合について表 5 に示す.

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 31 と図 32 に示す. これは通過域の下端付近での伝達率が 1 に近い. もしも原点に最も近い $g(t) = g_p$ と

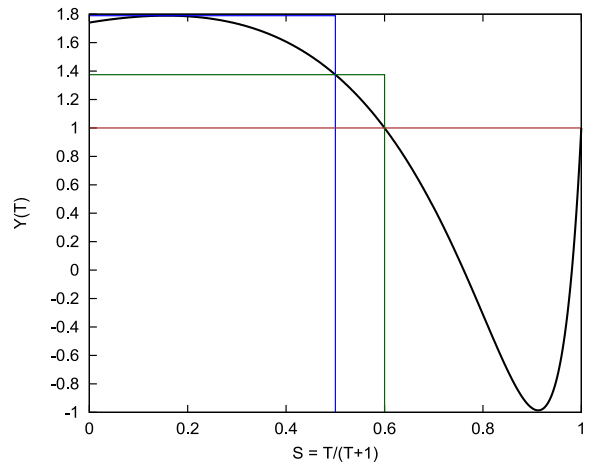


図 31 Ex3-I-1: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$) ($\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 20$)

Fig. 31 Ex3-I-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

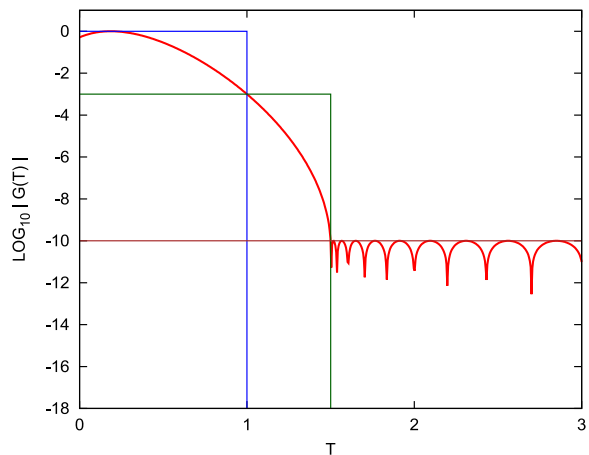


図 32 Ex3-I-1: フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 20$)

Fig. 32 Ex3-I-1: Log plot of the magnitude of $g(t)$.

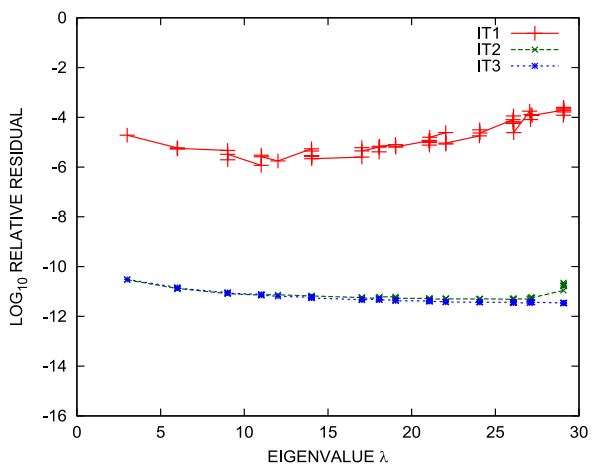


図 33 Ex3-I-1: 近似固有対の相対残差の大きさの対数 ($\mu = 1.5, g_s = 1E-10, g_p = 1E-3, n = 20$)

Fig. 33 Ex3-I-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

なる負の座標を $t = t_p$ とするとき、区間 $t \in [t_p, 1]$ と区間 $\tilde{t} \in [0, 1]$ を同じ向きに対応させる 1 次変換で座標変換を定義して $\tilde{g}(\tilde{t}) \equiv g(t)$ とすれば、新しい伝達関数 $\tilde{g}(\tilde{t})$ の通過域の下端 $\tilde{t} = 0$ での伝達率は減少して g_p になるが、元の伝達関数 $g(t)$ の遷移域の相対幅 $\mu - 1$ に比べて $\tilde{g}(\tilde{t})$ の遷移域の相対幅 $\tilde{\mu} - 1$ は $1/(1 - t_p)$ 倍になり、幅を狭くできる。

3.3.2 フィルタ Ex3-I-2

パラメタの組は $\mu = 1.5$, $g_s = 3E-10$, $g_p = 1E-2$, $n = 30$ である。それを満たす解として式 (28) が得られた。

$$\begin{cases} \beta = -0.65, \tau_1 = -2.5, \tau_2 = -4.0, \tau_3 = -8.0, \\ z = 0.41863\ 03472\ 33265\ 64, \\ \alpha_1 = -46.94773\ 12232\ 16199, \\ \alpha_2 = 138.44834\ 07901\ 2331, \\ \alpha_3 = -111.96218\ 15278\ 0179. \end{cases} \quad (28)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 34 と図 35 に示す。

フィルタ Ex3-I-2 は Ex3-I-1 に比べて g_p の値を 1 桁大きくすることで通過域における伝達率の均一性を良くすることを狙ったが、それと引き換えに Chebyshev 多項式の次数 n が 20 から 30 へと 1.5 倍に増大したことが欠点である。

3.3.3 フィルタ Ex3-I-3

パラメタの組は $\mu = 1.5$, $g_s = 1E-12$, $g_p = 1E-3$, $n = 30$ である。それを満たす解として式 (29) が得られた。

$$\begin{cases} \beta = 0.2, \tau_1 = -2.5, \tau_2 = -4.0, \tau_3 = -8.0, \\ z = 0.30267\ 47675\ 72432\ 76, \\ \alpha_1 = -61.72673\ 49701\ 25999, \\ \alpha_2 = 194.11248\ 93001\ 2096, \\ \alpha_3 = -181.08421\ 32370\ 6878. \end{cases} \quad (29)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 37 と図 38 に示す。

フィルタ Ex3-I-2 の場合と比べて、フィルタ Ex3-I-3 では g_s の値を $3E-10$ から $1E-12$ に下げると引き換えに g_p の値を $1E-2$ から $1E-3$ に下げた。Chebyshev 多項式の次数 n はどちらも 30 で同じである。

3.3.4 フィルタ Ex3-I-4

パラメタの組は $\mu = 1.5$, $g_s = 1E-12$, $g_p = 5E-3$, $n = 40$ である。それを満たす解として式 (30) が得られた。

$$\begin{cases} \beta = 1.0, \tau_1 = -3.65, \tau_2 = -5.5, \tau_3 = -10.5, \\ z = 0.43751\ 32482\ 75965\ 45, \\ \alpha_1 = -130.27325\ 48867\ 5910, \\ \alpha_2 = 374.74537\ 92283\ 5478, \\ \alpha_3 = -338.87131\ 83307\ 7839. \end{cases} \quad (30)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 40 と図 41 に示す。

フィルタ Ex3-I-3 の場合と比べて、フィルタ Ex3-I-4 で

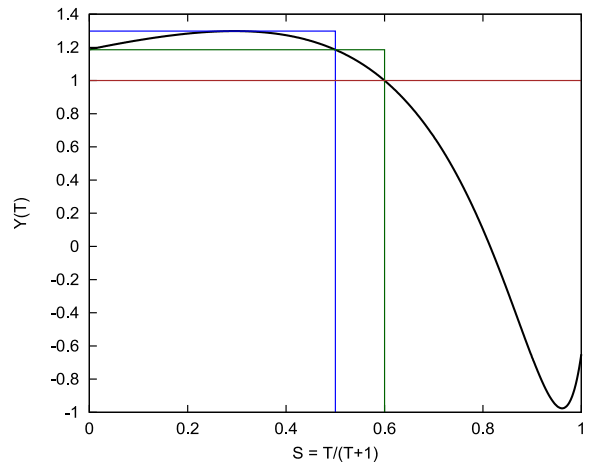


図 34 Ex3-I-2: 有理関数 $y(t)$ の値 (横軸 $s = t/(t + 1)$) ($\mu = 1.5$, $g_s = 3E-10$, $g_p = 1E-2$, $n = 30$)

Fig. 34 Ex3-I-2: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t + 1)$).

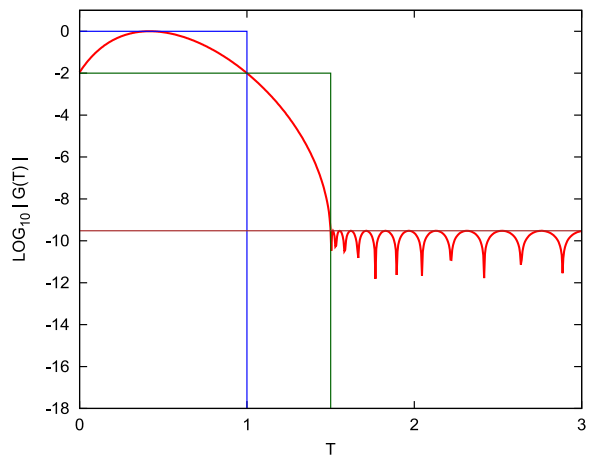


図 35 Ex3-I-2: フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5$, $g_s = 3E-10$, $g_p = 1E-2$, $n = 30$)

Fig. 35 Ex3-I-2: Log plot of the magnitude of $g(t)$.

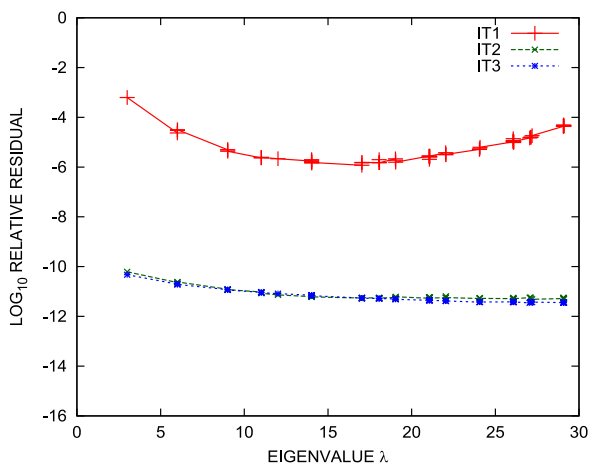


図 36 Ex3-I-2: 近似固有対の相対残差の大きさの対数 ($\mu = 1.5$, $g_s = 3E-10$, $g_p = 1E-2$, $n = 30$)

Fig. 36 Ex3-I-2: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

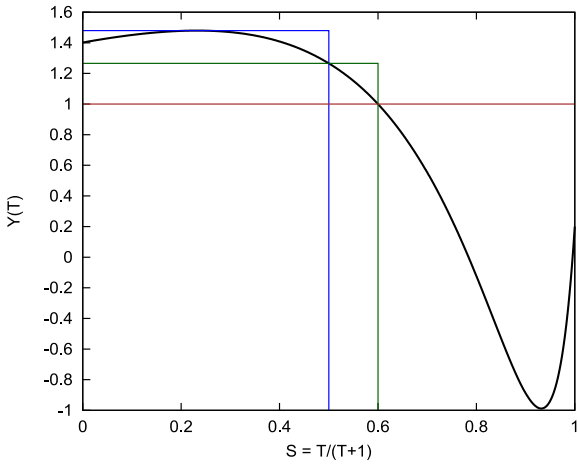


図 37 Ex3-I-3: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 30)$

Fig. 37 Ex3-I-3: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

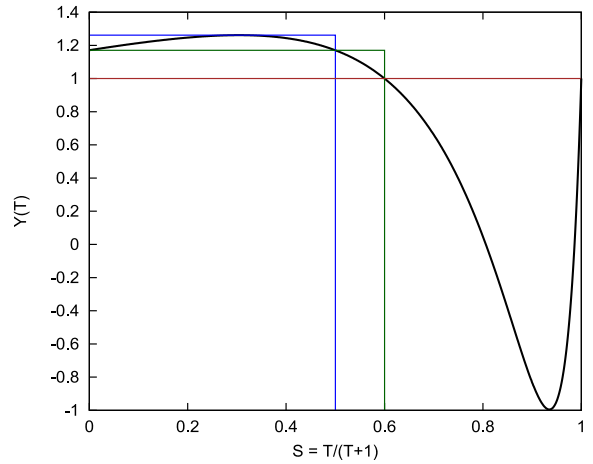


図 40 Ex3-I-4: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-12, g_p = 5E-3, n = 40)$

Fig. 40 Ex3-I-4: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

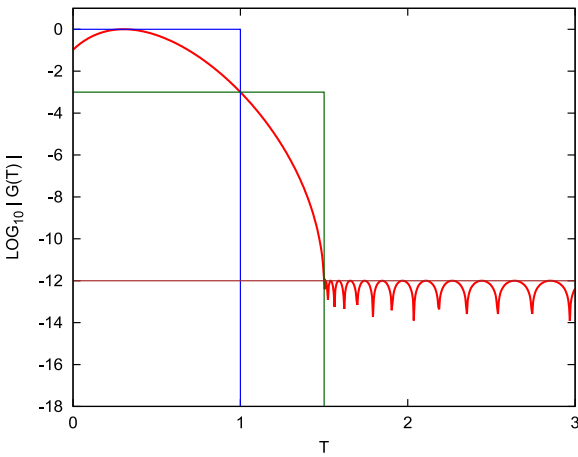


図 38 Ex3-I-3: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 30)$

Fig. 38 Ex3-I-3: Log plot of the magnitude of $g(t)$.

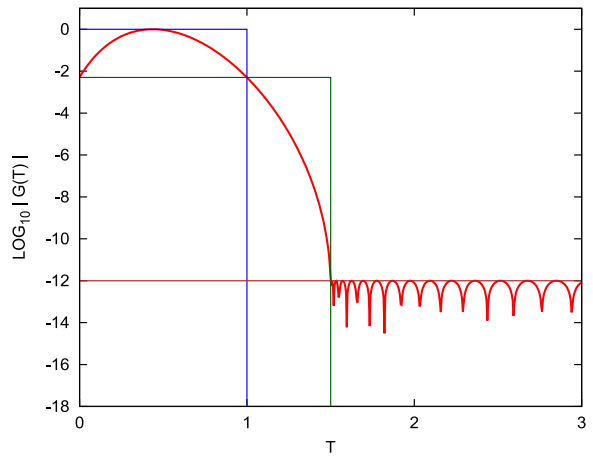


図 41 Ex3-I-4: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 5E-3, n = 40)$

Fig. 41 Ex3-I-4: Log plot of the magnitude of $g(t)$.

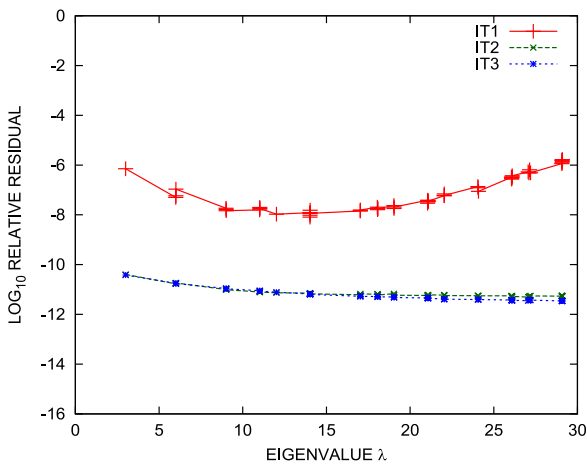


図 39 Ex3-I-3: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 1E-3, n = 30)$

Fig. 39 Ex3-I-3: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

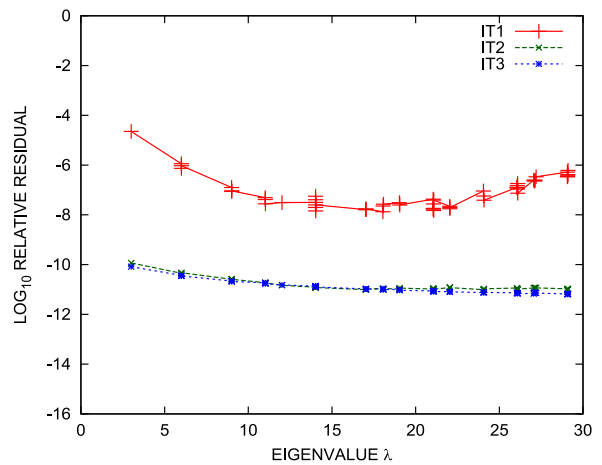


図 42 Ex3-I-4: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-12, g_p = 5E-3, n = 40)$

Fig. 42 Ex3-I-4: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

表 6 Ex3-I-5: Remez 法と同様の反復による z の値の収束状況
Table 6 Ex3-I-5: Convergence of z by Remez-like iteration.

反復回数	z
0	0.25
1	0.34808106022048742
2	0.36663347095716745
3	0.36726156850974075
4	0.36726227523852140
5	0.36726227523945272
6	0.36726227523945272

は g_p の値を 1E-3 から 5E-3 に増したのと引き換えに, Chebyshev 多項式の次数 n の値は 30 から 40 に増えている.

3.3.5 フィルタ Ex3-I-5

パラメタの組は $\mu = 1.5$, $g_s = 1.2E-12$, $g_p = 1E-3$, $n = 20$ である. それを満たす解として式 (31) が得られた (パラメタ g_s の値が 1.2E-12 であって 1E-12 ではないのは, そうしなければ解が得られなかったからである). Chebyshev 多項式の次数 n が小さく 20 であることは長所ではあるが, 3つの極のうち τ_2 と τ_3 の位置が近く, それら 2つの (異符号である) 極の係数の大きさが 4桁になっていることは, それらのレゾルベントの作用を適用した結果の線形結合を計算する段階で数値相殺による丸め誤差が拡大されて数値の相対精度が減る傾向が強いと考えられる.

$$\begin{cases} \beta=1.0, \tau_1=-1.73, \tau_2=-2.65, \tau_3=-2.87, \\ z=0.36726227523945272, \\ \alpha_1=-207.17748653406107, \\ \alpha_2=2173.8199744641517, \\ \alpha_3=-2008.7596713432417. \end{cases} \quad (31)$$

Remez 法と同様の反復による収束の状況を z の初期値を 0.25 とした場合について表 6 に示す.

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 43 と図 44 に示す.

3.3.6 フィルタ Ex3-I-6

パラメタの組は $\mu = 1.5$, $g_s = 2E-14$, $g_p = 1E-3$, $n = 30$ である. それを満たす解として式 (32) が得られた.

$$\begin{cases} \beta=1.0, \tau_1=-3.3, \tau_2=-3.4, \tau_3=-3.6, \\ z=0.40544273717608303, \\ \alpha_1=-27705.174519370320, \\ \alpha_2=43785.112858715520, \\ \alpha_3=-16135.512395505596. \end{cases} \quad (32)$$

Remez 法と同様の反復による収束の状況を z の初期値を 0.25 とした場合について表 7 に示す.

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 46 と図 47 に示す.

フィルタ Ex3-I-5 に比べるとフィルタ Ex3-I-6 では g_s の

表 7 Ex3-I-6: Remez 法と同様の反復による z の値の収束状況
Table 7 Ex3-I-6: Convergence of z by Remez-like iteration.

反復回数	z
0	0.25
1	0.37617020656913519
2	0.40409760233387351
3	0.40543969192635268
4	0.40544273716025054
5	0.40544273717608303
6	0.40544273717608303

値は 2E-14 で約 2桁小さい値であり, g_p の値は 1E-3 であって同じになっている. それと引き換えに Chebyshev 多項式の次数は 20 から 30 に増えている. Ex3-I-6 では g_s の値自体はかなり小さくできたのであるが, 欠点としては 3つの極の相互の位置が近く, 各極の係数の大きさが 5桁にまでなっていて, そのことから各レゾルベントの作用を適用して得られたベクトルの組の間で線形結合を計算する段階で生じる数値相殺による丸め誤差の拡大により, 数値の相対精度を失う傾向が強いと考えられる.

3.4 レゾルベントを 3つ用いた構成の例 (方式 II)

レゾルベント 3つと恒等演算子の線形結合の伝達関数を式 (24) とする. $\alpha_j, \tau_j, j = 1, 2, 3$ と β はすべて実数であり, 3つの極 τ_j はすべて負である. 有理関数 $y(t)$ についての制約条件は式 (7) である.

フィルタの設計パラメタとして μ, g_s, g_p, n を指定する (ただし $1 < \mu, 0 < g_s \ll g_p < 1$). そうして $y(t)$ の閾値 y_l と y_h を式 (5) により計算する.

そうしてこの方式 II では通過域 $t \in [0, 1]$ において $y(t)$ が座標 $t = 0$ と $t = z$ ($0 < z < 1$) の 2カ所において最大値をとるものとする. 有理関数 $y(t)$ は少なくとも式 (33) の 4つの条件を満たす必要がある.

$$\begin{cases} y(0) = y_h, \\ y(z) = y_h, \\ y(1) = y_l, \\ y(\mu) = 1, \end{cases} \quad (33)$$

極 τ_1, τ_2, τ_3 の値を仮定して与えよこの 4つの条件を表す連立 1次方程式 (34) の数値が決まり, それを解くことで $\alpha_1, \alpha_2, \alpha_3, \beta$ の値が求まる.

$$\begin{bmatrix} 1/(0-\tau_1) & 1/(0-\tau_2) & 1/(0-\tau_3) & 1 \\ 1/(z-\tau_1) & 1/(z-\tau_2) & 1/(z-\tau_3) & 1 \\ 1/(1-\tau_1) & 1/(1-\tau_2) & 1/(1-\tau_3) & 1 \\ 1/(\mu-\tau_1) & 1/(\mu-\tau_2) & 1/(\mu-\tau_3) & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta \end{bmatrix} = \begin{bmatrix} y_h \\ y_h \\ y_l \\ 1 \end{bmatrix} \quad (34)$$

得られた $y(t)$ について $t = z$ が区間 (0, 1) において $y(t)$ の唯一の極大点となっていてさらに通過域での最小値が y_l 以

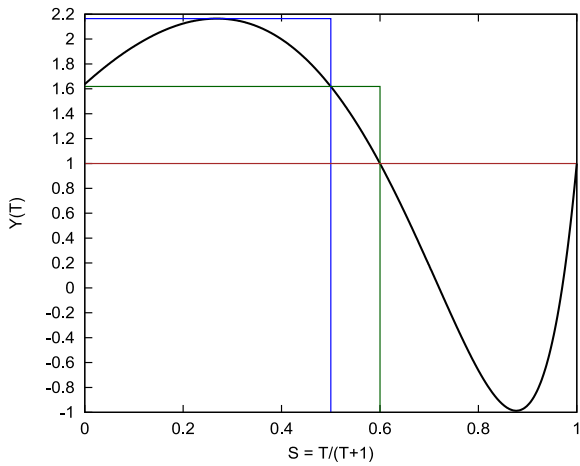


図 43 Ex3-I-5: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1.2E-12, g_p = 1E-3, n = 20)$

Fig. 43 Ex3-I-5: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

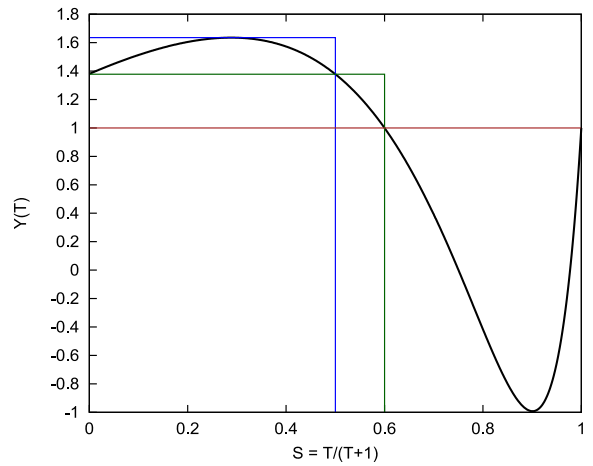


図 46 Ex3-I-6: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 2E-14, g_p = 1E-3, n = 30)$

Fig. 46 Ex3-I-6: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

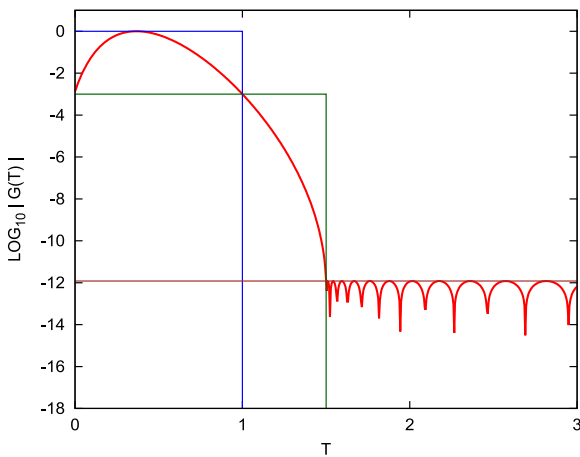


図 44 Ex3-I-5: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1.2E-12, g_p = 1E-3, n = 20)$

Fig. 44 Ex3-I-5: Log plot of the magnitude of $g(t)$.

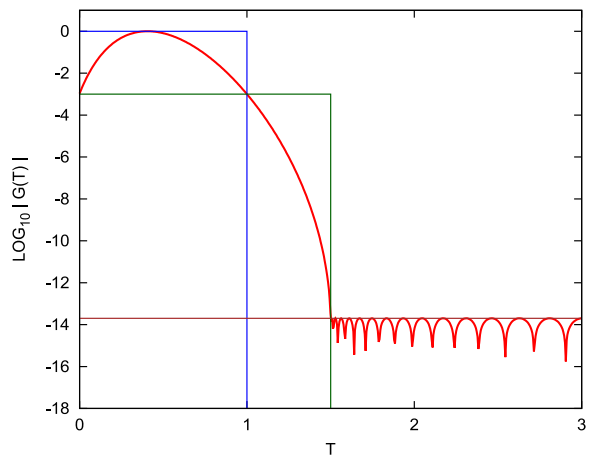


図 47 Ex3-I-6: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 2E-14, g_p = 1E-3, n = 30)$

Fig. 47 Ex3-I-6: Log plot of the magnitude of $g(t)$.

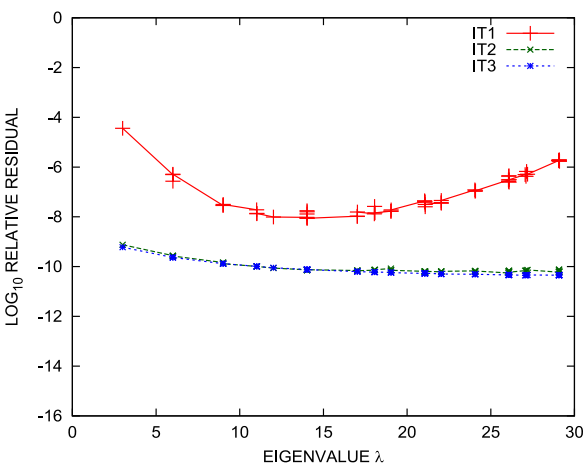


図 45 Ex3-I-5: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1.2E-12, g_p = 1E-3, n = 20)$

Fig. 45 Ex3-I-5: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

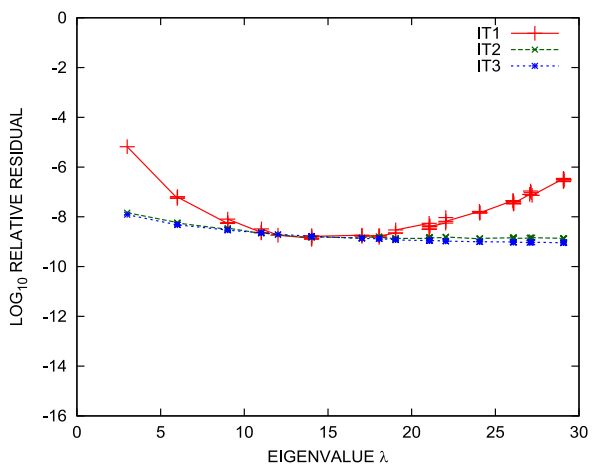


図 48 Ex3-I-6: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 2E-14, g_p = 1E-3, n = 30)$

Fig. 48 Ex3-I-6: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

表 8 レゾルベントを 3 つ用いたフィルタのパラメタの例 (方式 II)
Table 8 Parameters of filters which consist of 3 resolvents (method II).

フィルタ	μ	g_s	g_p	n
Ex3-II-1	1.5	1E-10	1E-2	30
Ex3-II-2	1.5	1E-11	1E-2	34
Ex3-II-3	1.5	1E-12	1E-2	41

表 9 Ex3-II-1 : Remez 法と同様の反復による z の値の収束状況
Table 9 Ex3-II-1: Convergence of z by Remez-like iteration.

反復回数	z
0	0.25
1	0.43889995782449841
2	0.49390779973752797
3	0.49926776890549809
4	0.49931829133629799
5	0.49931829581409693
6	0.49931829581409693

上であり, またさらに阻止域 $[\mu, \infty)$ で $|y(t)|$ が 1 を超えることがなければ $y(t)$ の構成は一応成功である.

もしも $y(t)$ が区間 $(0, 1)$ の中で $t = z$ で極大でなければ, 極大点の座標を求めてそれを次回の z の値にして繰り返しにより収束させる (Remez 法と同様の反復).

通過域での $y(t)$ の最小値が y_ℓ よりも小さかったり, 阻止域で $|y(t)|$ の値が 1 を超えることがあれば (特に $|\beta| > 1$ であれば) 構成は失敗であり, その場合は極の座標の値 τ_1, τ_2, τ_3 を変更して探索を行う. 今回はこの探索は人の手による試行錯誤で行っている. 指定した μ, g_s, g_p, n の値によっては数学的に適切な解が存在しない場合がある.

実数シフトのレゾルベントを 3 つ用いて構成できたフィルタの 3 通りの例について, 指定したパラメタ μ, g_s, g_p, n の値を表 8 に示す. これらは μ の値はすべて 1.5 で, g_p の値もすべて 1E-2 であり, 順に g_s の値を 1E-10, 1E-11, 1E-12 と下げて, それと引き換えに Chebyshev 多項式の次数 n を 30, 34, 41 と増している.

3.4.1 フィルタ Ex3-II-1

パラメタの組は $\mu = 1.5, g_s = 1E-10, g_p = 1E-2, n = 30$ である. それを満たす解として式 (35) が得られた.

$$\begin{cases} \tau_1 = -0.60, \tau_2 = -0.69, \tau_3 = -0.81, \\ z = 0.49931829581409693, \\ \beta = -0.99901747981329442, \\ \alpha_1 = 89.321336271875544, \\ \alpha_2 = -210.39740961563449, \\ \alpha_3 = 128.29029527935843, \end{cases} \quad (35)$$

Remez 法と同様の反復による収束の状況を z の初期値を 0.25 とした場合について表 9 に示す.

得られた有理関数 $y(t)$ のグラフを図 49 に, 伝達関数の大きさ $|g(t)|$ の対数グラフを図 50 に示す.

3.4.2 フィルタ Ex3-II-2

パラメタの組は $\mu = 1.5, g_s = 1E-11, g_p = 1E-2, n = 34$ である. それを満たす解として式 (36) が得られた.

$$\begin{cases} \tau_1 = -0.63, \tau_2 = -0.64, \tau_3 = -0.74, \\ z = 0.54377026128768921, \\ \beta = -0.98131600973886953, \\ \alpha_1 = 1805.9913175101005, \\ \alpha_2 = -2052.6571416016740, \\ \alpha_3 = 253.75537272985273. \end{cases} \quad (36)$$

得られた有理関数 $y(t)$ のグラフを図 52 に, 伝達関数の大きさ $|g(t)|$ の対数グラフを図 53 に示す.

このフィルタ Ex3-II-2 では 3 つの極のうち τ_1 と τ_2 の位置がかなり接近しており, それにともないそれら 2 つの (異符号である) 極の係数の絶対値が 4 桁程度と大きい. それにより数値相殺による丸め誤差の拡大を招いて計算結果の相対精度を低下させる傾向が強いことが懸念される.

3.4.3 フィルタ Ex3-II-3

パラメタの組は $\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 41$ である. それを満たす解として式 (37) が得られた.

$$\begin{cases} \tau_1 = -0.722, \tau_2 = -0.820, \tau_3 = -0.837, \\ z = 0.55822496403008699, \\ \beta = -0.99988504442730867, \\ \alpha_1 = 236.05420176688206, \\ \alpha_2 = -2101.8656374776586, \\ \alpha_3 = 1873.6697420694337. \end{cases} \quad (37)$$

得られた有理関数 $y(t)$ のグラフを図 55 に, 伝達関数の大きさ $|g(t)|$ の対数グラフを図 56 に示す.

このフィルタ Ex3-II-3 では 3 つの極のうち τ_2 と τ_3 の位置がかなり接近しており, それにともないそれら 2 つの (異符号である) 極の係数の絶対値が 4 桁程度と大きい. そのことから数値相殺による丸め誤差の拡大を招いて計算結果の相対精度を低下させる傾向が強いであろうことが懸念される.

3.5 レゾルベントを 4 つ用いた構成の例 (方式 I)

レゾルベント 4 つと恒等演算子の線形結合に対する伝達関数である $y(t)$ を式 (38) とする. ここで $\alpha_j, \tau_j, j=1, 2, 3, 4$ と β はすべて実数とし, 4 つの極 τ_j はすべて負とする.

$$y(t) = \sum_{j=1}^4 \frac{\alpha_j}{t - \tau_j} + \beta. \quad (38)$$

有理関数 $y(t)$ についての制約条件は式 (7) になる. フィルタの設計パラメタとして μ, g_s, g_p, n の値を指定する (ただし $1 < \mu, 0 < g_s \ll g_p < 1$ である). $y(t)$ の閾値 y_ℓ と y_h は式 (5) から計算で求める.

通過域 $[0, 1]$ において $y(t)$ は $t = 0$ で最大で, そのほか

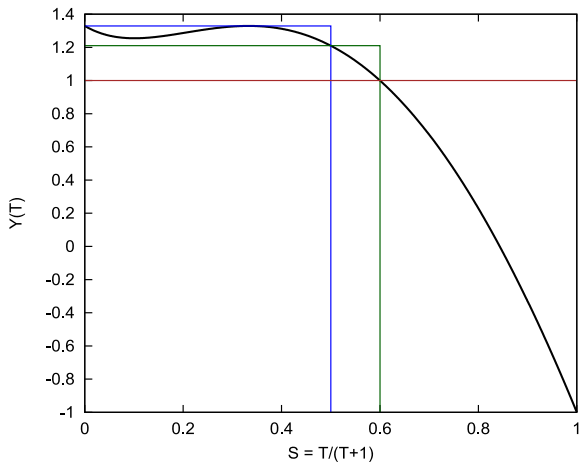


図 49 Ex3-II-1: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-2, n = 30)$

Fig. 49 Ex3-II-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

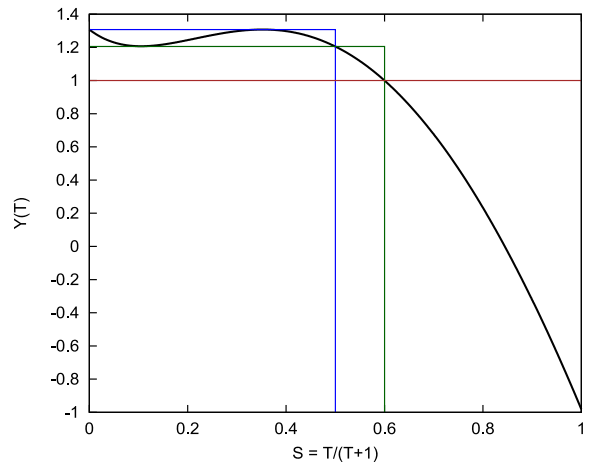


図 52 Ex3-II-2: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
 $(\mu = 1.5, g_s = 1E-11, g_p = 1E-2, n = 34)$

Fig. 52 Ex3-II-2: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

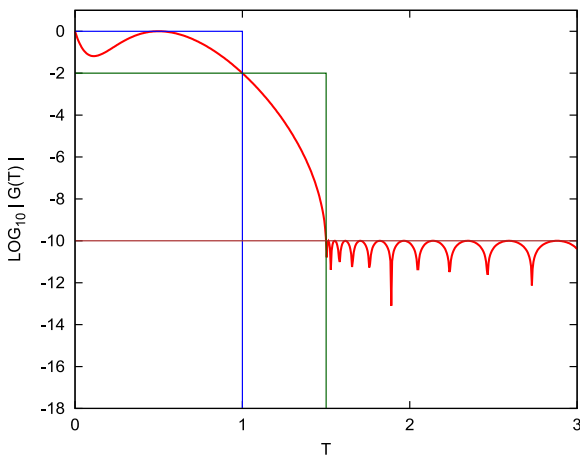


図 50 Ex3-II-1: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-2, n = 30)$

Fig. 50 Ex3-II-1: Log plot of the magnitude of $g(t)$.

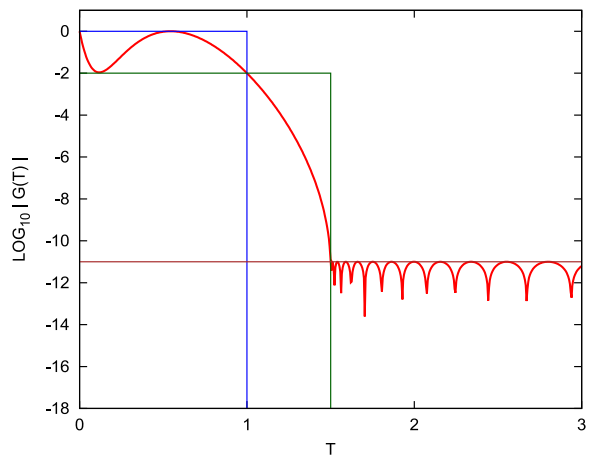


図 53 Ex3-II-2: フィルタ伝達関数の大きさ $|g(t)|$ の対数
 $(\mu = 1.5, g_s = 1E-11, g_p = 1E-2, n = 34)$

Fig. 53 Ex3-II-2: Log plot of the magnitude of $g(t)$.

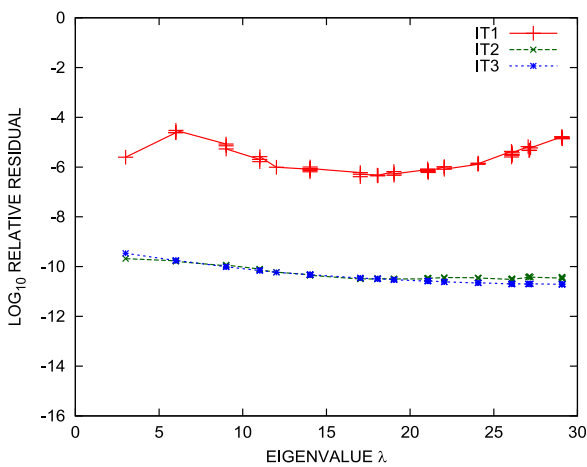


図 51 Ex3-II-1: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-10, g_p = 1E-2, n = 30)$

Fig. 51 Ex3-II-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

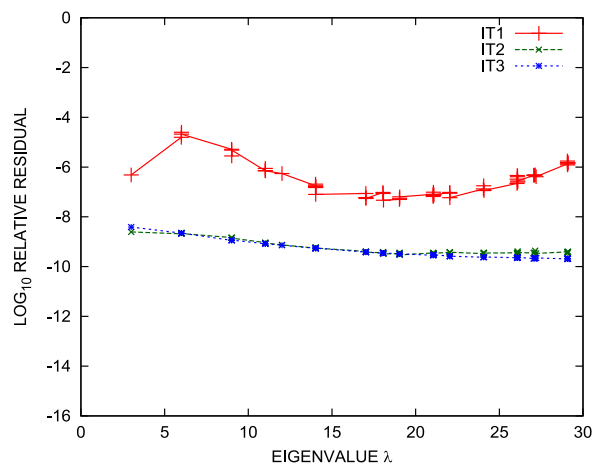


図 54 Ex3-II-2: 近似固有対の相対残差の大きさの対数
 $(\mu = 1.5, g_s = 1E-11, g_p = 1E-2, n = 34)$

Fig. 54 Ex3-II-2: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

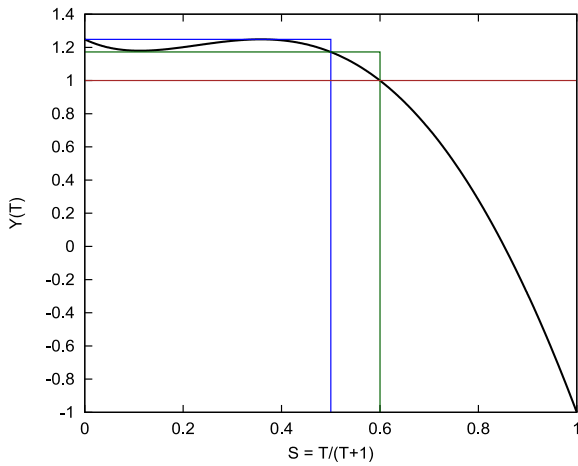


図 55 Ex3-II-3 : 有理関数 $y(t)$ の値 (横軸 $s = t/(t + 1)$) ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 41$)

Fig. 55 Ex3-II-3: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t + 1)$).

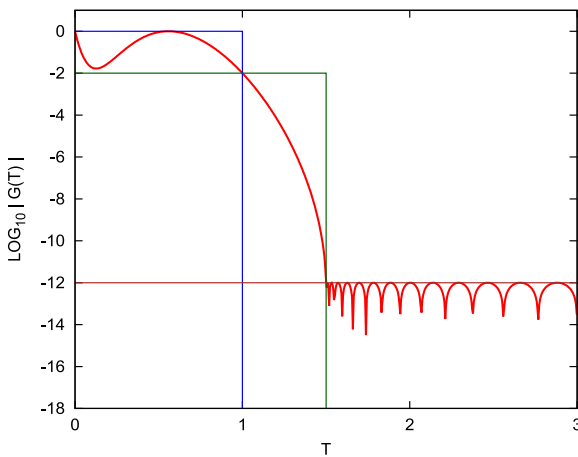


図 56 Ex3-II-3 : フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 41$)

Fig. 56 Ex3-II-3: Log plot of the magnitude of $g(t)$.

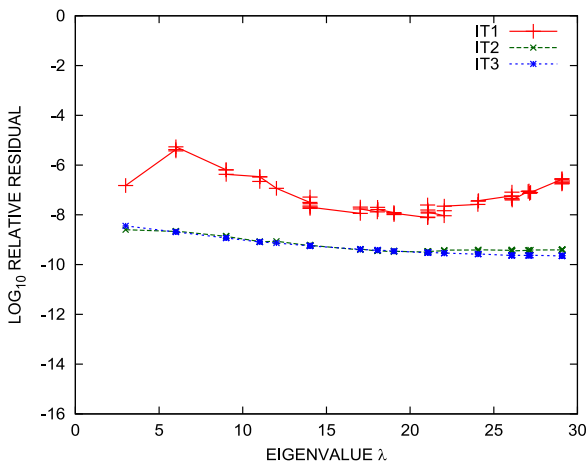


図 57 Ex3-II-3 : 近似固有対の相対残差の大きさの対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 41$)

Fig. 57 Ex3-II-3: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

に最小と極大になる点の座標をそれぞれ $t = z_2$ と $t = z_1$ とする ($0 < z_2 < z_1 < 1$). そうして $y(t)$ に式 (39) の 5 つの等式条件を課した (本来は極小点 z_2 と極大点 z_1 については, それぞれ等式よりも緩い不等式制約 $y(z_2) \geq y_\ell$ と $y(z_1) \leq y_h$ を満たせばそれで十分である).

$$\begin{cases} y(0) = y_h, \\ y(z_2) = y_\ell, \\ y(z_1) = y_h, \\ y(1) = y_\ell, \\ y(\mu) = 1. \end{cases} \quad (39)$$

仮定した 4 つの極の位置 $\tau_j, j=1, 2, 3, 4$ と極値点の座標 z_2, z_1 から連立 1 次方程式 (40) の数値が決まり, それを解くことにより $\alpha_j, j=1, 2, 3, 4$ と β の数値がそれぞれ決まり, 式 (38) の $y(t)$ の定数の数値がすべて決まる.

$$\begin{bmatrix} 1/(0-\tau_1) & 1/(0-\tau_2) & 1/(0-\tau_3) & 1/(0-\tau_4) & 1 \\ 1/(z_2-\tau_1) & 1/(z_2-\tau_2) & 1/(z_2-\tau_3) & 1/(z_2-\tau_4) & 1 \\ 1/(z_1-\tau_1) & 1/(z_1-\tau_2) & 1/(z_1-\tau_3) & 1/(z_1-\tau_4) & 1 \\ 1/(1-\tau_1) & 1/(1-\tau_2) & 1/(1-\tau_3) & 1/(1-\tau_4) & 1 \\ 1/(\mu-\tau_1) & 1/(\mu-\tau_2) & 1/(\mu-\tau_3) & 1/(\mu-\tau_4) & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta \end{bmatrix} = \begin{bmatrix} y_h \\ y_\ell \\ y_h \\ y_\ell \\ 1 \end{bmatrix} \quad (40)$$

得られた $y(t)$ の通過域の内部 $(0, 1)$ での最大点と最小点の座標がそれぞれ z_1 と z_2 にちょうど一致して, 通過域での最小値が y_ℓ になっていてさらに阻止域 $[\mu, \infty)$ で $|y(t)|$ の値が 1 を超えることがなければ条件を満足する $y(t)$ の構成は成功であるが, 通過域の内部における最大点と最小点の座標が z_1 と z_2 でない場合には, 最大点と最小点の座標を求めて次回の z_1 と z_2 に用いて繰り返して収束させる (Remez 法と同様の反復). 収束を得たときに阻止域で $|y(t)|$ の値が 1 を超えていれば構成は失敗であり, その場合には極の位置 $\tau_j, j=1, 2, 3, 4$ を変更してさらに探索を行う. この探索は今回は人の手により試行錯誤で行った. 指定したフィルタのパラメタ μ, g_s, g_p, n の値により条件を満たす解が存在しない場合があるが, 解が存在するかしないかを判定する手段は得られていない.

実数シフトのレゾルベントを 4 つ用いて実際に構成できた方式 I のフィルタの 3 通りの例について, 指定したパラメタ μ, g_s, g_p, n の値を表 10 に示す (これらの例はまだ改良できる可能性がある). これらはすべて $\mu = 1.5$ と $g_p = 1E-2$ であり, 順に g_s の値を $1E-12, 1E-15, 1E-16$ と小さくして, それにともなって Chebyshev 多項式の次数 n は 23, 40, 50 と増加させている.

表 10 レゾルベントを 4 つ用いたフィルタのパラメタの例 (方式 I)
Table 10 Parameters of filters which consist of 4 resolvents (method I).

フィルタ	μ	g_s	g_p	n
Ex4-I-1	1.5	1E-12	1E-2	23
Ex4-I-2	1.5	1E-15	1E-2	40
Ex4-I-3	1.5	1E-16	1E-2	50

表 11 Ex4-I-1 : Remez 法と同様の反復での z_2 と z_1 の収束状況
Table 11 Ex4-I-1: Convergence of z_2 and z_1 by Remez-like method.

反復回数	z_2	z_1
0	0.25	0.50
1	0.12414214464760153	0.58253424709988758
2	0.12495779967692215	0.56005494319787249
3	0.12521108901675325	0.56045273980731145
4	0.12521115661819931	0.5604526995554977
5	0.12521115661819931	0.5604526995554977

3.5.1 フィルタ Ex4-I-1

パラメタの組は $\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 23$ である. それを満たす解として式 (41) が得られた.

$$\left\{ \begin{array}{l} \tau_1 = -0.46, \tau_2 = -1.3, \tau_3 = -1.6, \tau_4 = -2.2, \\ z_2 = 0.12521115661819931, \\ z_1 = 0.5604526995554977, \\ \beta = 0.72830157094489884, \\ \alpha_1 = 11.103609016858801, \\ \alpha_2 = -747.44207460551708, \\ \alpha_3 = 1358.7632395652740, \\ \alpha_4 = -654.01397054308215. \end{array} \right. \quad (41)$$

Remez 法と同様の反復の収束状況を z_2 と z_1 の初期値をそれぞれ 0.25 と 0.5 とした場合について表 11 に示す.

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 58 と図 59 に示す.

3.5.2 フィルタ Ex4-I-2

パラメタの組は $\mu = 1.5, g_s = 1E-15, g_p = 1E-2, n = 40$ である. それを満たす解として式 (42) が得られた.

$$\left\{ \begin{array}{l} \tau_1 = -0.8, \tau_2 = -0.85, \tau_3 = -3.4, \tau_4 = -4.4, \\ z_2 = 0.14464146269159392, \\ z_1 = 0.60332442698162049, \\ \beta = 0.73890252752857741, \\ \alpha_1 = 175.18672767961545, \\ \alpha_2 = -206.28364424067430, \\ \alpha_3 = 413.29017809312039, \\ \alpha_4 = -427.58306042439581. \end{array} \right. \quad (42)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 61 と図 62 に示す.

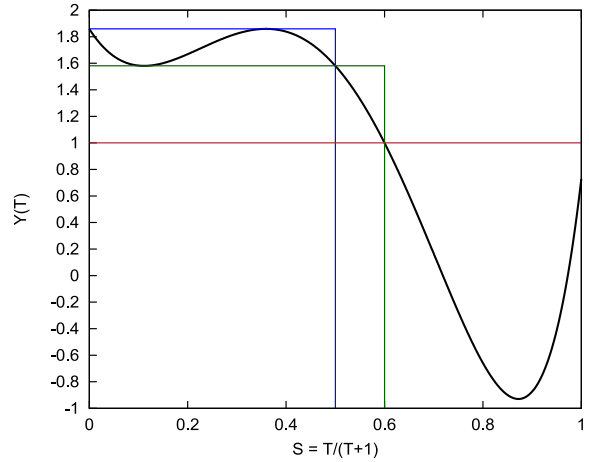


図 58 Ex4-I-1 : 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$) ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 23$)

Fig. 58 Ex4-I-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

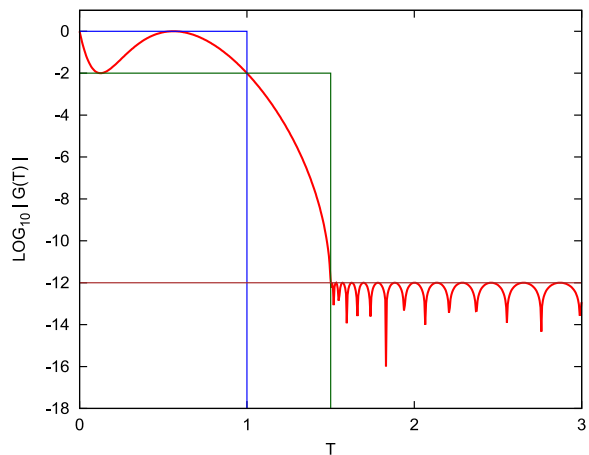


図 59 Ex4-I-1 : フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 23$)

Fig. 59 Ex4-I-1: Log plot of the magnitude of $g(t)$.

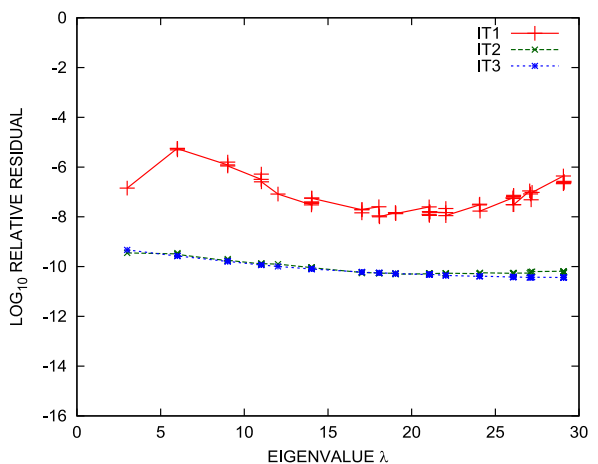


図 60 Ex4-I-1 : 近似固有対の相対残差の大きさの対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-2, n = 23$)

Fig. 60 Ex4-I-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

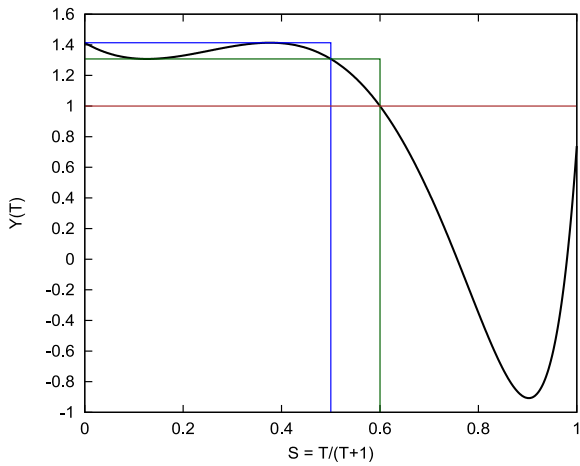


図 61 Ex4-I-2: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$)
($\mu = 1.5, g_s = 1E-15, g_p = 1E-2, n = 40$)

Fig. 61 Ex4-I-2: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

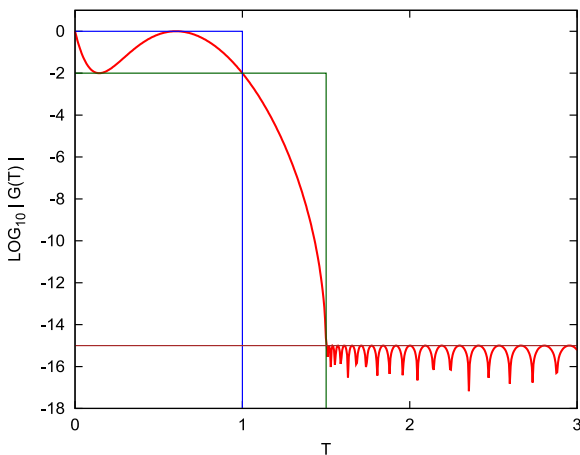


図 62 Ex4-I-2: フィルタ伝達関数の大きさ $|g(t)|$ の対数
($\mu = 1.5, g_s = 1E-15, g_p = 1E-2, n = 40$)

Fig. 62 Ex4-I-2: Log plot of the magnitude of $g(t)$.

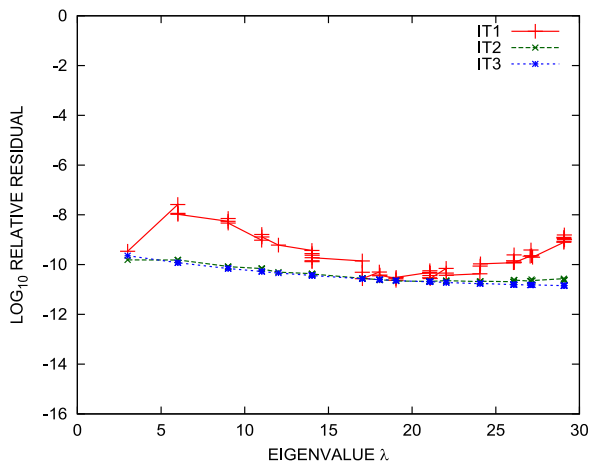


図 63 Ex4-I-2: 近似固有対の相対残差の大きさの対数
($\mu = 1.5, g_s = 1E-15, g_p = 1E-2, n = 40$)

Fig. 63 Ex4-I-2: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

3.5.3 フィルタ Ex4-I-3

パラメタの組は $\mu = 1.5, g_s = 1E-16, g_p = 1E-2, n = 50$ である. それを満たす解として式 (43) が得られた.

$$\begin{cases} \tau_1 = -0.5, \tau_2 = -1.4, \tau_3 = -4.9, \tau_4 = -5.9, \\ z_2 = 0.14592721809167415, \\ z_1 = 0.61110848092101514, \\ \beta = 0.93709121223898141, \\ \alpha_1 = 1.9546378129488793, \\ \alpha_2 = -26.920052689841116, \\ \alpha_3 = 639.14392557923146, \\ \alpha_4 = -677.08425058642284. \end{cases} \quad (43)$$

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 64 と図 65 に示す.

3.6 レゾルベントを 4 つ用いた構成の例 (方式 II)

レゾルベント 4 つと恒等演算子の線形結合に対する伝達関数を式 (38) とする. そうして $\alpha_j, \tau_j, j=1,2,3,4$ と β はすべて実数とし, 4 つの極 τ_j はすべて負であるとする. 有理関数 $y(t)$ についての制約条件は式 (7) になる. フィルタの設計パラメタとして μ, g_s, g_p, n の値を指定する (ただし $1 < \mu, 0 < g_s \ll g_p < 1$ である). $y(t)$ の閾値 y_ℓ と y_h は式 (5) から計算で求める.

通過域 $[0, 1]$ において $y(t)$ が最大値 y_h をとる点の座標を $t = z_3$ と $t = z_1$ とし, 最小値 y_ℓ をとる点の座標を $t = z_2$ とする (ただし $0 \leq z_3 < z_2 < z_1 < 1$ とする). すると $y(t)$ は式 (44) の 5 つの条件を満たす必要がある.

$$\begin{cases} y(z_3) = y_h, \\ y(z_2) = y_\ell, \\ y(z_1) = y_h, \\ y(1) = y_\ell, \\ y(\mu) = 1. \end{cases} \quad (44)$$

仮定する 4 つの極の位置 $\tau_j, j=1,2,3,4$ と極値点の初期値の座標 z_3, z_2, z_1 から連立 1 次方程式 (45) の数値が決まり, それを解くことにより $\alpha_j, j=1,2,3,4$ と β の数値がそれぞれ求まる. それにより式 (38) の $y(t)$ の定数の数値がすべて決まる.

$$\begin{bmatrix} 1/(z_3 - \tau_1) & 1/(z_3 - \tau_2) & 1/(z_3 - \tau_3) & 1/(z_3 - \tau_4) & 1 \\ 1/(z_2 - \tau_1) & 1/(z_2 - \tau_2) & 1/(z_2 - \tau_3) & 1/(z_2 - \tau_4) & 1 \\ 1/(z_1 - \tau_1) & 1/(z_1 - \tau_2) & 1/(z_1 - \tau_3) & 1/(z_1 - \tau_4) & 1 \\ 1/(1 - \tau_1) & 1/(1 - \tau_2) & 1/(1 - \tau_3) & 1/(1 - \tau_4) & 1 \\ 1/(\mu - \tau_1) & 1/(\mu - \tau_2) & 1/(\mu - \tau_3) & 1/(\mu - \tau_4) & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta \end{bmatrix} = \begin{bmatrix} y_h \\ y_\ell \\ y_h \\ y_\ell \\ 1 \end{bmatrix} \quad (45)$$

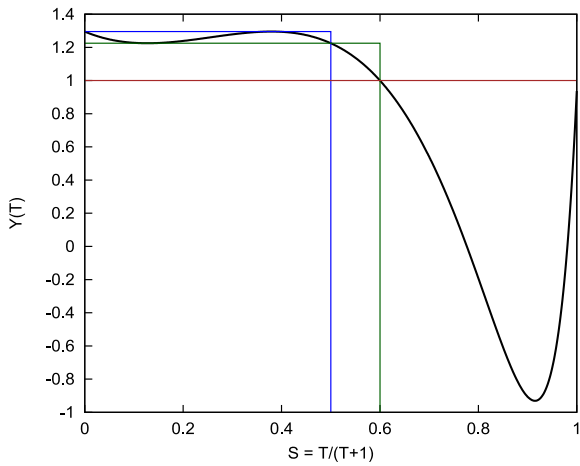


図 64 Ex4-I-3: 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$) ($\mu = 1.5, g_s = 1E-16, g_p = 1E-2, n = 50$)

Fig. 64 Ex4-I-3: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

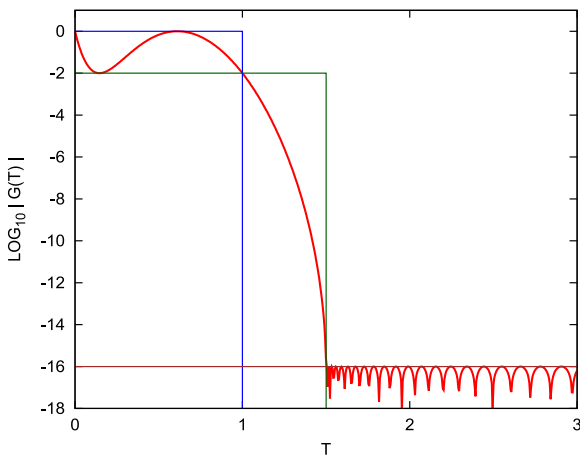


図 65 Ex4-I-3: フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5, g_s = 1E-16, g_p = 1E-2, n = 50$)

Fig. 65 Ex4-I-3: Log plot of the magnitude of $g(t)$.

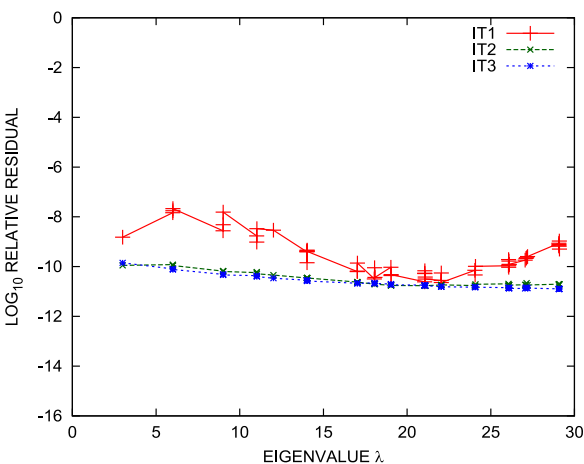


図 66 Ex4-I-3: 近似固有対の相対残差の大きさの対数 ($\mu = 1.5, g_s = 1E-16, g_p = 1E-2, n = 50$)

Fig. 66 Ex4-I-3: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

表 12 レゾルベントを 4 つ用いたフィルタのパラメタの例 (方式 II)
Table 12 Parameters of a filter which consists of 4 resolvents (method II).

フィルタ	μ	g_s	g_p	n
Ex4-II-1	1.5	1E-12	1E-1	63

こうして得られた $y(t)$ の通過域の内部 $[0, 1)$ での最大点, 最小点, 最大点の座標値がそれぞれ仮定した z_3, z_2, z_1 にちょうど一致していて, 通過域での最小値が y_ℓ になっていて, さらに阻止域 $[\mu, \infty)$ で $|y(t)|$ の値が 1 を超えることがなければ条件を満足する $y(t)$ の構成は成功であるが, もしもそうではなくて通過域の内部における極値点の座標がずれて一致していない場合には区間 $[0, 1)$ での最大点と最小点の座標を求めてそれらを次回の z_1 と z_2, z_3 として用いて繰り返しを行って収束させる (Remez 法と同様の反復). 収束が得られたときに阻止域での $|y(t)|$ の値の最大値が 1 を超えていたら構成は失敗であり, その場合には 4 つの極の位置 $\tau_j, j=1, 2, 3, 4$ を変更してさらに探索を行う. この探索は現在は人手による試行錯誤で行っている. 指定したフィルタのパラメタ μ, g_s, g_p, n の値によっては, 条件を満たす解が存在しない場合があるが, 解が存在するかどうかを判定する手段は得られていない.

実数シフトのレゾルベントを 4 つ用いる場合の構成 (方式 II) で得られたフィルタの例を 1 つだけ (Ex4-II-1) 示す (表 12).

3.6.1 フィルタ Ex4-II-1

パラメタの組は $\mu = 1.5, g_s = 1E-12, g_p = 1E-1, n = 63$ である. それを満たす解として式 (46) が得られた.

$$\left\{ \begin{array}{l} \tau_1 = -0.420, \tau_2 = -0.617, \tau_3 = -0.995, \tau_4 = -1.075, \\ z_3 = 0.06554\ 14265\ 33483\ 900, \\ z_2 = 0.29234\ 04793\ 99673\ 64, \\ z_1 = 0.70842\ 49925\ 96358\ 06, \\ \beta = -0.97891\ 04015\ 29457\ 29, \\ \alpha_1 = -8.42245\ 36308\ 00969\ 2, \\ \alpha_2 = 49.95291\ 83198\ 95732, \\ \alpha_3 = -390.72185\ 12457\ 4242, \\ \alpha_4 = 358.88151\ 30449\ 7295, \end{array} \right. \quad (46)$$

z_3, z_2, z_1 の初期値をそれぞれ 0.2, 0.5, 0.8 とした場合の Remez 法と同様の反復による収束状況を表 13 に示す.

有理関数 $y(t)$ とフィルタの伝達関数 $g(t)$ のグラフをそれぞれ図 67 と図 68 に示す.

このフィルタ Ex4-II-1 は g_s の値は $1E-12$ なのであまり小さいとはいえないが, g_p の値は $1E-1$ とかなり 1 に近くして通過域での伝達率の均一性をきわめて高くした例である. しかしそれと引き換えに Chebyshev 多項式の次数 n は 63 とずいぶん大きい. 4 つある極のうちで, τ_3 と τ_4 の

表 13 Ex4-II-1 : Remez 法と同様の反復での z_3, z_2, z_1 の値の収束

Table 13 Ex4-II-1: Convergences of z_3, z_2 and z_1 by Remez-like iteration.

反復	z_3	z_2	z_1
0	0.2	0.5	0.8
1	0.0	0.32123334169460461	0.73899058362834156
2	0.042297743912916631	0.29247225118458271	0.70773677106447518
3	0.062360888396385126	0.29322036388333887	0.70854338823251426
4	0.065474324389370158	0.29235708406649530	0.70842775508902967
5	0.065541396354204044	0.29234048703301698	0.70842499381303787
6	0.065541426533483900	0.29234047939967364	0.70842499259635806
7	0.065541426533483900	0.29234047939967364	0.70842499259635806

2つが比較的近い位置にあり、それらの極の係数の絶対値も3桁程度ある。

4. フィルタを用いて近似対を求める実験

4.1 実験に使用した計算機システムとプログラム

計算機システムは、東京大学情報基盤センターの Oakbridge-CX の 1 ノードである。ノードは 2 つの CPU から成り、各 CPU は Intel Xeon Platinum 8280 (Cascade Lake, 28 コア, 2.7 GHz) である。ノード内の共有主記憶は DDR4 で容量は 192 GiByte であり、ノードの理論最大性能は約 4.8 TFlops (倍精度 binary64) である。計算に使用したコンパイラは Intel Fortran version 19.1.3.304 である。

計算に用いたプログラムは全体を Fortran90 言語を用いて書いたものであり、外部のライブラリコードはまったく利用していない。レゾルベントの作用を実現するための連立 1 次方程式の解法は直接法であり、実対称正定値の帯行列に対するピボット選択をしない修正 Cholesky 法 (LDL^T 分解と前進後退代入) を用いている。現在のプログラムは並列化については Fortran のソースコードに OpenMP の指示を挿入することにより行っており、なるべくブロック化をしたが、それ以外には何も特別なことはしていない。

4.2 実対称定値一般固有値問題の例題

今回の例題に用いた実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ は、解析対象領域を 1 辺の長さ π の立方体とし、立法体の表面における値を零とする境界条件を課した連続関数 $\Psi(x, y, z)$ の Laplace 固有値問題 $-\Delta\Psi = \lambda\Psi$ を有限要素法で離散化して得られたものであり、実験ではその固有対のうちで固有値が固有値分布の下端付近の区間 $[a, b]$ にあるものをフィルタを利用して近似して求めた。有限要素法の離散化には立方体の領域を各辺の方向にそれぞれ N_1+1, N_2+1, N_3+1 に等分して得られる直方体の要素を用いて、要素内の基底関数としては 3 重線形関数を用いた。

領域の分割数は $(N_1, N_2, N_3) = (50, 60, 70)$ とした。一般固有値問題の係数である実対称行列 A と B の次数は $N_1N_2N_3 = 210000$ でその半帯幅は $1 + N_1 + N_1N_2 = 3051$

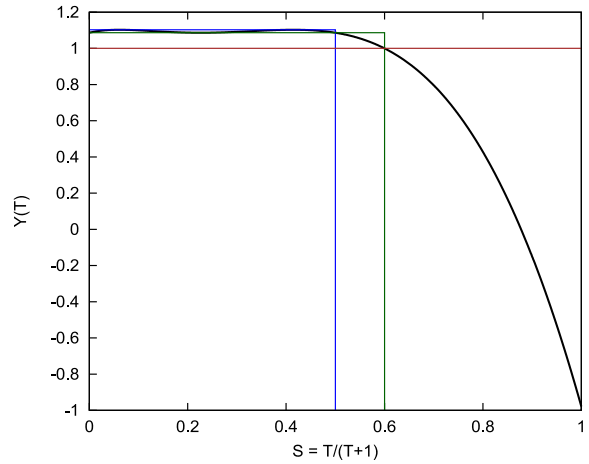


図 67 Ex4-II-1 : 有理関数 $y(t)$ の値 (横軸 $s = t/(t+1)$) ($\mu = 1.5, g_s = 1E-12, g_p = 1E-1, n = 63$)

Fig. 67 Ex4-II-1: The plot of the rational function $y(t)$ (The horizontal axis is $s = t/(t+1)$).

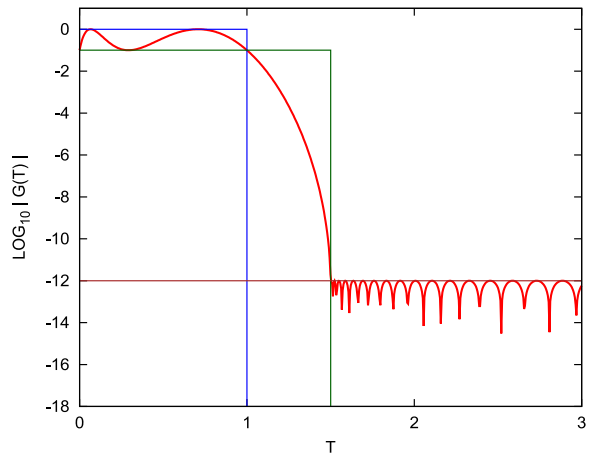


図 68 Ex4-II-1 : フィルタ伝達関数の大きさ $|g(t)|$ の対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-1, n = 63$)

Fig. 68 Ex4-II-1: Log plot of the magnitude of $g(t)$.

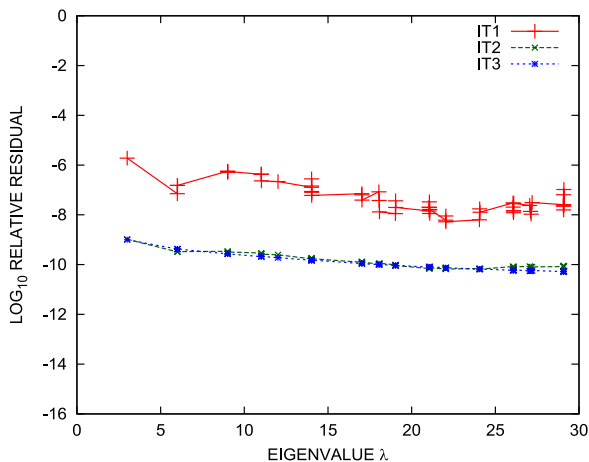


図 69 Ex4-II-1 : 近似固有対の相対残差の大きさの対数 ($\mu = 1.5, g_s = 1E-12, g_p = 1E-1, n = 63$)

Fig. 69 Ex4-II-1: Log plot of the magnitudes of relative residuals of approximate eigenpairs.

である (行列 A と B は両方とも行あるいは列あたりの非零要素数はただか 27 であり, 帯内部はきわめて疎であるが, 今回の実験においてはレゾルベントの作用を実現するための帯行列係数の連立 1 次方程式を修正 Cholesky 法を用いて解く際には, 帯内は密であるかのように扱っている).

この問題の固有値は数学的に必ず 3 以上になるので, 固有値が区間 $[a, b] = [3, 30]$ にある固有対をすべて求めた. 該当する固有対は全部で 54 個ある. 求める固有値の区間を $[0, 30]$ にしなかった理由は, フィルタの伝達関数の値が通過域の下端で下がっているときには固有値が下端に近い固有対の近似の精度が低下するのでなるべくそのような状況を作るためである.

4.3 近似固有対の相対残差の大きさ

近似固有対の品質評価には相対残差の大きさを使用した. 近似固有対 (λ, \mathbf{v}) に対する相対残差の大きさ Θ として, ベクトルの 2-ノルム $\|\cdot\|_2$ により, 式 (47) で表されるものを用いた.

$$\Theta = \frac{\|A\mathbf{v} - \lambda B\mathbf{v}\|_2}{\|\lambda B\mathbf{v}\|_2}. \quad (47)$$

4.4 実験結果

すべての計算は浮動小数点数とその演算に IEEE 754 規格の倍精度 (binary64) を用いて行った. 実験に用いたフィルタとそのパラメータを表 14 に示す. 表中で K は用いたレゾルベントの個数, n は Chebyshev 多項式の次数, そうして μ, g_s, g_p は伝達関数 $g(t)$ の主要な 3 つの形状を表すパラメータの値である.

各フィルタを用いた場合について, 得られた近似固有対の固有値を横軸にとり相対残差の大きさの対数を縦軸にとり, 1 つの図の中にフィルタの反復回数が 1, 2, 3 の各場合に対応する赤線, 緑線, 青線で描いた 3 つの折れ線のグラフ IT1, IT2, IT3 をプロットしたものを示す (Ex1-1 (図 5), Ex1-2 (図 8), Ex1-3 (図 11), Ex1-4 (図 14), Ex1-5 (図 17), Ex2-1 (図 21), Ex2-2 (図 24), Ex2-3 (図 27), Ex2-4 (図 30), Ex3-I-1 (図 33), Ex3-I-2 (図 36), Ex3-I-3 (図 39), Ex3-I-4 (図 42), Ex3-I-5 (図 45), Ex3-I-6 (図 48), Ex3-II-1 (図 51), Ex3-II-2 (図 54), Ex3-II-3 (図 57), Ex4-I-1 (図 60), Ex4-I-2 (図 63), Ex4-I-3 (図 66), Ex4-II-1 (図 69)). なお, フィルタを反復して近似固有対を改良する方法は, 文献 [6], [9] で用いたものと同様である.

表 15 には各フィルタを用いて得られた近似固有対で区間 $[a, b] = [3, 30]$ に固有値があるものすべてについての「相対残差の大きさ」の最大値を示す. この表中のすべての場合について, 得られた近似固有対の数は正しい数 54 と一致していた. この表中の IT1, IT2, IT3 の各列の値がそれ

表 14 実験に用いたフィルタのパラメータ

Table 14 Parameters of filters used in experiments.

フィルタ	K	n	μ	g_s	g_p
Ex1-1	1	27	2.0	1E-12	1E-4
Ex1-2	1	28	1.75	1E-15	1E-6
Ex1-3	1	33	1.5	1E-15	1E-7
Ex1-4	1	25	1.5	1E-16	1E-8
Ex1-5	1	109	1.3	1E-15	1E-8
Ex2-1	2	23	1.5	1E-10	1E-3
Ex2-2	2	38	1.5	1E-12	1E-3
Ex2-3	2	38	1.5	1E-15	1E-5
Ex2-4	2	40	1.5	1.1E-15	1E-4
Ex3-I-1	3	20	1.5	1E-10	1E-3
Ex3-I-2	3	30	1.5	1E-10	1E-2
Ex3-I-3	3	30	1.5	1E-12	1E-3
Ex3-I-4	3	40	1.5	1E-12	5E-3
Ex3-I-5	3	20	1.5	1.2E-12	1E-3
Ex3-I-6	3	30	1.5	2E-14	1E-3
Ex3-II-1	3	30	1.5	1E-10	1E-2
Ex3-II-2	3	34	1.5	1E-11	1E-2
Ex3-II-3	3	41	1.5	1E-12	1E-2
Ex4-I-1	4	23	1.5	1E-12	1E-2
Ex4-I-2	4	40	1.5	1E-15	1E-2
Ex4-I-3	4	50	1.5	1E-16	1E-2
Ex4-II-1	4	63	1.5	1E-12	1E-1

表 15 「相対残差の大きさ」の最大値 (倍精度計算)

$(N_1, N_2, N_3) = (50, 60, 70), [a, b] = [3, 30], nv=110$

Table 15 The largest magnitudes of relative residuals (D-P calculation).

フィルタ	K	n	IT1	IT2	IT3
Ex1-1	1	27	1.4E-03	1.0E-06	3.4E-10
Ex1-2	1	28	4.2E-05	8.5E-11	6.4E-12
Ex1-3	1	33	2.0E-04	8.3E-12	4.6E-12
Ex1-4	1	25	4.8E-03	9.7E-12	1.1E-11
Ex1-5	1	109	3.2E-03	3.6E-12	3.1E-11
Ex2-1	2	23	9.0E-03	3.8E-10	1.5E-11
Ex2-2	2	38	3.4E-05	1.5E-11	1.3E-11
Ex2-3	2	38	1.9E-07	3.7E-12	3.5E-12
Ex2-4	2	40	2.3E-06	3.1E-10	2.9E-10
Ex3-I-1	3	20	2.5E-04	2.9E-11	3.0E-11
Ex3-I-2	3	30	6.2E-04	6.2E-11	4.7E-11
Ex3-I-3	3	30	1.7E-06	4.0E-11	3.8E-11
Ex3-I-4	3	40	2.2E-05	1.2E-10	8.3E-11
Ex3-I-5	3	20	3.6E-05	7.5E-10	6.0E-10
Ex3-I-6	3	30	6.5E-06	1.5E-08	1.2E-08
Ex3-II-1	3	30	3.0E-05	2.1E-10	3.4E-10
Ex3-II-2	3	34	2.5E-05	2.5E-09	3.9E-09
Ex3-II-3	3	41	5.4E-06	2.5E-09	3.6E-09
Ex4-I-1	4	23	5.7E-06	3.6E-10	4.7E-10
Ex4-I-2	4	40	2.6E-08	1.5E-10	2.3E-10
Ex4-I-3	4	50	2.1E-08	1.2E-10	1.4E-10
Ex4-II-1	4	63	1.9E-06	1.0E-09	1.0E-09

表 16 フィルタ対角化全体の経過時間 (単位 sec, 倍精度計算)
 $(N_1, N_2, N_3) = (50, 60, 70)$, $[a, b] = [3, 30]$, $nv=110$

Table 16 Elapsed time of whole filter diagonalization (in sec, D-P calculation).

フィルタ	K	n	IT1	IT2	IT3
Ex1-1	1	27	103.8	181.6	258.3
Ex1-2	1	28	106.2	186.8	265.5
Ex1-3	1	33	120.7	214.2	302.9
Ex1-4	1	25	98.1	167.4	235.0
Ex1-5	1	109	332.6	627.5	915.8
Ex2-1	2	23	155.3	265.5	378.7
Ex2-2	2	38	225.7	404.2	586.7
Ex2-3	2	38	223.9	404.3	579.7
Ex2-4	2	40	234.2	429.9	606.9
Ex3-I-1	3	20	199.2	332.6	474.0
Ex3-I-2	3	30	265.2	464.5	667.1
Ex3-I-3	3	30	265.1	465.0	673.8
Ex3-I-4	3	40	331.6	599.3	872.2
Ex3-I-5	3	20	200.8	334.1	471.4
Ex3-I-6	3	30	267.0	467.7	671.0
Ex3-II-1	3	30	264.4	467.4	674.3
Ex3-II-2	3	34	291.3	518.9	752.2
Ex3-II-3	3	41	338.0	611.3	894.1
Ex4-I-1	4	23	282.9	476.8	677.4
Ex4-I-2	4	40	425.2	771.8	1,106.9
Ex4-I-3	4	50	509.5	946.4	1,373.9
Ex4-II-1	4	63	617.4	1,150.2	1,708.8

ぞれフィルタの反復回数を 1, 2, 3 とした場合の「相対残差の大きさ」の最大値である。列 K はフィルタを構成する実数シフトのレゾルベントの数を、列 n は Chebyshev 多項式の次数を示している。

表 16 には、乱数で生成された $nv=110$ 個のベクトルの組から始めてフィルタを利用して必要な近似固有対を得るまでの経過時間を示す。表中で列 IT1, IT2, IT3 の値はそれぞれフィルタの反復回数が 1, 2, 3 の場合の経過時間 (単位は秒) である (各回のフィルタを適用する前に、ベクトルの組が一般固有値問題の係数行列 B について正規直交基底となるようにベクトルの組に対して B -直交化を施している)。与えられたベクトルの組に対してフィルタを 1 回適用する処理の内側では、 K 個のレゾルベントをベクトルの組に適用して得られた結果の線形結合をとるという操作が、順次に n 回繰り返されている。

表 17 には、フィルタの適用 1 回 (IT=1) により必要な近似固有対を求めるまでの経過時間の内訳の主要なものを、秒を単位として四捨五入により小数点以下 1 桁目までを示す。表中で「行列分解」は、使用する K 個のレゾルベントの作用を実現するために、必要な連立 1 次方程式に対応する K 個のシフト行列 $(A - \rho_i B)$ をそれぞれ修正 Cholesky 分解して主記憶上に保持するまでに消費した経過時間である (半帯幅が w_L の実対称帯行列 1 つの Cholesky 分解に

表 17 フィルタ適用 1 回 (IT=1) の場合の経過時間内訳 (単位 sec, 倍精度) $(N_1, N_2, N_3)=(50, 60, 70)$, $[a, b]=[3, 30]$, $nv=110$

Table 17 Breakdown of elapsed times of an application (IT=1) of a filter (in sec, D-P calculation).

フィルタ	n	経過時間全体	行列分解	フィルタ処理	B の乗算	前進後退代入
Ex1-1	27	103.8	19.3	78.7	(28.0,	50.5)
Ex1-2	28	106.2	19.7	81.4	(28.8,	52.3)
Ex1-3	33	120.7	19.6	95.3	(33.5,	61.6)
Ex1-4	25	98.1	19.6	72.7	(25.9,	46.6)
Ex1-5	109	332.6	19.6	307.3	(104.4,	202.1)
Ex2-1	23	155.3	39.4	110.2	(24.1,	85.9)
Ex2-2	38	225.7	39.3	180.6	(38.1,	142.1)
Ex2-3	38	223.9	38.1	180.7	(38.4,	141.9)
Ex2-4	40	234.2	38.7	189.8	(40.2,	149.1)
Ex3-I-1	20	199.2	58.5	135.0	(21.2,	112.3)
Ex3-I-2	30	265.2	58.7	200.7	(30.7,	167.9)
Ex3-I-3	30	265.1	59.2	200.9	(30.6,	168.1)
Ex3-I-4	40	331.6	58.9	267.0	(39.9,	224.3)
Ex3-I-5	20	200.8	60.0	135.1	(21.4,	112.2)
Ex3-I-6	30	267.0	60.2	201.1	(30.6,	168.4)
Ex3-II-1	30	264.4	58.2	201.2	(30.8,	168.2)
Ex3-II-2	34	291.3	57.7	227.8	(34.5,	190.9)
Ex3-II-3	41	338.0	59.5	273.4	(41.0,	229.5)
Ex4-I-1	23	282.9	80.4	196.7	(24.3,	172.1)
Ex4-I-2	40	425.2	79.8	339.7	(40.0,	299.2)
Ex4-I-3	50	509.5	79.5	424.3	(49.6,	374.1)
Ex4-II-1	63	617.4	78.6	533.0	(61.7,	470.5)

かかる計算量のオーダーは $O(Nw_L^2)$ である、ただし N は固有値問題の行列の次数)。表中で「フィルタ処理」とあるのは、 K 個のレゾルベントの線形結合の作用の n 次の Chebyshev 多項式を 3 項漸化式を利用して計算するために消費した経過時間である。またさらにその内訳として、各レゾルベントの作用を実現するためにベクトルの組に行列 B を乗じたものを作ってそれを各レゾルベントに対応する連立 1 次方程式の共通する右辺ベクトルの組とする操作である「 B の乗算」(全部で n 回) と、あらかじめ分解して保持しておいた行列因子を用いることで $nv=110$ 個の右辺ベクトルの組について各シフト行列を係数とする連立 1 次方程式を解く「前進後退代入」(全部で nK 回の前進後退代入) の経過時間のそれぞれの総計が示してある (すでに Cholesky 分解済みの行列因子を用いて nv 個の右辺ベクトルに対して前進後退代入で連立 1 次方程式を解くための計算量のオーダーは $nv \times O(Nw_L)$ である)。

結局、 n 次の Chebyshev 多項式と K 個のレゾルベントの線形結合を合成したフィルタ \mathcal{F} を nv 個のベクトルの組に対して 1 回だけ適用するための演算量は、「行列分解」が全部で $K \times O(Nw_L^2)$, 「前進後退代入」が全部で $K \times n \times nv \times O(Nw_L)$, 「行列 B の乗算」(B が帯内が疎であることを利

用しなければ)が全部で $n \times nv \times O(Nw_L)$ となる。ただし、どのように演算処理を並列に実施可能であるかや、行列分解や前進後退代入においてブロック化などにより記憶参照の局所性をどれだけ高めて計算できるか、あるいは特殊なハードウェアの機能の利用などにより、演算量の経過時間への寄与の程度は大きく変化しうる。

4.5 実験結果からの考察

有理関数 $y(t)$ の極に近接がある場合には、それらの極の係数の絶対値も大きくなり、そのことからレゾルベントと恒等演算子の線形結合をベクトルの組に作用させた結果のベクトルの組を求める際には、数値相殺(桁落ち)による丸め誤差の拡大により計算結果の相対精度が失われてしまう。それにより、フィルタを用いて得られる近似固有対の精度が低下することになる。このレゾルベントと恒等演算子の線形結合を適用する際の桁落ちの影響による計算結果の数値の精度の低下は、得られた近似固有対を改良しようとして同じフィルタを反復適用しても改善されない。

今回のフィルタのうちで桁落ちによる精度低下の効果が最も著しかったものとしてフィルタ Ex3-I-6 (3.3.6 項)があり、その極の係数 α_i は 5 桁の数である。そのフィルタ Ex3-I-6 を用いて得られた近似固有対の相対残差の大きさのグラフは図 48 である。フィルタの反復による近似固有対の相対残差の大きさは 2 回目と 3 回目がほぼ同じで収束しており改良はこれ以上進まないと思わせるが、その収束したときの相対残差の大きさは固有値の下端付近で最も大きくて $1E-8$ 程度、そうして $\lambda = 30$ 付近で最も小さくて $1E-9$ 程度である。

これに比べて、桁落ちの影響が少ないと考えられるフィルタ Ex2-2 (3.2.2 項) を例にとると、そのフィルタを用いて得られた近似固有対の相対残差の大きさのグラフは図 24 であり、収束したと見なせるときの相対残差の大きさは固有値の下端付近で最も大きくて $1E-11$ 程度、 $\lambda = 30$ 付近で最も小さく $1E-12$ 程度である。これらの相対残差の大きさはフィルタ Ex3-I-6 の場合に比べて約 3 桁小さい。

以上のことから、レゾルベントの線形結合の作用の計算において、必要な固有ベクトルに対する数値相殺による丸め誤差の拡大傾向を抑えて計算結果の相対精度をなるべく保つためには、有理関数の極が近接することを避けることで極の係数の絶対値があまり大きくならないようにする配慮が必要であろう。

なお、フィルタを反復する際に B -正規直交化を間に必ず挟む場合は、 g_p の値を大きくして 1 に近づける(つまり通過域での伝達率の均一性を高める)よりも、なるべく g_s の値を充分小さくしたり、Chebyshev 多項式の次数 n を下げの方が計算の効率を良くできるであろう。

フィルタを反復しないで 1 回だけ使用する場合には、単

一のレゾルベントからなるものを用いるよりも複数のレゾルベントからなるフィルタを用いる方が近似固有対の相対残差の大きさを小さくできる(たとえば今回のフィルタでは表 15 の中の Ex2-3, Ex3-I-3, Ex3-II-1, Ex4-I-1 の結果がそのことを示している)。ただし複数のレゾルベントを使用すると、それだけ多くの記憶容量と行列分解に掛かる演算量が必要になる。レゾルベントの作用を行列分解や前進後退代入により実現する計算処理をレゾルベントごとに複数のノードに分散させて行うのであれば、レゾルベントの数が増えてもノードあたりの記憶量や演算量を一定に保てる。それにより、計算を複数のレゾルベントから構成されたフィルタで行う方が単一のレゾルベントから構成されたもので行うよりも経過時間が短縮する可能性が出てくる。しかしもしもレゾルベントの作用を行列分解と前進後退代入で実現する処理を複数ノードに分散できて、しかもそれらの多くの部分が並列に計算できるのであれば、資源の総量が制限されているのであれば単一のレゾルベントからなるフィルタを用いて分散計算を行う方が経過時間も少なく済むであろう。そうして近似固有対の精度が足りない(相対残差の大きさが十分小さくない)場合にはフィルタを反復することで精度を改良する方が良いと思われる。

5. まとめ

実対称定値一般固有値問題の固有値が下端付近にある少数の固有対の近似を求めるのに、フィルタを利用して必要な固有値を持つ固有ベクトルの全体が張る不変部分空間の基底の近似を構成する方法がある。

本論文で扱ったフィルタは、Chebyshev 多項式を少数のレゾルベントと恒等演算子の線形結合の作用に合成したものであり、さらに問題を固有値が固有値分布の下端付近にある少数の固有対だけを求める場合に限定することで、レゾルベントのシフトをすべて最小固有値未満の実数にして虚数のシフトの使用を避けることを試みた。そうしてフィルタの伝達特性を代表する 3 つの形状パラメータ μ , g_s , g_p の値が指定されたときに、それを満たすフィルタを構成する方法を与えた。

レゾルベントの個数がいくつでも、有理最良近似の理論により楕円有理関数の性質を用いることで、レゾルベントと恒等演算子の線形結合の最適なシフトの値や線形結合の係数は、解析的な表式に数値を入れて計算すれば決まる [7]。しかしその方法では複数のレゾルベントを用いる場合にはレゾルベントのシフトに虚数のものが必ず現れるので、今回のシフトを実数に制限する場合には適用ができない。

我々は以前の論文 [8] で実数シフトのレゾルベントが 1 つと 2 つの場合については、連立代数方程式の実数解を代数的な操作で求めることにより、指定された条件を満たす伝達関数が存在する場合にはそれを具体的に構成できる手順を得ていた。しかし実数シフトのレゾルベントが 3 つ以

上の場合については、実数解を代数的な操作で求める方法は構成できていない。

そこで我々は本論文で、実数シフトのレゾルベントが2つから4つまでの場合について、連立代数方程式の実数解を代数的な操作で求めるのではなく、Remez法と同様の反復法と人の手による試行錯誤により求める方法を用いた。その方法により、フィルタのパラメタとして指定した値の組を満たすような連立代数方程式の実数解をうまく数値で得た場合には、その解に基づいてフィルタを構成できる。そのようにして得られたフィルタを利用して一般固有値問題の下端付近の近似固有対を実際に求めて、それらの相対残差の大きさを求めた。

ただし、指定した3つの形状パラメタ μ , g_s , g_p の値と Chebyshev の次数 n の組に対して連立代数方程式を満たす実数解が求まらない場合に、探索が不十分であるからなのか、それともそのような実数解が存在しないからなのかを区別できていない。また今後は実数解の探索は人手による試行錯誤になるべく依らずに自動化することが望ましい。

また、数値相殺による必要な固有ベクトルの精度低下をなるべく抑えるため、レゾルベントの線形結合の係数の大きさがあまり大きくならないよう制御することが望ましい。

参考文献

- [1] Polizzi, E.: Density-Matrix-Based Algorithm for Solving Eigenvalue Problems, *Phys. Rev. B*, Vol.79, No.11, pp.115112-115118 (2009).
- [2] Pachón, R.: Chap.6.3: Location of Extrema: From a trial function to a reference, *Algorithms for Polynomial and Rational Approximation*, Thesis of Oxford University (2010).
- [3] Guttel, S., Polizzi, E., Tang, P.T.P. and Viaud, G.: Zolotarev Quadrature Rules and Load Balancing for the FEAST Eigensolver, *SIAM J. Sci. Comput.*, Vol.37, No.4, pp.A2100-A2122 (2015).
- [4] 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).
- [5] Murakami, H.: Filters Consist of a Few Resolvents to Solve Real Symmetric-Definite Generalized Eigenproblems, *JJIAM*, Vol.36, No.2, pp.579-618 (2019).
- [6] 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).
- [7] 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 (ACS), Vol.13, No.1 (ACS67), pp.1-27 (2020).
- [8] 村上 弘:「実数シフトのレゾルベント少数から構成されたフィルタによる実対称定値一般固有値問題の下端側固有値を持つ固有対の解法について」, 情報処理学会研究報告: ハイパフォーマンスコンピューティング (HPC), Vol.2020-HPC-175, No.1, pp.1-26 (2020).
- [9] Murakami, H.: Single-Precision Calculation of Iterative Refinement of Eigenpairs of a Real Symmetric-Definite

Generalized Eigenproblem by Using a Filter Composed of a Single Resolvent, *Proc. HPC Asia 2021: The international Conference on High Performance Computing in Asia-Pacific Region Companion*, pp.11-20 (2021).



村上 弘 (正会員)

1960年生。1992年北海道大学大学院理学博士号(化学第二学専攻)取得。現在、東京都立大学理学研究科数理学専攻の准教授。数理的な問題の数値的あるいは記号的方法による効率あるいは精度の良い解法やその並列化手

法の研究に従事。日本応用数理学会、日本数式処理学会、SIAM、ACM各会員。