

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

村上 弘^{1,a)}

受付日 2019年7月12日, 採録日 2019年11月26日

概要: 実対称定値一般固有値問題に対して, 固有値が指定された区間にある固有対すべての近似を得るためのフィルタとして, その区間に対応する不変部分空間への射影を近似する線形作用素を構成して用いる. 本研究で扱うフィルタは少数 2~4 個の複素シフトを持つレゾルベントの線形結合 (の実部) の多項式である. その構成方法は, まず単一のレゾルベントの多項式を元のフィルタとして, その伝達関数を表す有理関数に, うまく選んだ低次の有理関数を合成して, 改善された伝達特性を持つ新しい有理関数を得て, その新しい有理関数を伝達関数として持つ線形作用素が複素数をシフトとする少数のレゾルベントの線形結合 (の実部) の多項式により表せることを用いる. 合成用の有理関数の構成法は, アナログ電気回路の 4 種類の典型フィルタであるバターワース型, チェビシェフ型, 逆チェビシェフ型, 楕円型のを模倣する. これまでの研究で我々はすでに, 最初の 3 種類に対応するものについて合成用の有理関数とフィルタの構成法を具体的に示し, さらにそれらのフィルタを用いた実験により実対称定値一般固有値問題の近似固有対がうまく求められることを示したが, 最後の楕円型に対応するものについてはまだ扱っていなかった. そこで今回の論文で我々は, 追加として楕円型を模倣した合成用の有理関数とフィルタの構成法を示し, さらにそのフィルタを用いた実験により実対称定値一般固有値問題の近似固有対がうまく求められることを示す.

キーワード: フィルタ対角化, 固有値問題, レゾルベント, 多項式, 楕円型

Solution of Real Symmetric Definite Generalized Eigenproblems by Using a Filter Composed of a Small Number of Resolvents

HIROSHI MURAKAMI^{1,a)}

Received: July 12, 2019, Accepted: November 26, 2019

Abstract: For a real symmetric-definite generalized eigenproblem to obtain all those approximate eigenpairs whose eigenvalues are in a specified interval, we construct a linear operator as the filter which is an approximate projector to the invariant subspace corresponds to the interval. In this study, we use filters which are polynomials (of the real part) of a linear combination of a small number (2 to 4) of resolvents whose shifts are complex numbers. To construct the filter, we first consider an original filter which is a polynomial of a single resolvent. And to the transfer function of the original filter, which is a rational function, we compose a well chosen low degree rational function to make a new rational function as a transfer function whose property is improved. Then, from the improved transfer function, we can construct a linear operator which is a polynomial of (the real part of) a linear combination of a few resolvents whose shifts are complex numbers. We select the low degree rational function for the composition, by analogies to those four typical types of filters in the theory of electronic circuits namely *Butterworth*, *Chebyshev*, *inverse-Chebyshev* and *elliptic*. In previous studies, we have already shown those rational function compositions which are analogies of the first three types, and also shown by experiments that approximate eigenpairs of real symmetric-definite generalized eigenproblems can be solved well by uses of corresponding filters. But we have not yet shown the analogy of the last one, the elliptic type. Thus, in the present paper, we show the rational function composition which is the analogy of the elliptic type as a supplement, and show by experiments that approximate eigenpairs of real symmetric-definite generalized eigenproblems can be solved well by the corresponding filter.

Keywords: filter diagonalization, eigenproblem, resolvent, polynomial, elliptic

1. はじめに

本論文の内容の多くは以前の我々の文献 [14] に出ている。しかしその時点ではアナログ電気回路の(伝達率に着目した)4種類の典型フィルタの類似のうちの3つまでしか記述ができていなかった。我々は、残されていた第4番目の楕円フィルタの類似で構成が数学的に最も高度で潜在的な性能が高いもの(E-合成,楕円型合成)を今回の本論文で新たに追加し記述する(6章)。なるべく4種類を同等に統一して扱うために,すでに文献[14]に記述されたものと共通・類似する内容を含むが,記号や記法を(追加部分との折り合いをつけるために)変更した箇所が多い。

行列 A と B が実対称で B は正定値である一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対で,固有値 λ が指定された区間 $[a, b]$ にあるものをフィルタ対角化法 [5] により近似して解く技法を扱う。フィルタとして用いる線形作用素 \mathcal{F} をうまく構成して,固有値が区間 $[a, b]$ の近傍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ている固有ベクトルは強く減衰させるようにする。

1.1 レゾルベントの線形結合型のフィルタ

フィルタの構成としては,複素関数論の Cauchy の積分定理から導かれる複素平面内の閉曲線に沿った線積分の離散化近似に相当するものを用いる方法がある(文献 [6], [7], [8], [9], [11], [13])。これはフィルタの作用を曲線上の積分の離散化点に対応するシフト ρ_j を持つ複数のレゾルベント $\mathcal{R}(\rho_j) \equiv (A - \rho_j B)^{-1}B$ の線形結合として実現している。

別のアプローチには,固有値 λ の固有ベクトル \mathbf{v} に対する伝達関数 $f(\lambda)$ の特性が,なるべく望ましいものとなるように構成された線形作用素 \mathcal{F} をフィルタとして用いる方法がある。いまフィルタの構成をレゾルベント(と恒等作用素)の線形結合 $\mathcal{F} = c_\infty I + \sum_j \gamma_j \mathcal{R}(\rho_j)$ (あるいはその実部)に限ると,その伝達関数 $f(\lambda)$ は有理関数で,無限遠で有界であり重複する極を持たないものになる。そこでそのような性質を持つ有理関数であって理想的なフィルタの伝達特性を近似するものを定めてやり(文献 [2], [3]),それから伝達関数 $f(\lambda)$ を持つフィルタ \mathcal{F} を逆に構成すれば,各レゾルベントのシフト ρ_j と線形結合の係数 c_j が決まる。このようにして固有値が区間 $[a, b]$ にある固有ベクトルを抽出するのに適した特性を持つフィルタを設計できる(文献 [12], [15], [16], [19])。たとえば構成を「楕円フィルタ」にすれば,8個から16個程度の複素数のシフトを持つレゾルベントの線形結合の実部として特性の十分に良い

フィルタが得られ,しかも求めたい固有対の固有値を含む区間を固有値分布の任意の位置に設定できるので,中間固有対用である。

また,シフトが実数のレゾルベントの線形結合でフィルタを構成することにより,実数だけを用いて計算を行う方法も一応考えられる(文献 [17], [18], [20])。しかしそのようなフィルタは,求めたい固有対の固有値が固有値分布の端にある場合にだけ安全に適用ができるものであり,また得られるフィルタの伝達特性もシフトとして虚数の使用を許す場合に比べると良くない。

上記の方法はレゾルベントのシフトとして実数だけを用いるかそうでないかの違いはあるものの,どちらの場合もフィルタはレゾルベントの線形結合であり,レゾルベントを複数用いることになる。

1.2 レゾルベントと連立1次方程式

与えられたベクトルの組 V に対するレゾルベント $\mathcal{R}(\rho)$ の作用は,係数行列が $C(\rho) = A - \rho B$ で右辺ベクトルの組が BV の連立1次方程式を解いて実現する(行列 A と B が実対称であることから,シフト行列 $C(\rho)$ はシフト ρ が実数ならば実対称で, ρ が虚数ならば複素対称である。 $C(\rho)$ の非零要素の分布は A と B の非零要素の分布の合併になるので, A と B がともに疎であれば $C(\rho)$ も疎行列になり, A と B がともに帯行列であれば $C(\rho)$ も帯行列になる)。いま係数が $C(\rho)$ の連立1次方程式を直接法で(行列分解を利用して)解くことにすると,フィルタの作用の実現には少なくともフィルタの構成に用いるレゾルベントと同数の行列分解が必要である(連立1次方程式を反復法で解く場合にも,行列の不完全分解を前処理に用いるならば,フィルタの構成に用いるレゾルベントと同数の行列の不完全分解が必要である)。

シフト ρ が実数で,それが固有値分布の下限あるいは上限であれば,シフト行列 $C(\rho)$ は実対称定値になり, $C(\rho)$ を係数とする連立1次方程式はピボット交換なしの改訂コレスキ法(modified Cholesky method)を用いて安定に解ける。さらに $C(\rho)$ が帯行列であれば帯用に特化した算法を用いることができる。

シフト ρ が虚数のときには,行列 $C(\rho)$ は正則で複素対称($C(\rho)^T = C(\rho)$)になるので,係数が $C(\rho)$ の連立1次方程式は複素対称版の改訂コレスキ法を用いて解く。もしもその際に,ピボット交換をしない分解の計算では数値的な不安定性が生じたなら,行列 $C(\rho)$ の複素対称性を利用することは止めて,連立1次方程式を片側ピボット交換付きの複素版の LU 分解を用いて安定に解くことができる(ただし必要な記憶量と計算量がかなり増えてしまう)。さらに $C(\rho)$ が帯行列である場合には,改訂コレスキ法あるいは片側ピボット交換付きの LU 分解には,それぞれ帯用に特殊化したものを利用できる。

¹ 首都大学東京・数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

a) mrkmhrsh@tmu.ac.jp

1.3 単一のレゾルベントの多項式型フィルタ

行列分解の数を最少にするために、単一のレゾルベントを用いて、レゾルベントのシフトが実数の場合にはその作用の多項式で、シフトが複素数の場合にはその虚部の作用の多項式で、フィルタを構成する方法を我々は試みてきた [21].

レゾルベントの作用を与える連立1次方程式を直接法を用いて解くと仮定する場合には、レゾルベントを複素用いる方法に比べて1つだけ用いる方法では行列分解を行う数が複素ではなくて1つになるので行列分解にかかる演算量を減らせる利点がある。

いまフィルタ \mathcal{F} を単一のレゾルベント $\mathcal{R}(\rho)$ から作られる実作用素の実多項式とすれば、その伝達関数 $f(\lambda)$ は実有理関数になる。そこで、必要な固有値を持つ固有ベクトルの抽出に適したフィルタの伝達関数を表すように、実有理関数 $f(\lambda)$ をうまく調整する。そうしてその実有理関数 $f(\lambda)$ を伝達関数を持つフィルタ \mathcal{F} を構成する。

本論文で扱う実対称定値一般固有値問題の固有値はすべて実数なので、求めたい固有値の区間 $[a, b]$ が固有値分布の下端（上端）にある場合には、レゾルベントのシフトを実数に制限しても、そのシフトの値を最小固有値よりも小さく（あるいは最大固有値よりも大きく）設定すれば、シフトと固有値が一致したり極端な近接をすることはないことが保証される。シフトに実数を用いた場合にはフィルタ対角化法の計算は実数とその演算だけを用いて構成できる。

求めたい固有対の固有値の区間が固有値分布に対して一般的な位置にある場合には、シフトに虚数を選べば固有値との一致や極端な近接は生じない。互いに複素共役なシフトを持つレゾルベントを実ベクトルに作用させた結果の2つの複素ベクトルは互いに複素共役なものになることを用いると、実作用素であるフィルタはシフトが互いに複素共役であるレゾルベントの片方（たとえばシフトの虚部が正であるもの）だけを用いて構成できる。この場合にも必要なレゾルベントは1つだけとなり、レゾルベントに対応する連立1次方程式を直接法で解く場合に必要係数行列の分解は1つになる。

しかし単一のレゾルベントの多項式として構成されたフィルタには、（多項式を Chebyshev 多項式で表す簡易型の場合は特に）難点がある。それは通過域における伝達関数の最大最小比を小さく抑えることが難しいことである（シフトが実数の場合の方が虚数の場合よりも難しい）。この最大最小比が大きければ（たとえば3~6桁）、近似固有ベクトルの精度の不均一さも同程度になる可能性がある。この最大最小比は遷移域を広くとれば小さくできるが、遷移域の広いフィルタ特性はあまり望ましいものではない。なぜならば遷移域が広いとそこに入る不要な固有値の数も増える傾向があり、そのような不要な固有値が増えるとそれに対応してフィルタを適用した結果のベクトルに含まれ

る不要な固有ベクトルの数も増えるので、フィルタを適用するベクトルの数を多くすることが必要になり、作業の効率が悪くなるからである。

1.4 関数合成による特性の改良

本論文では、特性が改良されたフィルタを得るために、元にする単一のレゾルベントの多項式であるフィルタの伝達関数に対してうまく選んだ有理関数を合成することにより、遷移域の幅が縮小した特性を持つ伝達関数を表す有理関数を作る。そうしてその合成で得られた有理関数を伝達関数とするフィルタを、少数のレゾルベントの線形結合の多項式で実現するという方法をとる。

しかしレゾルベントの作用を与える連立1次方程式を直接法を用いて解く場合には、用いるレゾルベントの数が元の1つから少数（たとえば2~4個）に増えるので、行列分解にかかる演算量と行列分解を保持するための記憶量も（元の2~4倍に）増える。この不利益はフィルタの伝達特性を改良することで得られる利益との交換になる。

伝達率の特性に着目したアナログ電気回路の典型フィルタは、バターワース型 (Butterworth)、チェビシェフ型 (Chebyshev)、逆チェビシェフ型 (inverse Chebyshev)、楕円型 (elliptic) の4種類である。本論文で扱う4種類の有理関数による関数合成の方法は、それらのアナログ回路の場合の構成手法を模倣するものであり、それぞれ B-合成、C-合成、I-合成、E-合成と呼ぶことにする。

注記

すでに文献 [14] において、我々は上記4種類のうちの最初の3種類 (B, C, I) の合成については導出を行い、さらにそれらから導かれるフィルタを用いて一般固有値問題の固有対を求める数値実験を行ってそれらの動作を確認しているが、E-合成についての記述は含まれていない。

その後の文献 [22], [24] において、我々は E-合成の導出を示した（なお文献 [22] の中で示した E-合成の場合の有理関数 $h(t)$ の式は、文献 [24] や本論文で用いている式 (59) とは異なる形をしているが、有理楕円関数についての恒等関係式 (55) を用いて式を変形すれば両者は一致する）。

さらにその後の文献 [25] の中で、我々は E-合成も含めた4種類の合成 (B, C, I, E) から導かれたフィルタを用いて、一般固有値問題の固有対を求める数値実験の例も示した。ただしこの文献 [25] からは C-合成の場合の有理関数 $h(t)$ の定義の方法が文献 [14] およびそれ以前のものとは異なっている（以前の文献では、C-合成の場合には有理関数 $h(t)$ の定義をその次数 l の偶奇に従って切り替えていたが、文献 [25] からは l の偶奇で扱いを変える箇所を他の場所に移して、 $h(t)$ の定義をその次数 l の偶奇にはよらないものにした）。なお本論文は文献 [25] とほぼ同じ内容である。

2. 単一のレゾルベントの多項式によるフィルタ

実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対 (λ, \mathbf{v}) の固有値 λ はすべて実数である。また固有ベクトル \mathbf{v} もすべて実ベクトルに選べるのでそのように選ぶことにする。さらに固有ベクトルの全体は全空間を張る B -正規直交系にとれる。この固有値問題に対するシフト ρ のレゾルベント \mathcal{R} を式 (1) で定義する。

$$\mathcal{R}(\rho) \equiv (A - \rho B)^{-1} B. \quad (1)$$

そのとき固有対 (λ, \mathbf{v}) に対しては式 (2) が成り立つ。

$$\mathcal{R}(\rho)\mathbf{v} = \frac{1}{\lambda - \rho} \mathbf{v}. \quad (2)$$

すると $Q(x)$ が x の多項式であれば、式 (3) が成り立つ。

$$Q(\mathcal{R}(\rho))\mathbf{v} = Q\left(\frac{1}{\lambda - \rho}\right)\mathbf{v}. \quad (3)$$

いまフィルタ \mathcal{F} がレゾルベント $\mathcal{R}(\rho)$ の多項式として式 (4) の形に書けたとする。

$$\mathcal{F} = Q(\mathcal{R}(\rho)). \quad (4)$$

そのとき固有対 (λ, \mathbf{v}) のベクトル \mathbf{v} に対するフィルタ \mathcal{F} の作用は式 (5) で与えられるので、 $f(\lambda)$ はフィルタ \mathcal{F} の伝達関数である。

$$\mathcal{F}\mathbf{v} = f(\lambda)\mathbf{v}, \quad f(\lambda) = Q\left(\frac{1}{\lambda - \rho}\right). \quad (5)$$

いま座標 λ に対する正規化座標 t を、指定した固有値の区間 $\lambda \in [a, b]$ から標準区間 $t \in [0, 1]$ への 1 次変換 $t = (\lambda - a)/(b - a)$ により定義する。そうして式 (6) で表される 2 つのパラメータ $\mu > 1$ と $\sigma > 0$ を持つ非負領域 $t \in [0, \infty)$ を定義域とする有理関数 $x(t)$ を導入する。

$$x(t) \equiv \frac{\mu + \sigma}{t + \sigma}. \quad (6)$$

いまフィルタ \mathcal{F} の伝達関数 $f(\lambda)$ を正規化座標 t で表した有理関数 $g(t) = f(\lambda)$ が n 次多項式 P を用いて $g(t) = P(x(t))$ と書けたら、 λ で表した伝達関数は $f(\lambda) = P(\gamma/(\lambda - \rho))$ であり、フィルタは $\mathcal{F} = P(\gamma\mathcal{R}(\rho))$ となる。ただし $\gamma \equiv (b - a)(\mu + \sigma)$ 、 $\rho \equiv a - (b - a)\sigma$ である。

この場合のレゾルベントのシフト ρ は a 未満の実数になる。シフト ρ がどの固有値にも一致や極端に近接しないことを保証するために、区間の下端 a は最小固有値 λ_{\min} 以下とする。つまり a は固有値分布の 1 つの下界であり、区間 $[a, b]$ は最小固有値を含むように設定されていると仮定する。そうして $x(t)$ が持つ 2 つのパラメータ σ と μ 、および n 次多項式 P をうまく調整することでなるべく伝達関数 $g(t) = P(x(t))$ の特性を都合の良いものにする。つまり $g(t)$ に対応して構成されたフィルタ \mathcal{F} は固有値 λ が区間 $[a, b]$ にある固有ベクトルは良く伝達するが、固有値が区間から離れた固有ベクトルは強く阻止するものにする。

3. 伝達関数 $g(t)$ と有理関数 $h(t)$ の合成

いま正規化座標 t で表したフィルタ \mathcal{F} の伝達関数 $g(t)$ が ($t = \infty$ も含めた) 非負の実数全体の区間 $[0, \infty]$ で定義された有界な実有理関数であるとし、その定義域を $[0, 1]$ 、 $(1, \mu)$ 、 $[\mu, \infty]$ (ただし $\mu > 1$) の 3 つに分割してそれぞれ通過域 (pass-band)、遷移域 (transition-band)、阻止域 (stop-band) とする。そうして $g(t)$ の通過域における最大値は 1 に規格化し、 $g(t)$ の通過域での最小値を g_p とし、さらに $|g(t)|$ の阻止域における上限値を g_s とする。また g_p と g_s の 2 つを伝達率の閾値と呼ぶことにする。通過域における伝達関数の最大最小比は $1/g_p$ である。

3.1 有理関数 $h(t)$ の合成

関数の合成に用いるための実有理関数 $h(t)$ を導入する。 $h(t)$ の有理関数としての次数を ℓ 、ただし $\ell \geq 2$ とする。そうして $h(t)$ は ($t = \infty$ も含めた) 区間 $t \in [0, \infty]$ において、以下の 3 つの性質を満たすものとする。ただし $1 < \xi$ であるとして、関数の値として $\pm\infty$ も含めて考える。

- (1) 区間 $[0, 1]$ から $[0, 1]$ への全射。
- (2) 区間 $(1, \xi)$ から $(1, \mu)$ への全射で、単調増加。
- (3) 区間 $[\xi, \infty]$ から $[\mu, \infty]$ への全射。

(上記の条件 (2) から $h(1) = 1$ 、 $h(\xi) = \mu$ が出る。)

このとき $g(t)$ と $h(t)$ の合成関数 $\hat{g}(t) = g(h(t))$ を定義すると、 $\hat{g}(t)$ も実有理関数であり、 $\hat{g}(t)$ の定義域である $[0, \infty]$ の通過域、遷移域、阻止域への分割をそれぞれ $[0, 1]$ 、 $(1, \xi)$ 、 $[\xi, \infty]$ とすれば、 $g(t)$ と $\hat{g}(t)$ の伝達率の 2 つの閾値はそれぞれ一致する、つまり通過域における伝達率の最小値はどちらも g_p で、阻止域における伝達率の大きさの上限値もどちらも g_s である。ただし遷移域の幅を与えるパラメータはそれぞれ μ と ξ であって異なる。つまり伝達関数の形状パラメータの 3 つ組は、合成前の伝達関数 $g(t)$ では (g_p, g_s, μ) であるが、合成後の伝達関数 $\hat{g}(t) = g(h(t))$ では (g_p, g_s, ξ) になる。関係 $h(\xi) = \mu$ を満たす ξ が μ よりも 1 に近ければ合成後の伝達関数 $\hat{g}(t)$ の遷移域の幅 $\xi - 1$ は合成前の伝達関数 $g(t)$ の遷移域の幅 $\mu - 1$ よりも狭くなる。

いま $h(t)$ が偶関数であるときには、合成後の伝達関数 $\hat{g}(t)$ も偶関数になるので $\hat{g}(t)$ の定義域を自然に実数全体に広げることができて、 $\hat{g}(t)$ の通過域、遷移域、阻止域も原点 $t = 0$ について対称に広げることができて、それぞれ通過域 $[-1, 1]$ 、遷移域 $1 < |t| < \xi$ 、阻止域 $\xi \leq |t|$ となる。また $h(t)$ が偶関数でなくても、 $h(t)$ の $[-1, 0]$ における値域が $[0, 1]$ に対する値域に含まれる場合には、 $\hat{g}(t)$ の通過域を $[-1, 1]$ に広げることができる。

3.1.1 4 種類の各合成用の有理関数 $h(t)$

本論文ではアナログ電気回路の典型フィルタの 4 種類としてよく知られているバターワース (Butterworth) 型、チェビシェフ (Chebyshev) 型、逆チェビシェフ (inverse-

Chebyshev) 型, 楕円 (elliptic) 型の各構成法 (文献 [2], [3]) の模倣から得られる 4 種類の合成 (B, C, I, E) を利用してフィルタを構成する. それらの合成に用いる l 次 (ただしつねに $l \geq 2$ とする) の実有理関数 $h(t)$ は, それぞれ式 (7) で与えられるものである.

$$h(t) = \begin{cases} t^\ell & \text{(B-合成)}, \\ \frac{1 + T_\ell(t)}{2} & \text{(C-合成)}, \\ \frac{1 + T_\ell(\xi)}{1 + T_\ell(\xi/t)} & \text{(I-合成)}, \\ \frac{L+1}{2} \times \frac{1 + R_\ell(\xi, t)}{L + R_\ell(\xi, t)} & \text{(E-合成)}. \end{cases} \quad (7)$$

ここで C-合成と I-合成に現れる $T_\ell(z)$ は変数 z の次数 l の第 1 種 Chebyshev 多項式を表す. また I-合成と E-合成に現れるパラメータ ξ は実数で 1 よりも大きい. E-合成の場合に現れる $R_\ell(\xi, z)$ はパラメータ ξ を持ち変数が z である次数 l の「楕円有理関数」を表し, L は実数で $L = R_\ell(\xi, \xi)$ を満たし 1 より大きい (E-合成で用いる楕円有理関数については 6.1 節に詳しく記述する).

上記の 4 通りの有理関数 $h(t)$ は次数 l が 2 の場合にはどれも同じ多項式 t^2 となり一致する. また有理関数 $h(t)$ は B-合成と C-合成ではつねに多項式であるが, I-合成と E-合成では ($l \geq 3$ の場合には) 多項式ではない.

注記

以前の文献 [14] では, C-合成の場合の $h(t)$ の式を l の偶奇で区別して定義をしていたが, 本論文では各合成用の関数 $h(t)$ の定義は l の偶奇にはよらないものに改める. そうして通過域とする固有値の区間 $\lambda \in [a, b]$ に対応する規格化座標 t で表した区間をたとえば $t \in [-1, 1]$ にするか $t \in [0, 1]$ にするかは, レゾルベントの線形結合である式 (15) のシフト ρ_i と結合係数 γ_i を有理関数 $\hat{x}(t)$ の部分分数分解 (14) の極 t_i とその係数 c_i に対応させる線形変換の式を適切に与えることにより行うようにした.

3.1.2 4 種類の各合成における μ と ξ の対応

4 種類の合成 (B, C, I, E) の各場合に, ξ の値から μ の値は式 (8) を用いて計算で求めることができる (C-合成と I-合成の場合の式は一致する). ただし式 (8) で E-合成の場合は $L \equiv 1/q^{-1}(\{q(1/\xi)\}^\ell)$ である. ここで $q(k)$ は楕円 nome 関数 $q(k) \equiv \exp\{-\pi K'(k)/K(k)\}$ を表し, $q^{-1}(x)$ はその逆関数を表す. また $K(k)$ は母数 k の第 1 種の完全楕円積分であり, $K'(k) \equiv K(k')$ であり, $k' \equiv \sqrt{1-k^2}$ は母数 k に対する補母数を表す (これらの関数や楕円積分については 6.1 節に詳しく記述する).

$$\mu = \begin{cases} \xi^\ell & \text{(B-合成)}, \\ \frac{1 + T_\ell(\xi)}{2} & \text{(C-合成と I-合成)}, \\ \frac{(L+1)^2}{4L} & \text{(E-合成)}. \end{cases} \quad (8)$$

あるいは逆に μ の値から ξ の値を式 (9) の計算で求めることができる (これも C-合成と I-合成に対する式は一致する). ただし式 (9) で E-合成の場合は $L \equiv (\sqrt{\mu} + \sqrt{\mu-1})^2$ である.

$$\xi = \begin{cases} \mu^{1/\ell} & \text{(B-合成)}, \\ \cosh((2/\ell) \sinh^{-1} \sqrt{\mu-1}) & \text{(C-合成と I-合成)}, \\ 1/q^{-1}(\{q(1/L)\}^{1/\ell}) & \text{(E-合成)}. \end{cases} \quad (9)$$

3.1.3 各合成による遷移域の幅を縮小する能力の比較

図 1, 図 2, 図 3, 図 4, 図 5, 図 6 までの 6 枚の図はそれぞれ, 合成用の有理関数の次数 l を 3 から 8 まで変えた各場合について 4 種類の合成 (B, C, I, E) の各場合について μ に対する ξ の関係を, 横軸に μ の常用対数の値を, 縦軸に ξ の常用対数の値をとって描いたグラフである. ただし C-合成と I-合成の場合のグラフはつねに同じになるのでまとめて 1 つにしている. なお次数が $l = 2$ の場合には, 4 種類の合成すべてで $h(t) = t^2$ になるので, $\xi = \sqrt{\mu}$ であり, 各合成の場合のグラフがすべて一致するので図は省略した. パラメータ μ と ξ はそれぞれ正規化座標により合成前と合成後の遷移域の端の位置を表したものであり ($\mu - 1$ と $\xi - 1$ がそれぞれ遷移域の幅である), 4 種類の合成はどの場合も, $\log_{10} \xi$ の値が $\log_{10} \mu$ の値よりも小さくて零に近い, つまり ξ の値は μ の値よりも小さくなっていて 1 に近いことを示しており, 伝達関数の遷移域の幅が合成で縮小されることを表している. またこれらのグラフでは, 曲線が下側にあるものから順番に E-合成の場合, C-合成と I-合成の場合, そうして最後が B-合成となっていることから, 4 種類の合成のうちで遷移域の幅を縮小する能力が最も優れているのは E-合成であり, その次が C-合成および I-合成であり, 最後が B-合成であることも分かる.

別の図 7, 図 8, 図 9 の各グラフからも, B-合成, C-合成 (I-合成), E-合成それぞれについて, 次数 l を増すときに遷移域の幅 $\xi - 1$ が縮小する様子を読みとることができる. そうしてこれらの比較から (l が 2 以外では) 遷移域の幅を縮小する能力は E-合成が最も優れていて, その次は C-合成と I-合成であり, 最後が B-合成であることが確認できる.

3.1.4 4 種類の各合成による伝達関数 $\hat{g}(t)$ について

4 種類の合成 (B, C, I, E) の各場合の有理関数 $h(t)$ は, B-合成と I-合成では $h(t)$ の原点付近での振舞いは両方とも $O(t^\ell)$ であることから, 合成された伝達関数 $\hat{g}(t) = g(h(t))$ は, 原点 $t = 0$ において $(l - 1)$ 階までの導関数がすべて零になるので, 原点付近で高度な平坦性 (いわゆるバターワース特性) を持つ.

式 (7) で定義された有理関数 $h(t)$ は, どの合成の場合にも次数 l が偶数の場合には偶関数である. しかし, 次数 l

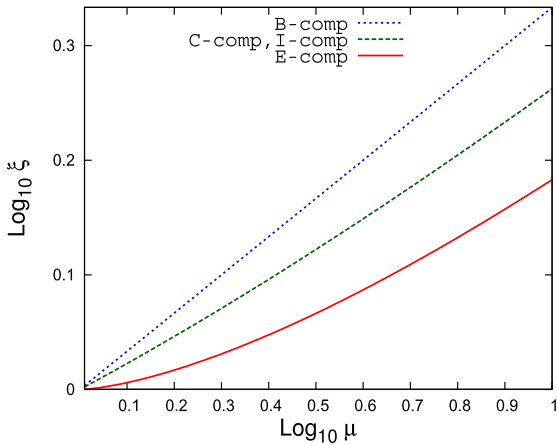


図 1 μ に対する ξ の両対数グラフ (各種合成, $\ell = 3$ の場合)
Fig. 1 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 3$).

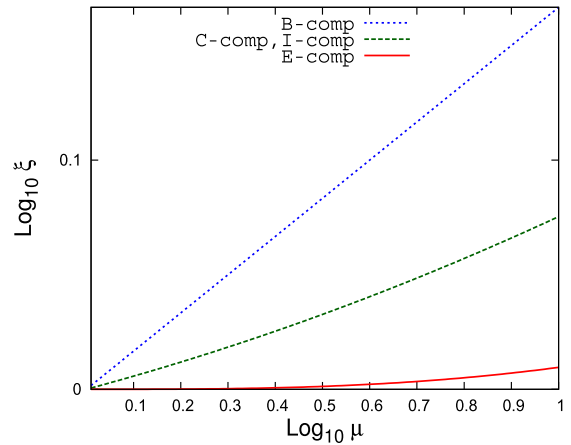


図 4 μ に対する ξ の両対数グラフ (各種合成, $\ell = 6$ の場合)
Fig. 4 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 6$).

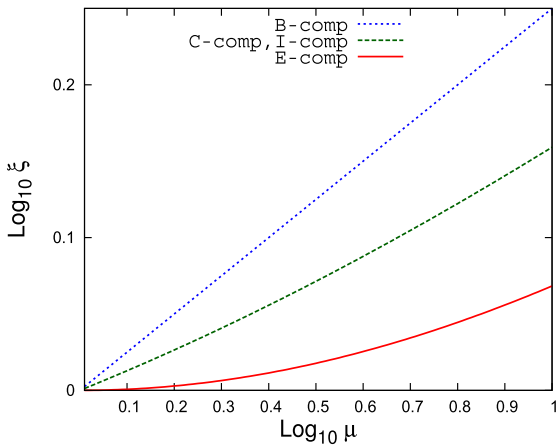


図 2 μ に対する ξ の両対数グラフ (各種合成, $\ell = 4$ の場合)
Fig. 2 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 4$).

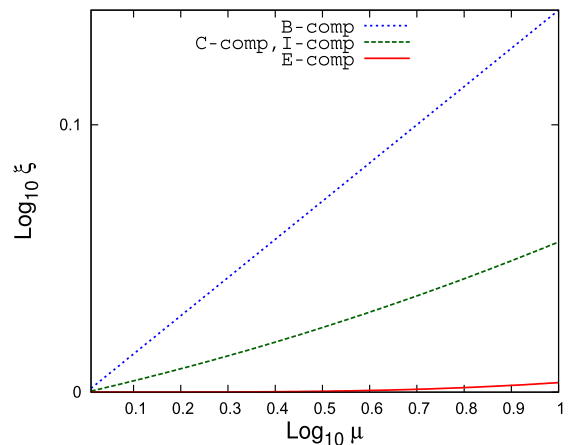


図 5 μ に対する ξ の両対数グラフ (各種合成, $\ell = 7$ の場合)
Fig. 5 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 7$).

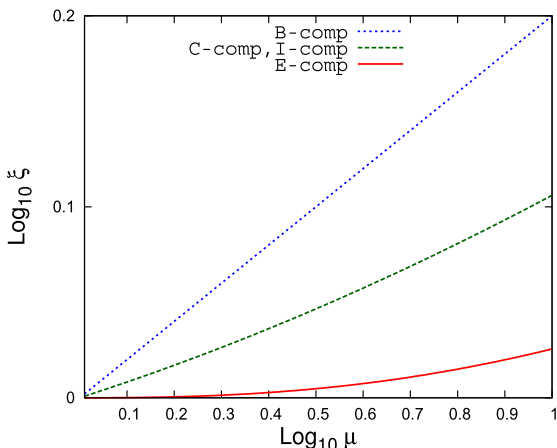


図 3 μ に対する ξ の両対数グラフ (各種合成, $\ell = 5$ の場合)
Fig. 3 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 5$).

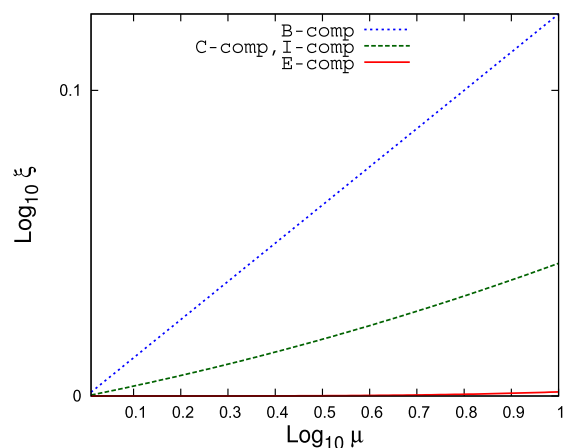


図 6 μ に対する ξ の両対数グラフ (各種合成, $\ell = 8$ の場合)
Fig. 6 Log-log graphs of μ vs. ξ (for each composition, for the case $\ell = 8$).

が奇数である場合には、C-合成とE-合成については $h(t)$ の区間 $[-1, 0]$ における値域は $h(t)$ の区間 $[0, 1]$ における値域に含まれることを示せるが、B-合成とI-合成については

$h(t)$ の区間 $[-1, 0]$ における値域は区間 $[0, 1]$ における値域には含まれずにはみ出すことが分かる。そのため ℓ が奇数である場合のB-合成やI-合成についてはたとえば $h(t)$ が

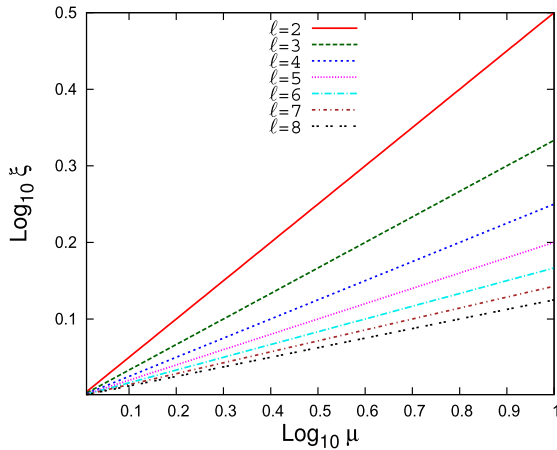


図 7 μ に対する ξ の両対数グラフ (各次数 l , B-合成)
 Fig. 7 Log-log graphs of μ vs. ξ (for each degree l , for B-composition).

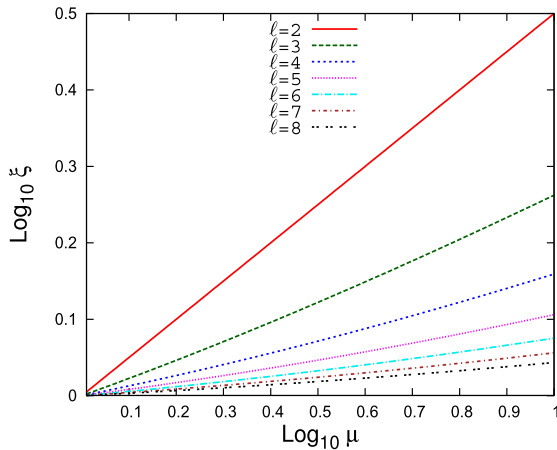


図 8 μ に対する ξ の両対数グラフ (各次数 l , C-合成と I-合成)
 Fig. 8 Log-log graphs of μ vs. ξ (for each degree l , for both C and I-compositions).

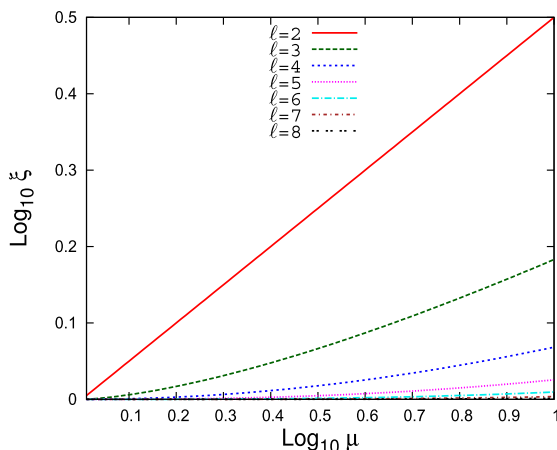


図 9 μ に対する ξ の両対数グラフ (各次数 l , E-合成)
 Fig. 9 Log-log graphs of μ vs. ξ (for each degree l , for E-composition).

区間 $[-1, 1]$ に極を持たない場合に通過域を $t \in [-1, 1]$ に拡張しようとするれば、それらの合成に対する関数 $h(t)$ の定義式 (7) を修正する必要がある。

以上のことから、次数 l が偶数である場合には、4種類の合成のどれについても合成された伝達関数 $\hat{g}(t)$ の通過域、遷移域、阻止域を原点对称に自然に拡張してそれぞれ $t \in [-1, 1]$, $1 < |t| < \xi$, $\xi \leq |t|$ にすることができる。他方で、次数 l が奇数である場合には、B-合成とI-合成では、原点 $t = 0$ から $t = -1$ に向かうときに $\hat{g}(t)$ の値は 1 を超えるので、通過域を自然に拡張して $t \in [-1, 1]$ とすることはできない。そこでたとえば通過域、遷移域、阻止域の各区域はそれぞれ拡張せずに元のままの $[0, 1]$, $(1, \xi)$, $[\xi, \infty)$ とする。他方でC-合成とE-合成については、 l の偶奇によらずに通過域を自然に拡張して $t \in [-1, 1]$ に変更できるので、通過域、遷移域、阻止域の各区域をそれぞれ t について $[-1, 1]$, $(1, \xi)$, $[\xi, \infty)$ にできることが分かる (なおC-合成とE-合成で l が奇数である場合には、合成された伝達関数 $\hat{g}(t)$ の実数の極はただ 1 つでその値は -1 よりも小さい)。

3.1.4.1 次数 l が偶数の場合

次数 l が偶数の場合は、4種類の合成で得られるフィルタの伝達関数 $\hat{g}(t)$ は実数の極を持たないので、求めたい固有対の固有値の区間 $[a, b]$ は位置を任意に設定できる。そうして4種類の合成のどの場合にも、拡張された通過域 $t \in [-1, 1]$ を固有値の区間 $\lambda \in [a, b]$ に対応させることができる。ただしそのように対応させると、B-合成とI-合成の場合の伝達関数の原点 $t = 0$ 付近における高度な平坦性 (バターワース特性) は、指定した固有値の区間の中央 $(a + b)/2$ 付近の固有値に対するものとなる。またB-合成とI-合成の伝達関数 $\hat{g}(t)$ は原点 $t = 0$ で最大値をとり $t \in [0, 1]$ では単調減少であるから、通過域を元の $t \in [0, 1]$ に制限するときには固有値が区間 $[a, b]$ の下側にある固有対ほど良く近似される傾向を持つ。そこでB-合成やI-合成で固有対の固有値が指定した区間の下側にあるものほど良い近似で求めたいという場合には、通過域を区間 $t \in [-1, 1]$ にせずにあえて区間 $t \in [0, 1]$ にすることが考えられる。すなわち l が偶数である場合でも、目的によっては固有値の区間 $\lambda \in [a, b]$ に区間 $t \in [0, 1]$ を対応させることも考えられる。

3.1.4.2 次数 l が奇数の場合

次数 l が奇数の場合には、4種類のフィルタの伝達関数は実数の極を 1 つ持つので、フィルタを安全に適用できるのは固有値分布の下端付近 (あるいは上端付近) の固有値だけを求める場合である。下端付近の固有値を求める場合は、固有値の座標で最小固有値よりも下側の位置にその実数の極を対応させる。上端付近の固有値を求める場合には、最大固有値よりも上側の位置にその実数の極を対応させる。以下では下端付近に対する場合に限って記述する。

いま a が最小固有値 λ_{\min} 以下の実数とする。C-合成とE-合成では拡大された通過域 $t \in [-1, 1]$ を固有値の区間 $\lambda \in [a, b]$ に対応させることができる。他方でB-合成とI-合成では $t \in [-1, 1]$ を $\lambda \in [a, b]$ に対応させると伝達関数

$\hat{g}(t)$ の値は $t \in [-1, 0)$ で 1 を超える. すると, 伝達関数の式を最大値を 1 にするように再規格化を行ってそれにとまって g_p と g_s の値も変更しない限りは, 通過域における伝達関数の最大値は 1 と決めたことを守る立場では, 固有値の区間 $\lambda \in [a, b]$ に対応させる通過域は区間 $t \in [0, 1]$ に制限せざるを得ない. その場合には, B-合成と I-合成は区間 $t \in [0, 1]$ の下端付近でバターワース特性を持ち, フィルタの伝達関数は区間 $\lambda \in [a, b]$ において単調減少になる.

注記: あるいは, 固有値が存在する領域において伝達関数の値が閾値 g_p 以上である場所は通過域であると定義しなおして, 伝達関数の最大値を 1 とする制限を緩和するからである. そのような定義の変更をした結果, 伝達関数の通過域における最大値が g_{\max} であれば, 通過域における伝達関数の最大最小比は g_{\max}/g_p になる. この比の値が大きいかいほど近似固有対の精度の不均一さの傾向も大きくなる.

有理関数の合成により伝達関数 $\hat{g}(t) = g(h(t))$ を得たなら, 次はそれに対応するフィルタ \hat{F} をレゾルベントの線形結合 (の実部) の多項式 P として構成する. 本論文ではフィルタは簡易型とする, つまり n 次多項式 P が Chebyshev 多項式 T_n を用いて表される場合に限ることにする.

3.2 簡易型の伝達関数と有理関数 $h(t)$ の合成

伝達関数 $g(t)$ がレゾルベントの線形結合 (の実部) の多項式で, その多項式が 1 つの Chebyshev 多項式 T_n を用いて表されるとき, ここではそのような伝達関数を簡易型であるということにする.

合成を施す前の簡易型の伝達関数を式 (10) とする. この式が含むパラメータの組は (n, μ, σ) であり, それを本論文では標準パラメータの 3 つ組と呼ぶことにする.

$$g(t) = g_s T_n(2x(t) - 1), \quad x(t) = \frac{\mu + \sigma}{t + \sigma}. \quad (10)$$

有理関数 $h(t)$ を合成した後の伝達関数 $\hat{g}(t) = g(h(t))$ は式 (11) で与えられる.

$$\hat{g}(t) = g_s T_n(2\hat{x}(t) - 1), \quad \hat{x}(t) = \frac{\mu + \sigma}{h(t) + \sigma}. \quad (11)$$

標準パラメータの 3 つ組 (n, μ, σ) から閾値 g_s と g_p を求めるには, 式 (12) を上側から順に計算すればよい.

$$\begin{cases} g_s \leftarrow 1 / \cosh \left(2n \sinh^{-1} \sqrt{\frac{\mu}{\sigma}} \right), \\ g_p \leftarrow g_s \cosh \left(2n \sinh^{-1} \sqrt{\frac{\mu - 1}{\sigma + 1}} \right). \end{cases} \quad (12)$$

有理関数 $h(t)$ の合成により, 閾値 g_s と g_p をまったく変えずに遷移域の幅だけを関係式 (9) に従って $\mu - 1$ から $\xi - 1$ に縮小できる.

4. 合成された伝達関数に対応するフィルタの構成

この 4 章の内容は文献 [14] に対して E-合成の記述が加

わった以外にも, C-合成の場合の有理関数 $h(t)$ の定義式を次数 ℓ の偶奇による場合分けを含まないものに変更したことなどにより本質的に異なる. また記号や記法は本論文に適合するように改めた.

いま P を多項式とすると, 合成関数 $\hat{g}(t) = P(\hat{x}(t))$ を伝達関数とするフィルタ \hat{F} は以下の手順で構成できる.

遠方で有界であって重複する極を持たない有理関数 $\hat{x}(t)$ の式 (13) から, 式 (14) の右辺の形をした複素数範囲での部分分数分解を構成する (4.1 節参照).

$$\hat{x}(t) \equiv x(h(t)) = \frac{\mu + \sigma}{h(t) + \sigma}. \quad (13)$$

$$\hat{x}(t) = c_\infty + \sum_{j=1}^{\ell} \frac{c_j}{t - t_j}. \quad (14)$$

すると $\hat{x}(t)$ に対応する作用素 $\hat{\mathcal{X}}$ は, $\hat{x}(t)$ の部分分数分解と対応する形の式 (15) で表されて, 恒等作用素 I と ℓ 個のレゾルベントの線形結合になる. ここで各シフト ρ_j には正規化座標による極 t_j が対応している.

$$\hat{\mathcal{X}} = c_\infty I + \sum_{j=1}^{\ell} \gamma_j \mathcal{R}(\rho_j). \quad (15)$$

さらに, いま扱っている実対称定値一般固有値問題では, $\hat{\mathcal{X}}$ は実線形作用素, つまり実ベクトル上の作用素であることから, 複素共役に対する対称性を利用して複素ベクトルの実部を取り出す操作を用いて, たとえば式 (15) をシフトの虚部が非負であるレゾルベントの項だけを含む式に書き換えることができる. それにより, 4 種類の合成 (B, C, I, E) の各場合の $\hat{\mathcal{X}}$ の作用は, シフトの虚部が正であるレゾルベントの項 $\lfloor \ell/2 \rfloor$ 個, それに ℓ が奇数のときにだけシフトが実数であるレゾルベントの項を 1 つ追加して表すことができる (4.2 節参照).

伝達関数が $\hat{g}(t) = P(\hat{x}(t))$ である場合のフィルタは $\hat{F} = P(\hat{\mathcal{X}})$ で与えられる. 特に伝達関数が簡易型で $\hat{g}(t) = g_s T_n(2\hat{x}(t) - 1)$ である場合のフィルタは $\hat{F} = g_s T_n(2\hat{\mathcal{X}} - I)$ で与えられる (4.3 節参照).

4.1 各合成に対する有理関数 $\hat{x}(t)$ の部分分数分解の構成

4 種類の合成 (B, C, I, E) で用いる関数 $h(t)$ は, 次数 ℓ の有理関数 (B-合成と C-合成の場合には特に多項式) である. それぞれの場合に, フィルタをレゾルベントを用いて構成するためには, 式 (13) で表される実有理関数 $\hat{x}(t)$ から式 (14) の右辺の形の部分分数分解に現れる数値を具体的に構成する必要がある (いまの 4 種類の合成の場合には有理関数 $\hat{x}(t)$ の極は全部で ℓ 個で, 重複するものはない). ここで定数項の係数 c_∞ は実数であり, 極 t_j とその係数 $c_j, j=1, 2, \dots, \ell$ は複素数である.

- 定数項の係数 c_∞ は (式 (14) から容易に導かれる) 式 (16) の計算をすることにより求められる.

$$c_\infty = \lim_{t \rightarrow \infty} \hat{x}(t). \quad (16)$$

- 極 t_j は方程式 (17) (これは式 (13) で $\hat{x}(t)$ の分母が零になる条件である) の解 t として求められる。

$$h(t) + \sigma = 0. \quad (17)$$

- 極 t_j の係数 c_j は (これは式 (14) の形から容易に導かれることであるが) 式 (18) を計算することにより求められる。

$$c_j = \lim_{t \rightarrow t_j} \{(t - t_j)\hat{x}(t)\}. \quad (18)$$

4.1.1 部分分数分解の構成 (B-合成の場合)

ここでは B-合成の場合の $\hat{x}(t)$ の部分分数分解 (14) を構成する。その際に有理関数 $\hat{x}(t) = x(h(t))$ の式に現れている μ と σ と次数 ℓ の 3 つは既知とする。

B-合成に用いる有理関数 $h(t)$ は多項式であることから、定数項の係数 c_∞ はつねに零になる。

複素数の ℓ 個の極 $t = t_j$ は方程式 $t^\ell + \sigma = 0$ の解なので、式 (19) で表されて、重複するものはない。なお、虚部が正である極の添字はちょうど前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ である。実数の極 t_R は ℓ が奇数のときに限りただ 1 つ存在してその添字は $j = (\ell+1)/2$ であり、値は $t_R \equiv t_{(\ell+1)/2} = -\sigma^{1/\ell}$ で負である (その存在範囲は $\sigma > 1$, $\sigma = 1$, $\sigma < 1$ に従ってそれぞれ $t_R < -1$, $t_R = -1$, $t_R > -1$ となる)。

$$t_j = \sigma^{1/\ell} \omega^{2j-1}, j=1, 2, \dots, \ell, \quad \omega \equiv \exp\left(\frac{\pi}{\ell} \sqrt{-1}\right). \quad (19)$$

極 t_j の係数 c_j , $j=1, 2, \dots, \ell$ は式 (20) で与えられる。そうして ℓ が奇数のときに限りただ 1 つ存在する実数の極 $t_R \equiv t_{(\ell+1)/2}$ に対する係数 $c_R \equiv c_{(\ell+1)/2}$ は実数である。

$$c_j = -\frac{\mu + \sigma}{\sigma \ell} \times t_j. \quad (20)$$

4.1.2 部分分数分解の構成 (C-合成の場合)

ここでは、C-合成の場合の $\hat{x}(t)$ の部分分数分解 (14) を構成する。その際に有理関数 $\hat{x}(t) = x(h(t))$ の式に現れている μ と σ と次数 ℓ の 3 つは既知とする。

C-合成に用いる有理関数 $h(t)$ は多項式であることから、定数項の係数 c_∞ はつねに零になる。

極 t_j を求めるには $\hat{x}(t)$ の分母が零になる条件 $h(t) + \sigma = 0$, つまり方程式 $T_\ell(t) = -(1 + 2\sigma)$ を解けばよい。そこでまず 3 つの正の実数 α , d_R , d_I をそれぞれ式 (21) を用いて順に計算する (式の形から $d_R > 1$ であることも分かる)。

$$\begin{cases} \alpha \leftarrow (\sqrt{\sigma} + \sqrt{1 + \sigma})^{2/\ell}, \\ d_R \leftarrow (\alpha + 1/\alpha)/2, \\ d_I \leftarrow (\alpha - 1/\alpha)/2. \end{cases} \quad (21)$$

あるいは α の計算の中で $2/\ell$ 乗乗を対数と指数の計算を

用いて求めるのであれば、むしろ $\log \alpha \leftarrow (2/\ell) \log(\sqrt{\sigma} + \sqrt{1 + \sigma})$, $d_R \leftarrow \cosh(\log \alpha)$, $d_I \leftarrow \sinh(\log \alpha)$ として計算する方がよいであろう。

するとすべての複素数の極 t_j , $j = 1, 2, \dots, \ell$ は式 (22) により表されて、重複するものはない。なお、虚部が正である極に対する添字はちょうど前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ である。実数の極は ℓ が奇数のときに限りただ 1 つ存在してその添字は $j = (\ell+1)/2$ であり、値は $t_R \equiv t_{(\ell+1)/2} = -d_R$ であって -1 よりも小さい。

$$t_j = d_R \cos \frac{(2j-1)\pi}{\ell} + \sqrt{-1} \cdot d_I \sin \frac{(2j-1)\pi}{\ell}, \quad (22) \\ j=1, 2, \dots, \ell.$$

極 $t = t_j$ の係数 c_j は式 (23) を用いて計算できる。その式中の $(d/dt)h(t)|_{t=t_j}$ の値は、公式 $(d/dz)T_m(z) = mU_{m-1}(z)$ を用いて計算できる。ただし $U_m(z)$ は m 次の第 2 種 Chebyshev 多項式を表し、その値は漸化式 (24) を用いて計算できる。すると $(d/dt)h(t)|_{t=t_j}$ の値は $(\ell/2)U_{\ell-1}(t_j)$ となるので、極 t_j の係数 c_j は最終的には式 (25) により計算できる。

$$c_j = \lim_{t \rightarrow t_j} \{(t - t_j)\hat{x}(t)\} = \frac{\mu + \sigma}{\frac{d}{dt}h(t)|_{t=t_j}}. \quad (23)$$

$$\begin{cases} U_0(z) = 1, \quad U_1(z) = 2z, \\ U_m(z) = 2zU_{m-1}(z) - U_{m-2}(z) \quad (m \geq 2). \end{cases} \quad (24)$$

$$c_j = \frac{2(\mu + \sigma)}{\ell} \times \frac{1}{U_{\ell-1}(t_j)}. \quad (25)$$

4.1.3 部分分数分解の構成 (I-合成の場合)

ここでは、I-合成の場合の $\hat{x}(t)$ の部分分数分解 (14) を構成する。その際に有理関数 $\hat{x}(t) = x(h(t))$ の式に現れている μ と σ と次数 ℓ と ξ の 4 つは既知とする。

I-合成に用いる有理関数 $h(t)$ は $\ell = 2$ 以外では多項式ではないので、定数項の係数 c_∞ の値は必ずしも零にはならず、次数 ℓ の 4 を法とする剰余で場合が分かれて、式 (26) により与えられる。

$$c_\infty = \begin{cases} \frac{\mu + \sigma}{2\mu + \sigma} & (\text{mod}(\ell, 4) = 1, 3 \text{ の場合}), \\ 1 & (\text{mod}(\ell, 4) = 0 \text{ の場合}), \\ 0 & (\text{mod}(\ell, 4) = 2 \text{ の場合}). \end{cases} \quad (26)$$

極 t_j は $h(t) + \sigma = 0$ の解 t であるから $T_\ell(\xi/t) = -(1 + 2\mu/\sigma)$ を満たす。そこで $z = \xi/t$ とおくと、 z の方程式は $T_\ell(z) = -(1 + 2\mu/\sigma)$ になる。すると C-合成の場合と同様に、まず 3 つの正の実数 α , d_R , d_I をそれぞれ式 (27) を用いて順に計算する (式の形から $d_R > 1$ であることも分かる)。

$$\begin{cases} \alpha \leftarrow (\sqrt{\mu/\sigma} + \sqrt{1 + \mu/\sigma})^{2/\ell}, \\ d_R \leftarrow (\alpha + 1/\alpha)/2, \\ d_I \leftarrow (\alpha - 1/\alpha)/2. \end{cases} \quad (27)$$

あるいは (これも C-合成の場合と同様に), $\log \alpha \leftarrow (2/\ell) \log(\sqrt{\mu/\sigma} + \sqrt{1 + \mu/\sigma})$, $d_R \leftarrow \cosh(\log \alpha)$, $d_I \leftarrow \sinh(\log \alpha)$ として計算することもできる.

するとすべての複素数の極 t_j , $j=1, 2, \dots, \ell$ は式 (28) により表されて, 重複がないことも分かる. z_j の式の中では虚部の符号を C-合成の場合とは反対にとった. すると虚部が正である極に対する添字はちょうど前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ である. 実数の極は ℓ が奇数の場合に限りただ 1 つ存在してその添字は $j = (\ell + 1)/2$ であり, 値は $t_R \equiv t_{(\ell+1)/2} = -\xi/d_R$ となる.

$$\begin{cases} z_j = d_R \cos \frac{(2j-1)\pi}{\ell} - \sqrt{-1} \cdot d_I \sin \frac{(2j-1)\pi}{\ell}, \\ t_j = \xi/z_j, \quad j=1, 2, \dots, \ell. \end{cases} \quad (28)$$

ℓ が奇数のとき t_R は代数方程式 $T_\ell(\xi/t) + 1 + 2\mu/\sigma = 0$ のただ 1 つの実根である. いまこの方程式の左辺を $\phi(t)$ と書けば, $\phi(-\xi) = 2\mu/\sigma$ は正である. また $\phi(-1) = -T_\ell(\xi) + 1 + 2\mu/\sigma$ であるが, $\mu = \frac{1}{2}\{1 + T_\ell(\xi)\}$ であることから, $\phi(-1) = 2\{1 - \mu(1 - 1/\sigma)\}$ となり, もしも $\mu > 1$ や $\sigma > 0$ がある程度大きければこの値は負になり, そのとき方程式 $\phi(t) = 0$ のただ 1 つの実根 t_R は区間 $(-\xi, -1)$ に含まれるので $t_R < -1$ である.

極 t_j の係数 c_j は式 (29) で表せることが数式の計算により示せる. ただし第 2 種 Chebyshev 多項式 $U_m(z)$ は式 (24) で計算する. なお ℓ が奇数のときに限りただ 1 つ存在する実数の極 $t_R \equiv t_{(\ell+1)/2}$ に対する係数 c_R は実数で, $c_R \equiv c_{(\ell+1)/2}$ である.

$$c_j = \frac{2(\sigma + \mu)\mu}{\ell\sigma^2\xi} \times \frac{t_j^2}{U_{\ell-1}(\xi/t_j)}. \quad (29)$$

4.1.4 部分分数分解の構成 (E-合成の場合)

E-合成の場合の $\hat{x}(t)$ の部分分数分解 (14) を構成する計算手順の導出の詳細は, その準備に用いる楕円有理関数などについての記述が長いので, 後の 6.3 節と 6.4 節に移して, ここでは計算方法だけを記述する.

有理関数 $\hat{x}(t) = x(h(t))$ の部分分数分解を行う際には, その式に含まれる $\mu, \sigma, \ell, L, \xi$ の 5 つの値は既知とする (実際には, μ から L が式 (60) により決まり, また「次数方程式」(49) を式 (53) を用いて解くと L と ℓ から ξ も決まるので, $\hat{x}(t)$ は 3 つの値 μ, σ, ℓ から完全に決まるのであるが).

まず定数項の係数 c_∞ は以下の式 (30) で与えられる.

$$c_\infty = \begin{cases} \frac{2(\mu + \sigma)}{L + (2\sigma + 1)} & (\text{mod}(\ell, 4) = 1, 3 \text{ の場合}), \\ 1 & (\text{mod}(\ell, 4) = 0 \text{ の場合}), \\ 0 & (\text{mod}(\ell, 4) = 2 \text{ の場合}). \end{cases} \quad (30)$$

次に, 実数 y を式 (73) で定義される第 1 種楕円積分 $F(\varphi, k)$ を用いて, 以下の式 (31) で求める. ただし $k' = (L^{-1})'$ は母数 $k = L^{-1}$ に対する補母数 ($k' = \sqrt{1 - k^2}$) を表す. k の値は小さいので精度の面から, 実際の数値計算では $F(\varphi, k')$ の値を求めるのには第 1 種楕円積分の値を求める計算ルーチンを直接用いて $F(\varphi, \sqrt{1 - k^2})$ とするのではなくて, 与えられた φ と k から (k' の値を作らずに) $F(\varphi, k')$ の値を直接求める計算ルーチン $F'(\varphi, k)$ を特別に用意して, それを用いて $y \leftarrow F'(\varphi, L^{-1})$ として計算する.

$$\begin{cases} \varphi \leftarrow \cos^{-1} \frac{\sqrt{L^2 + 2(2\sigma + 1)L + 1}}{(2\sigma + 1)L + 1}, \\ y \leftarrow F(\varphi, (L^{-1})'). \end{cases} \quad (31)$$

実数 y を求めたら複素数の極 t_j とその係数 c_j , $j=1, 2, \dots, \ell$ は以下の式 (32) により計算できる. ただし, $K(k)$ は母数 k の第 1 種完全楕円積分 (48) を表す. t_j を求める式中の Jacobi の楕円関数 sn の引数は複素数なので, それを実数引数の Jacobi の楕円関数 $\text{sn}, \text{cn}, \text{dn}$ の値を組み合わせて式 (76) により計算する. また関数 $\Psi(t)$ は楕円有理関数 $R_\ell(\xi, t)$ の対数微分を表し, その値は複素数の引数 t に対して式 (78) を用いて計算できる. なお式 (78) の中に現れる実数の値 x_i と \tilde{x}_i は, 式 (57) により実数引数に対する楕円関数 sn を計算して得られる.

$$\begin{cases} \omega_j = \frac{1}{\ell} \left\{ 4j - 2 - \sqrt{-1} \times \frac{y}{K(L^{-1})} \right\}, \\ t_j = -\text{sn} \{ (\omega_j - 1)K(\xi^{-1}), \xi^{-1} \}, \\ c_j = -\frac{2(\mu + \sigma)(L^2 - 1)}{\{L + (2\sigma + 1)\}\{(2\sigma + 1)L + 1\}} \times \frac{1}{\Psi(t_j)}. \end{cases} \quad (32)$$

4.2 $\hat{x}(t)$ の部分分数分解に対応する作用素 $\hat{\mathcal{X}}$ の構成

いま 4 種類の各合成 (B, C, I, E) について, 次数 ℓ の有理関数 $h(t)$ を $x(t)$ に合成した式 $\hat{x}(t) = x(h(t))$ に対する部分分数分解 (14) をそれに現れる係数 c_∞ や極 t_j とその係数 c_j , $\ell=1, 2, \dots, \ell$ を具体的な数値を決めて構成したとする. なお虚部が正である極は添字が前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ とする. 実数である極は ℓ が奇数のときに限りただ 1 つ存在して添字は $j = (\ell + 1)/2$ で, その値を t_R と表す.

複素共役対称性を利用すると, 複素数から実部を取り出す操作 Re を用いて部分分数分解の式 (14) を書き換えることができる (複素共役な虚数の極の項の対の寄与の和は, どちらか片方の寄与の実部の 2 倍になる. 極の虚部の符号それ自身には意味はなくて, 複素共役な項の対からその片方だけを選出する手段に用いているだけである).

表 1 次数 ℓ のフィルタ構成に必要なレゾルベントのシフトの数
Table 1 Numbers of imaginary shifts and real shift for resolutions required to construct a filter of degree ℓ .

| ℓ | シフトが虚数のもの | シフトが実数のもの |
|--------|-----------|-----------|
| 2 | 1 | 0 |
| 3 | 1 | 1 |
| 4 | 2 | 0 |
| 5 | 2 | 1 |
| 6 | 3 | 0 |
| 7 | 3 | 1 |
| 8 | 4 | 0 |

$$\hat{x}(t) = \begin{cases} c_\infty + \sum_{j=1}^{\ell/2} \operatorname{Re} \left(\frac{2c_j}{t-t_j} \right) & (\ell \text{ が偶数}), \\ c_\infty + \frac{c_R}{t-t_R} + \sum_{j=1}^{(\ell-1)/2} \operatorname{Re} \left(\frac{2c_j}{t-t_j} \right) & (\ell \text{ が奇数}). \end{cases} \quad (33)$$

すると実有理関数 $\hat{x}(t)$ に対応する実線形作用素 $\hat{\mathcal{X}}$ は、 ℓ が偶数の場合には式 (34) と (36) の組により、 ℓ が奇数の場合には式 (37) と (39) の組により、それぞれ与えられる。

このように複素ベクトルの実部をとる操作 Re を用いることで $\hat{\mathcal{X}}$ の作用をレゾルベントの数を減らして表せば、 ℓ が偶数の場合にはシフトが虚数のものが $\ell/2$ 個になり、 ℓ が奇数の場合にはシフトが虚数のものが $(\ell-1)/2$ 個およびシフトが実数のものが 1 つになる (表 1)。

4.2.1 レゾルベントのシフトと係数 (次数 ℓ が偶数の場合)

ℓ が偶数の場合には、求めたい固有対の固有値の区間 $\lambda \in [a, b]$ はその位置を任意に設定できて、固有値分布の内側 (中間固有値) でも良い。固有値の区間 $\lambda \in [a, b]$ と正規化座標で表した通過域 $t \in [-1, 1]$ を線形変換 $\lambda = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t$ を用いて対応させるものとする。いま部分分数分解 (33) の中の極が虚数 t_j で係数が c_j である項 $c_j/(t-t_j)$ がレゾルベントの項 $\gamma_j \mathcal{R}(\rho_j)$ に対応するならば、 $c_j/(t-t_j) = \gamma_j/(\lambda - \rho_j)$ の関係があり、それと λ と t の間の線形関係をあわせれば、レゾルベントの虚数のシフト ρ_j とそのレゾルベントの複素数の係数 γ_j はそれぞれ式 (34) で与えられる。

$$\rho_j \equiv \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t_j, \quad \gamma_j \equiv \frac{1}{2}(b-a)c_j. \quad (34)$$

あるいは、 ℓ が偶数で通過域を $t \in [-1, 1]$ に選べる場合であっても、たとえば B-合成や I-合成でなるべく固有値が区間 $[a, b]$ の下端側にある固有対ほど良い精度で求めたいという理由から、あえて通過域との対応を元の $t \in [0, 1]$ のままにするのであれば、 λ と t の間の線形関係は $\lambda = a + (b-a)t$ となるので、部分分数分解の極 t_j とその係数 c_j に対応するレゾルベントの虚数のシフト ρ_j とそのレゾルベントの複素数の係数 γ_j はそれぞれ式 (35) で与えられる。

$$\rho_j \equiv a + (b-a)t_j, \quad \gamma_j \equiv (b-a)c_j. \quad (35)$$

レゾルベントのシフトと係数が決まれば、実線形作用素

$\hat{\mathcal{X}}$ は式 (36) で与えられる。

$$\hat{\mathcal{X}} = c_\infty I + \sum_{j=1}^{\ell/2} \operatorname{Re} \{ 2\gamma_j \mathcal{R}(\rho_j) \}. \quad (36)$$

4.2.2 レゾルベントのシフトと係数 (次数 ℓ が奇数の場合)

ℓ が奇数の場合には、求めたい固有対の固有値の区間 $\lambda \in [a, b]$ の位置の設定は固有値分布の下端に限られ、 a は最小固有値 λ_{\min} 以下である。

4.2.2.1 B-合成と I-合成の場合

B-合成と I-合成の場合は通過域を (通常は) $t \in [0, 1]$ にとり、それを固有値の区間 $\lambda \in [a, b]$ に線形変換 $\lambda = a + (b-a)t$ で対応させる。いま部分分数分解 (33) の中の虚数の極 t_j とその係数の複素数 c_j の項にレゾルベントの項 $\gamma_j \mathcal{R}(\rho_j)$ が対応すれば、 ℓ が偶数である場合と同様に $c_j/(t-t_j) = \gamma_j/(\lambda - \rho_j)$ の関係がある。すると、レゾルベントの虚数のシフト ρ_j とその係数の複素数 γ_j 、および (同様に) 実数の極 ρ_R とその係数の実数 γ_R は、それぞれ式 (37) で与えられる。

$$\begin{cases} \rho_j \equiv a + (b-a)t_j, & \gamma_j \equiv (b-a)c_j, \\ \rho_R \equiv a + (b-a)t_R, & \gamma_R \equiv (b-a)c_R. \end{cases} \quad (37)$$

4.2.2.2 C-合成と E-合成の場合

C-合成と E-合成の場合には、それらに対する有理関数 $h(t)$ の性質から、 ℓ が奇数の場合にも通過域を区間 $t \in [-1, 1]$ に広げることができる。そうしてそれを固有値の区間 $\lambda \in [a, b]$ に線形変換 $\lambda = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t$ により対応させることができる。その場合のレゾルベントの虚数のシフト ρ_j とその係数の複素数 γ_j 、および (同様に) 実数のシフト ρ_R とその係数の実数 γ_R はそれぞれ式 (38) で与えられる。

$$\begin{cases} \rho_j \equiv \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t_j, & \gamma_j \equiv \frac{1}{2}(b-a)c_j, \\ \rho_R \equiv \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t_R, & \gamma_R \equiv \frac{1}{2}(b-a)c_R. \end{cases} \quad (38)$$

レゾルベントのシフトと係数が決まれば、実線形作用素 $\hat{\mathcal{X}}$ は式 (39) で与えられる。

$$\hat{\mathcal{X}} = c_\infty I + \gamma_R \mathcal{R}(\rho_R) + \sum_{j=1}^{(\ell-1)/2} \operatorname{Re} \{ 2\gamma_j \mathcal{R}(\rho_j) \}. \quad (39)$$

4.3 合成から得られた簡易型フィルタ $\hat{\mathcal{F}}$ の作用の計算

有理関数を合成して得られた伝達関数 $\hat{g}(t) = g_s T_n (2\hat{x} - 1)$ に対応する簡易型のフィルタ $\hat{\mathcal{F}}$ は式 (40) で与えられる。

$$\hat{\mathcal{F}} = g_s T_n (2\hat{\mathcal{X}} - I). \quad (40)$$

いま $\hat{\mathcal{Y}} \equiv 2\hat{\mathcal{X}} - I$ とおいて、与えられた m 個のベクトルの組を列として並べた $N \times m$ 行列 V に対して $V^{(k)} \equiv T_k(\hat{\mathcal{Y}})V$

と定義すると、 V から開始して 3 項漸化式 (41) を用いることで $V^{(n)}$ が求められる。

$$\begin{cases} V^{(0)} = V, \\ V^{(1)} = \hat{Y}V, \\ V^{(k)} = 2\hat{Y}V^{(k-1)} - V^{(k-2)} \quad (k \geq 2 \text{ の場合}). \end{cases} \quad (41)$$

そうしてフィルタ \hat{F} の V に対する作用は、 V から始めて 3 項漸化式を用いて計算された $V^{(n)}$ に対して定数 g_s を乗じる形の式 (42) により与えられる。

$$\hat{F}V = g_s V^{(n)}. \quad (42)$$

5. 合成で得られる簡易型のフィルタの設計法

この 5 章の内容は、文献 [14] の 2.11 節にも部分的に含まれてはいるが、より網羅的なものである。

合成を施す前のフィルタ \mathcal{F} の伝達関数 $g(t)$ は、 n 次の Chebyshev 多項式を用いる簡易型の設計であれば式 (10) で与えられる。この $g(t)$ の式に含まれるパラメータの組 (n, μ, σ) を「標準パラメータの 3 つ組」と呼ぶことにする (係数 g_s の値は $g(t)$ の通過域における最大値を 1 とする条件、すなわちいまの場合には $g(0) = 1$ から決まる)。他方でフィルタの伝達特性を特徴づける 3 つの形状パラメータの組は (g_p, g_s, μ) である。

また、この伝達関数 $g(t)$ と l 次の有理関数 $h(t)$ の関数合成で得られる伝達関数は $\hat{g}(t) = g(h(t))$ であり、伝達率の閾値である g_p と g_s は $g(t)$ と $\hat{g}(t)$ の双方でそれぞれ共通の値になるが、阻止域の端の位置を指定するパラメータ (あるいは遷移域の幅を決めるパラメータ) は元の伝達関数 $g(t)$ では μ であるが、合成後の伝達関数 $\hat{g}(t)$ では ξ になる (ただし $\mu = h(\xi)$ の関係がある)。合成されたフィルタの伝達特性を特徴付ける 3 つの形状パラメータは (g_p, g_s, ξ) であり、フィルタの設計に際してこれらの 3 つすべてもしくはその一部を指定したい場合がある。

そこで簡易型の伝達関数あるいはそれに合成を施したものについて、指定されたパラメータの組から伝達関数とそれに対応するフィルタを具体的に決定する方法を以下に記述する。ただし合成の種類 (B, C, I, E) とその合成に用いる有理関数 $h(t)$ の次数 l は先に決めてそれから設計を行うことにする。

5.1 標準パラメータの組 (n, μ, σ) を指定する場合

Chebyshev 多項式の次数 n とパラメータ $\mu (> 1)$ と $\sigma (> 0)$ の 3 つ、すなわち標準パラメータの 3 つ組 (n, μ, σ) を指定する場合は、まず式 (12) を計算して g_s と g_p の値を求める。すると合成前の $g(t)$ とその形状パラメータの 3 つ組である (g_p, g_s, μ) が決まる。

元の伝達関数 $g(t)$ に $h(t)$ を合成して得られる $\hat{g}(t) =$

$g(h(t))$ は閾値の値である g_p と g_s は $g(t)$ と共通になるが、遷移域の幅に関係する ξ の値については μ から関係 (9) を用いて計算して求める。そうして合成された $\hat{g}(t)$ の形状パラメータの 3 つ組は (g_p, g_s, ξ) である。

μ と σ と合成の種類 (B, C, I, E) と合成の次数 l が決まれば $\hat{x}(t)$ が決まる。それから $\hat{x}(t)$ の部分分数分解 (14) が決まる。そうして部分分数分解に対応してレゾルベントの線形結合の実部で表された作用素 $\hat{\mathcal{X}}$ が式 (36) あるいは式 (39) として決まり、それに Chebyshev 多項式の次数 n と g_s の値を用いて式 (40) のフィルタ \hat{F} が決まる。

5.2 組 (n, g_p, g_s) を指定する場合

Chebyshev 多項式の次数 n と伝達関数の閾値 g_p, g_s を指定する場合は、式 (12) の関係を逆に解いて μ と σ の値を求める。それには式 (43) を上から順に計算すればよい。

$$\begin{cases} w_1 \leftarrow \frac{1}{2n} \cosh^{-1} \frac{1}{g_s}, \\ w_2 \leftarrow \frac{1}{2n} \cosh^{-1} \frac{g_p}{g_s}, \\ \sigma \leftarrow \frac{\cosh^2 w_2}{\sinh(w_1 + w_2) \sinh(w_1 - w_2)}, \\ \mu \leftarrow \sigma \sinh^2 w_1. \end{cases} \quad (43)$$

これにより、指定されたパラメータの 3 つ組 (n, g_p, g_s) から標準パラメータの 3 つ組 (n, μ, σ) が決まるので、フィルタ \hat{F} の構成法は 5.1 節の「標準パラメータの 3 つ組 (n, μ, σ) を指定する場合」に帰着する。

5.3 形状パラメータの組 (g_p, g_s, ξ) を指定する場合

合成されたフィルタ \hat{F} の伝達関数 $\hat{g}(t)$ の形状パラメータの 3 つ組 (g_p, g_s, ξ) が指定された場合には、それから条件を満たす標準パラメータの 3 つ組 (n, μ, σ) が決まれば、フィルタの具体的な構成法は 5.1 節の「標準パラメータの 3 つ組 (n, μ, σ) を指定する場合」に帰着する。

しかしながら、形状パラメータの組 (g_p, g_s, ξ) が指定されたときに、それを正確に実現する標準パラメータの組 (n, μ, σ) を決めることは、次数 n の値が整数であって連続的に変えられる量ではないことから、一般には不可能である。そのため、ここでは形状パラメータの 3 つのうち 2 つについては値を指定して正確に一致させるが、残りの 1 つの値については不等式による制約条件を指定することにする。

以下の 3 通りのそれぞれの場合について、標準パラメータの 3 つ組 (n, μ, σ) を求める方法を以下に示す。

- (1) g_p と ξ は値を、 g_s には上限値 G_s をそれぞれ指定
- (2) g_s と ξ は値を、 g_p には下限値 G_p をそれぞれ指定
- (3) g_p と g_s は値を、 ξ には上限値 Ξ をそれぞれ指定

5.3.1 g_p と ξ は値を、 g_s には上限値 G_s を指定する場合

この場合は、次数 n には条件 $g_s \leq G_s$ を満たせる最小の

ものを採用することにする．さらに n の上限値 n_{\max} (たとえば 50 とする) を適切に設定して, $n = 1$ から始めて, 以下の方法で条件を満たせる n の値を順次に探索する．

まず式 (8) を用いて ξ から計算して μ を求める．そうして, すでに決まっている g_p と μ の値および探索の中で仮定する整数 n の値を用いて, 非線形方程式 (44) を解いて σ の値を決める．

$$F(\sigma) \equiv \frac{\cosh\left(2n \sinh^{-1} \sqrt{\frac{\mu-1}{\sigma+1}}\right)}{\cosh\left(2n \sinh^{-1} \sqrt{\frac{\mu}{\sigma}}\right)} - g_p = 0. \quad (44)$$

この非線形方程式 $F(x) = 0$ は適切に設定された x の非負の初期区間 $[x_L, x_H]$ から開始して 2 分法を用いて解ける．

まず $F(0) = -g_p < 0$ である．そうして x の値を順次増大させてゆき (たとえば小さい正の値から始めて毎回 2 倍ずつにする), 初めて $F(x) \geq 0$ が成り立つ x の値が見つければその x の値を初期区間の上端 x_H とし, その直前のまだ $F(x) < 0$ であった最後の x の値を初期区間の下端 x_L にとればよい．あるいは簡単に $x_L = 0$ と設定することもできる．

そうして 2 分法により σ を求めたら, 式 (45) を用いて, g_s の値を n と μ と σ から計算して求める．

$$g_s \leftarrow 1 / \cosh\left(2n \sinh^{-1} \sqrt{\frac{\mu}{\sigma}}\right). \quad (45)$$

このようにして求めた g_s の値が G_s 以下であれば条件を満たせたことになるので探索を終える．そうでなければ n の値を 1 つ増やして, n の値があらかじめ設定した上限値を超えたら探索は失敗とするが, そうでなければ新しい n の値について上記の方法により繰り返し調べる．

5.3.2 g_s と ξ は値を, g_p には下限値 G_p を指定する場合

ここでは次数 n として, 指定された条件を満たすことのできる最小のものを採用することにする．さらに n の値の上限値 n_{\max} をあらかじめ適切に決めておき (たとえば 50 とする), 条件を満たす標準パラメータの 3 つ組 (n, μ, σ) をその範囲の n についてだけ探索することにする．

まず式 (8) を用いて ξ から μ の値を計算で求める．そうしてすでに決まっている g_s と μ の値と探索でいま仮定している整数 n の値から, 式 (46) を順に計算することにより σ と g_p の値を求める．その計算で得られた g_p の値が条件 $g_p \geq G_p$ を満たす, そのような n の最小値を探索で求める．

$$\begin{cases} \sigma \leftarrow \mu / \sinh^2\left(\frac{1}{2n} \cosh^{-1} \frac{1}{g_s}\right), \\ g_p \leftarrow g_s \cosh\left(2n \sinh^{-1} \sqrt{\frac{\mu-1}{1+\sigma}}\right). \end{cases} \quad (46)$$

最小値の探索法は, たとえば整数 n を 1 から始めて n_{\max} まで順に上昇させて行うことができる．あるいは n についての 2 分法を用いて最初 1 と n_{\max} をそれぞれ上端と下端にして探索することもできる．なお, もしも $n = n_{\max}$ において不等式 $g_p \geq G_p$ が満たせなければ, n_{\max} 以下には条件を満たせる n は存在しない．

最小値 n とそのときの σ の値および $\mu = h(\xi)$ を並べて作った標準パラメータの 3 つ組 (n, μ, σ) が得られたとすると, フィルタ \hat{F} の構成は 5.1 節の「標準パラメータの 3 つ組 (n, μ, σ) を指定する場合」に帰着される．

なお本論文の趣旨からは外れるが, 合成の次数 ℓ にある程度大きい値を選んで次数 n が 1 で条件が満たされる場合には, それから導かれるフィルタ \hat{F} の構成は選択した合成の種類 (B, C, I, E) による「レゾルベントの線形結合の実部」になる．

5.3.3 g_p と g_s は値を, ξ には上限値 Ξ を指定する場合

この場合も次数 n として条件 $\xi \leq \Xi$ を満たせる最小のものを採用することにする．またさらに n の上限値 n_{\max} を適切に設定して探索を行う．

まず Ξ から関係 $M = h(\Xi)$ を用いて μ の上限値 M を求める．そうして, すでに決まっている g_p と g_s の値および探索中に仮定する整数 n の値を用いて, 式 (43) を順に計算して μ と σ の値を求める．このようにして求めた μ の値が M 以下であれば条件を満たせたことになるが, そうでなければ n のより大きい値を調べにいく．関数合成により得られたフィルタの伝達関数を表すには, 次数 n とこのようにして求めた μ と σ を並べた標準パラメータの 3 つ組 (n, μ, σ) が決まれば十分である．式 (9) を用いれば μ から求めた ξ が上限値 Ξ を超えていないかを確認することができる．

6. E-合成について

この 6 章の内容は, 以前の我々の文献 [14] には含まれていなかったものである．

6.1 楕円有理関数の紹介

まず E-合成で利用する「楕円有理関数」(elliptic rational function) について簡単に紹介する．

6.1.1 楕円有理関数の楕円関数による媒介変数表示

楕円有理関数 $R_\ell(\xi, t)$ は t についての ℓ 次の実有理関数で, 次数以外に実数のパラメータ $\xi (> 1)$ を 1 つ持ち, Jacobi の楕円関数 cd による媒介変数表示 (47) を持つ [2], [3] (ここでは媒介変数として複素数 ϕ を用いている)．

$$\begin{cases} R_\ell(\xi, t) = \text{cd}\{\ell \phi K(L^{-1}), L^{-1}\}, \\ t = \text{cd}\{\phi K(\xi^{-1}), \xi^{-1}\}. \end{cases} \quad (47)$$

なお $K(k)$ は式 (48) により定義される母数 k の第 1 種完全楕円積分である．

$$K(k) \equiv \int_0^{\pi/2} \frac{d\theta}{\sqrt{1-k^2 \sin^2 \theta}}. \quad (48)$$

関数 cd は偶関数で、奇関数である Jacobi の楕円 sn 関数を用いたその実数周期の 4 分の 1 である $K(k)$ だけずらしたものとして $\text{cd}(u, k) \equiv \text{sn}(u + K(k), k)$ により定義される (式 (47) は、偶関数の三角関数 \cos と変数 θ を用いた Chebyshev 多項式の媒介変数表示 $T_n(t) = \cos n\theta$, $t = \cos \theta$ の類似になっている).

この媒介変数表示は 2 つの実数の定数パラメータ $\xi > 1$, $L > 1$ を含んでいるが、媒介変数表示が ℓ 次の有理関数を表すためにはそれら 2 つのパラメータは「次数方程式」と呼ばれる条件を満たす必要があるため独立なものではなくて、実際には両者は次数 ℓ に応じて決まる 1 対 1 の対応関係にある (6.1.2 項, 6.1.3 項).

6.1.2 パラメータ ξ と L が満たす関係 (次数方程式)

媒介変数表示 (47) による関数 $R_\ell(\xi, t)$ が ℓ 次の有理関数を表すためには、パラメータ ξ と L は「次数方程式」(degree equation) と呼ばれる式 (49) を満たす必要がある [2], [3].

$$K'(L^{-1})/K(L^{-1}) = \ell \times K'(\xi^{-1})/K(\xi^{-1}). \quad (49)$$

ここで $K'(k)$ は $K'(k) \equiv K(k')$ で定義される関数で、 $k' \equiv \sqrt{1-k^2}$ は母数 k の補母数を表している.

母数 k に対する楕円 nome 関数 $q(k)$ は式 (50) で定義される.

$$q(k) \equiv \exp \{-\pi K'(k)/K(k)\}. \quad (50)$$

この楕円 nome 関数 $q(k)$ を用いて「次数方程式」(49) を書き直せば式 (51) になる.

$$q(L^{-1}) = \{q(\xi^{-1})\}^\ell. \quad (51)$$

6.1.3 「次数方程式」の解法

楕円有理関数の次数が ℓ であるとき、楕円 nome 関数 $q(k)$ とその逆関数である $q^{-1}(z)$ を用いて「次数方程式」に等価な式 (51) を解けば、式 (52) により ξ から L を求めることができ、同様に式 (53) により L から ξ を求めることもできる (なお楕円 nome 関数の p 乗の値 $\{q(k)\}^p$ は、 $q(k)$ の値を求めてそれを p 乗するのではなく、 $q(k)$ の定義式 (50) の指数関数 \exp の引数を p 倍した式で計算する方がよい).

$$L = 1/q^{-1}(\{q(\xi^{-1})\}^\ell). \quad (52)$$

$$\xi = 1/q^{-1}(\{q(L^{-1})\}^{1/\ell}). \quad (53)$$

楕円 nome 関数の逆関数 $q^{-1}(z)$ は巾級数展開を用いた式 (54) を用いて計算できる. 正数 z (< 1) の値が小さければこの式に含まれる級数の収束は速い.

$$\begin{aligned} q^{-1}(z) &= \left\{ \frac{\sum_{m=1}^{\infty} z^{(m-\frac{1}{2})^2}}{1 + 2 \sum_{m=1}^{\infty} z^{m^2}} \right\}^2 \\ &= 4\sqrt{z} \times \left\{ \frac{\sum_{m=1}^{\infty} z^{m(m-1)}}{1 + 2 \sum_{m=1}^{\infty} z^{m^2}} \right\}^2 \\ &= 4\sqrt{z} \times \left(\frac{1 + z^2 + z^6 + z^{12} + \dots}{1 + 2z + 2z^4 + 2z^9 + \dots} \right)^2. \end{aligned} \quad (54)$$

なお、次数方程式の解法には楕円 nome 関数とその逆関数を使わないものもある [2], [3].

6.1.4 「楕円有理関数」の性質

楕円有理関数 $R_\ell(\xi, t)$ は t の ℓ 次の実有理関数で、以下の性質を満たす ($\xi > 1$ と $L > 1$ はすでに次数方程式を満たしているものとする) [2], [3].

- (1) $R_\ell(\xi, t)$ は次数 ℓ の偶奇に従って偶関数、奇関数.
- (2) $t \in [0, 1]$ で等リプル性を持ち、 $|R_\ell(\xi, t)| \leq 1$.
- (3) $t \in (1, \xi)$ で単調増加. $R_\ell(\xi, 1) = 1$, $R_\ell(\xi, \xi) = L$.
- (4) $t \in [\xi, \infty)$ では関数値の逆数が等リプル性を持ち、 $L \leq |R_\ell(\xi, t)|$.

それ以外の有用な性質として式 (55) を満たす.

$$R_\ell(\xi, \xi/t) = L/R_\ell(\xi, t). \quad (55)$$

区間 $[0, 1]$ で値の絶対値が 1 以下であり、 $t = 1$ での値が 1 であるすべての偶または奇の有理関数のうちで、 $t = 1$ における増加率が最大になるのが楕円有理関数であり、そのことを利用して遷移域 $(1, \xi)$ における値の変化が最も急峻な伝達関数を楕円有理関数を用いて実現できる.

6.1.5 楕円有理関数の明示的な有理関数表示

パラメータ ξ を持つ ℓ 次の楕円有理関数 $R_\ell(\xi, t)$ は、すべての零点と極を用いて因子に分解された明示的な有理関数表示 (56) で表せる [2], [3].

$$R_\ell(\xi, t) = C t^{\text{mod}(\ell, 2)} \prod_{j=1}^{\lfloor \ell/2 \rfloor} \frac{t^2 - x_j^2}{t^2 - \tilde{x}_j^2}. \quad (56)$$

ここで零点 $\pm x_j$ と極 $\pm \tilde{x}_j$, $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ はすべて実数であり、引数が実数の Jacobi 楕円関数 sn を用いた式 (57) で表される (ℓ が奇数の場合にはこれらに加えて零点 $t = 0$ と極 $t = \infty$ も持つ).

$$\begin{cases} x_j = \text{sn} \left[\frac{1}{\ell} \{2j - 1 + \text{mod}(\ell, 2)\} K(\xi^{-1}), \xi^{-1} \right], \\ \tilde{x}_j = \xi/x_j. \end{cases} \quad (57)$$

定数 C の逆数は、規格化条件 $R_\ell(\xi, 1) = 1$ を満たすように式 (58) で与えられる.

$$C^{-1} = \prod_{j=1}^{\lfloor \ell/2 \rfloor} \frac{1 - x_j^2}{1 - \tilde{x}_j^2}. \quad (58)$$

6.2 E-合成の場合の有理関数 $h(t)$

いま E-合成に用いる l 次の実有理関数 $h(t)$ を l 次の楕円有理関数を用いた式 (59) で定義する. ただし $l \geq 2$ であるとし, L と ξ は l 次の「次数方程式」を満たしているものとする.

$$h(t) \equiv \frac{L+1}{2} \times \frac{1+R_l(\xi, t)}{L+R_l(\xi, t)}. \quad (59)$$

条件 $h(\xi) = \mu$ を満たすことから, μ と L の関係を表す式 (60) が得られる.

$$\mu = \frac{(L+1)^2}{4L} \quad \text{あるいは} \quad L = (\sqrt{\mu} + \sqrt{\mu-1})^2. \quad (60)$$

6.2.1 E-合成の場合の有理関数 $h(t)$ の性質

式 (59) で表される関数 $h(t)$ は, 楕円有理関数 $R_l(\xi, t)$ の持つ性質から導かれる以下の 4 つの性質を持つ.

- (1) l が偶数のときは偶関数.
- (2) $t \in [0, 1]$ では等リプル性を持ち, $[0, 1]$ への全射.
- (3) $t \in (1, \xi)$ で単調増加で $(1, \mu)$ への全射.
- (4) $t \in [\xi, \infty]$ では振動して極を持ち, $[\mu, \infty]$ への全射で, すべての極小値は μ に一致.

すると $h(t)$ は合成用の有理関数として要請される性質 (3.1 節) を備えていることが分かる.

なお式 (59) の E-合成用の関数 $h(t)$ は, $l \geq 2$ であれば次数 l が偶数でも奇数でも, $t \in [0, 1]$ に対する値域と $t \in [-1, 0]$ に対する値域はどちらも $[0, 1]$ であり一致する. このことから, $l \geq 2$ である場合には, E-合成で得られるフィルタの通過域を $t \in [0, 1]$ から $t \in [-1, 1]$ に広げることが次数 l の偶奇によらずつねにできる.

6.3 E-合成の場合の伝達関数の設計方法の例

Chebyshev 多項式の次数 n と伝達率の閾値 g_p と g_s の 3 つを指定すると, 簡単な式 (43) の計算で σ と μ の値が決まる. すると 1 次分数関数 $x(t)$ は完全に決まる. そうして μ から L が式 (60) の左側の計算により決まる. あとは楕円有理関数の次数 l が決まれば「次数方程式」(49) を解いて L から ξ を求めれば E-合成の有理関数 $h(t)$ がさらに合成関数 $\hat{x}(t) = x(h(t))$ も完全に決まる.

次数 l の値を直接指定して決めるのであれば, それで $\hat{x}(t)$ の決定は完了する. ただし中間固有対用のフィルタを構成する場合には, 指定する l は偶数とする必要がある.

他方で, 次数 l の値を直接指定せずに, ξ の上限値 ξ_{\max} を指定する場合は, l を式 (61) で定義される実数 l_{\min} 以上の整数に選ぶと, その次数 l に対する「次数方程式」を解いて L から得られる ξ の値は必ず ξ_{\max} 以下となるので, それでもって $\hat{x}(t)$ の決定は完了する. ただし中間固有対用のフィルタを構成する場合には, l は l_{\min} 以上の「偶数」に選ぶ必要がある.

$$l_{\min} = \left\{ \frac{K'(L^{-1})}{K(L^{-1})} \right\} / \left\{ \frac{K'(\xi_{\max}^{-1})}{K(\xi_{\max}^{-1})} \right\}. \quad (61)$$

- 1) n, g_p, g_s を「指定する」.
- 2) σ と μ の値を式 (43) を用いて計算する.
- 3) $L \leftarrow (\sqrt{\mu} + \sqrt{\mu-1})^2$ とする.
- 4) l の決め方は, 最初から l の値を「指定する」, もしくは ξ の上限値 ξ_{\max} を「指定して」, 式 (61) で計算される実数 l_{\min} 以上の整数 (フィルタが中間固有対用ならば偶数) に l を設定.
- 5) $q(k)$ とその逆関数を用いて $\xi \leftarrow 1/q^{-1}(\{q(L^{-1})\}^{1/l})$.

図 10 次数 n と閾値 g_p と g_s および次数 l あるいは ξ_{\max} を与えて $\hat{x}(t) = x(h(t))$ を決定する手順 (E-合成の場合)

Fig. 10 Procedure to determine $\hat{x}(t) = x(h(t))$ from given degree n and both thresholds g_p and g_s and also either degree l or ξ_{\max} (for the case of E-composition).

以上の手順を図 10 に示す. この手順により, n, g_p, g_s の 3 つに加えて l または ξ_{\max} のどちらか一方を指定すれば, $\hat{x}(t)$ の表式が含む 5 つの値 σ, μ, L, l, ξ のすべてが確定するので, $\hat{x}(t)$ は完全に決まる.

6.4 E-合成の場合の $\hat{x}(t)$ の部分分数分解

E-合成の場合の実有理関数 $\hat{x}(t)$ の式は, 式 (13) と式 (59) の組合せである.

以下で, 実関数 $\hat{x}(t)$ の複素数の範囲における部分分数分解 (14) に現れる定数項の実数係数 c_∞ および複素数の極 t_j とその係数 $c_j, j=1, 2, \dots, l$ を求める方法を示す.

6.4.1 定数項の係数 c_∞ の計算

まず最初に定数項の係数 c_∞ を求める. それにはまず $\lim_{t \rightarrow \infty} R_l(\xi, t) = \lim_{t \rightarrow \infty} L/R_l(\xi, \xi/t) = L/R_l(\xi, 0)$ であるが, l が奇数の場合には $R_l(\xi, t)$ は奇関数なので $R_l(\xi, 0) = 0$ である. また l が偶数の場合にはさらに $l/2$ が偶数であれば $R_l(\xi, 0) = 1$ であり, $l/2$ が奇数であれば $R_l(\xi, 0) = -1$ であることも分かる.

すると $\lim_{t \rightarrow \infty} h(t)$ の値は, l が奇数の場合は $(L+1)/2$ であるが, l が偶数の場合にはさらに $l/2$ が偶数ならば μ になり, $l/2$ が奇数ならば ∞ になる. よって, $c_\infty = \lim_{t \rightarrow \infty} \hat{x}(t)$ の値は, l が奇数の場合は $(\mu + \sigma) / \{(L+1)/2 + \sigma\}$ であり, l が偶数の場合にはさらに $l/2$ が偶数ならば 1 であり, $l/2$ が奇数ならば 0 になる.

以上をまとめると, 定数項の係数 c_∞ の値は式 (62) とする.

$$c_\infty = \begin{cases} \frac{2(\mu + \sigma)}{L + (2\sigma + 1)} & (\text{mod}(\ell, 4) = 1, 3 \text{ の場合}), \\ 1 & (\text{mod}(\ell, 4) = 0 \text{ の場合}), \\ 0 & (\text{mod}(\ell, 4) = 2 \text{ の場合}). \end{cases} \quad (62)$$

6.4.2 極 t_j の計算

ここでは部分分数分解 (14) の極 t_j を複素数の範囲で求

める．極 t_j は $h(t) + \sigma = 0$ の解 t であり， $h(t)$ の式 (59) により，方程式 (63) の解になる．

$$R_\ell(\xi, t) = -\Gamma, \quad \Gamma \equiv 1 + \frac{2(L-1)\sigma}{L+(2\sigma+1)} > 1. \quad (63)$$

楕円有理関数 $R_\ell(\xi, t)$ の媒介変数表示 (47) を用いると，式 (64) が得られる（上側の式を満たす複素数 ϕ の値を求めて，それを下側の式中の ϕ にすると t の値が得られる）．ただし L と ξ は次数 ℓ に対する「次数方程式」(49) をすでに満たしているものとする．

$$\begin{cases} -\Gamma = \text{cd}\{\ell\phi K(L^{-1}), L^{-1}\}, \\ t = \text{cd}\{\phi K(\xi^{-1}), \xi^{-1}\}. \end{cases} \quad (64)$$

いま関数 cd の性質を用いることで (64) の上側の式を符号を変えて式の値が正となるように（引数を実数周期 $4K(L^{-1})$ の半分だけずらして）書き直せば式 (65) になる．

$$\text{cd}\{(\ell\phi + 2)K(L^{-1}), L^{-1}\} = \Gamma (> 1). \quad (65)$$

楕円関数 $\text{cd}(u, k)$ が変数 u について実数周期 $4K(k)$ を持つ周期関数であることから，式 (65) の左辺の式の値は任意の整数 j を用いて ϕ を $\phi + 4j/\ell$ に置き換えても変わらない．すると式 (65) を満たす ϕ の 1 つの値 ϕ_0 が得られたなら，それを元にして式 (66) により t_j の ℓ 個の値を生成すると，（生成された値 t_j がすべて異なれば）必要とする ℓ 個すべての解が得られる．

$$t_j = \text{cd}\{(\phi_0 + 4j/\ell)K(\xi^{-1}), \xi^{-1}\}, \quad j=1, 2, \dots, \ell. \quad (66)$$

Jacobi の楕円関数 $\text{cd}(u, k)$ は実数の引数 u に対する関数値は（実数で）絶対値が 1 以下になる性質を持つが，いま $\Gamma > 1$ であるから，式 (65) を満たす ϕ は実数ではないことが分かる．関数 $\text{cd}(u, k)$ は変数 u の偶関数なので，純虚数の引数に対する関数値は実数になる．そこでいま y を実数として，式 (65) の関数 cd の（第 1 の）引数が純虚数 $-y\sqrt{-1}$ であるとおいて式 (67) をたてる．

$$(\ell\phi + 2)K(L^{-1}) = -y\sqrt{-1}. \quad (67)$$

すると式 (65) から式 (68) が得られる．

$$\text{cd}(-y\sqrt{-1}, L^{-1}) = \Gamma. \quad (68)$$

Jacobi の楕円関数 $\text{cd}(u, k)$ に対しては虚数変換の公式 (69) [1], [10] がある．

$$\text{cd}(u\sqrt{-1}, k) = \frac{1}{\text{dn}(u, k')}. \quad (69)$$

この虚数変換の公式を $u = -y$, $k = L^{-1}$ として式 (68) にあてはめれば式 (70) が得られる．ここで $(L^{-1})'$ は母数 L^{-1} の補母数 $\sqrt{1-L^{-2}}$ を表す．

$$\text{dn}\{-y, (L^{-1})'\} = \Gamma^{-1}. \quad (70)$$

Jacobi の楕円関数の間の恒等関係式 $\text{dn}^2(u, k) = 1 - k^2\text{sn}^2(u, k)$ により， $\text{sn}^2(u, k) = (1/k^2)\{1 - \text{dn}^2(u, k)\}$ であることを用いると，式 (71) が得られる．ここで複号は任意だが以下ではマイナスの符号を選ぶ（そうすれば虚部が正である極 t_j の添字は前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ になる）．この式で関数 sn の絶対値は引数が実数なので 1 以下である．

$$\text{sn}\{-y, (L^{-1})'\} = \pm\sqrt{(1-\Gamma^{-2})/(1-L^{-2})}. \quad (71)$$

Jacobi の楕円関数 $\text{sn}(u, k)$ と第 1 種楕円積分 $F(\varphi, k)$ の間には， $\sin \varphi = \text{sn}(u, k)$ と $u = F(\varphi, k)$ が同値になる関係があるので，（さらに sn が奇関数であることも使うと）実数 y の値は式 (72) で表せる．

$$y = F(\sin^{-1} \sqrt{(1-\Gamma^{-2})/(1-L^{-2})}, (L^{-1})'). \quad (72)$$

ここで第 1 種楕円積分 $F(\varphi, k)$ の定義は式 (73) である．

$$\begin{aligned} F(\varphi, k) &\equiv \int_0^\varphi \frac{d\theta}{\sqrt{1-k^2\sin^2\theta}} \\ &= \int_0^{\sin\varphi} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}. \end{aligned} \quad (73)$$

さらに本論文の中だけでの記法として，関数 $F'(\varphi, k) \equiv F(\varphi, k')$ を定義する．ただし k' は通常のように母数 k の補母数 $k' \equiv \sqrt{1-k^2}$ を表している．算術幾何平均を利用して第 1 種楕円積分 $F(\varphi, k)$ の値を求める算法 [4] をわずかに変更することにより，母数 k の値が小さい場合にも $F'(\varphi, k)$ の値を精度良く求められる計算法が得られる．いまのように $L^{-1} \ll 1$ である場合には， $F(\varphi, (L^{-1})')$ の数値を求めるために，通常の $F(u, k)$ の計算ルーチンに第 2 引数として $(L^{-1})' = \sqrt{1-L^{-2}}$ の数値を与えるよりも，2 つの引数 φ と k に対して $F'(\varphi, k)$ を計算する特別に用意されたルーチンに第 2 引数として L^{-1} の数値を与えた方が，計算結果の精度が良くなる．

求めた実数 y の値を用いて式 (67) を解くと，式 (74) で表される ϕ の 1 つの解 ϕ_0 が得られる．

$$\phi_0 = \frac{1}{\ell} \left\{ -2 - \frac{y}{K(L^{-1})} \sqrt{-1} \right\}. \quad (74)$$

そうして式 (66) を用いると，部分分数分解 (14) に現れるすべての極 t_j は複素引数の楕円関数 cd を用いて，あるいは引数をずらすことで sn を用いて，式 (75) により表せる．

$$\begin{cases} \omega_j = \frac{1}{\ell} \left\{ 4j - 2 - \frac{y}{K(L^{-1})} \sqrt{-1} \right\}, \\ t_j = \text{cd}\{\omega_j K(\xi^{-1}), \xi^{-1}\} \\ = -\text{sn}\{(\omega_j - 1)K(\xi^{-1}), \xi^{-1}\}. \end{cases} \quad (75)$$

上記の式中で引数が複素数 $u + v\sqrt{-1}$ である Jacobi の

楕円関数 sn の値は、恒等式 (76) を用いて引数が実数の楕円関数の値を組み合わせる計算する。

$$\begin{aligned} & \text{sn}(u + v\sqrt{-1}, k) \\ &= \frac{\text{sn}(u, k) \text{dn}(v, k') + \text{cn}(u, k) \text{dn}(u, k) \text{sn}(v, k') \text{cn}(v, k')\sqrt{-1}}{\text{cn}(v, k')^2 + k^2 \text{sn}(u, k)^2 \text{sn}(v, k')^2}. \end{aligned} \quad (76)$$

引数が実数である場合の 3 つの楕円関数の値 $\text{sn}(u, k)$, $\text{cn}(u, k)$, $\text{dn}(u, k)$ は相加相乗平均に基づく方法により 3 つをまとめて計算できる (文献 [1], [10])。同様に $\text{sn}(v, k')$, $\text{cn}(v, k')$, $\text{dn}(v, k')$ の 3 つの値もまとめて計算できる。

式 (75) で楕円関数 sn の引数である複素数 $(\omega_j - 1)K(\xi^{-1})$ が虚部に添字 j を含まないことを用いると、複素数を引数とする楕円関数の値の計算にかかる手間を少なくできる。

式 (75) で表される複素数の極 t_j は、添字が $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ のものは (式 (71) の複号でマイナスの符号を選択したことにより) 虚部が正であり、 ℓ が奇数の場合に限りただ 1 つ存在する実数の極は添字が $j = (\ell + 1)/2$ のものであり値は -1 よりも小さい。

以上をまとめると、極を求める計算は式 (77) になる。まず $\sin \varphi$ あるいは $\cos \varphi$ の値を (Γ は使わずに) σ と L だけから計算して、それから逆三角関数を用いて第一象限の角である φ の値を求める。

逆三角関数は引数の絶対値が 1 に近づくときに変化率が大きくなり値を精度良く求めることが難しくなるので、なるべくそのような状況を避けるようにすべきであるが、実は式 (77) の中の $\sin \varphi$ と $\cos \varphi$ を表す式はどちらも 1 にきわめて近い値をとることがないことを式を用いて示せる。よって φ の値を求めるにはどちらを用いても良いが、 $\hat{x}(t)$ の計算に対して部分分数分解前の式 (13) を用いた場合と、部分分数分解後の式 (14) を用いた場合とで合成された伝達関数 $\hat{g}(t) = g_s T_n(2\hat{x}(t) - 1)$ の値の差の大きさを多数の分点 t において求めて比較してみたところ、 φ の値を $\sin \varphi$ の式の値から求めるよりも $\cos \varphi$ の式の値から求める方が比較の結果がごくわずかであるが良くなる傾向が感じられた。ただしその違いはあまりはっきりしたものではありません。なお φ の値を $\sin \varphi$ と $\cos \varphi$ の小さい方の値から求めるようにすることもできる。

次に実数 $y (> 0)$ の値を、第 2 引数が母数 k の補母数 k' である第 1 種楕円積分 $F(u, k')$ を求めるための専用ルーチン $F'(u, k)$ を用いて、 $y = F'(\varphi, L^{-1})$ を計算して実数 y を求める。そうして添字 j に対する複素数の値 ω_j を計算し、それを用いて $t_j = -\text{sn}\{(\omega_j - 1)K(\xi^{-1}), \xi^{-1}\}$ により複素数を引数とする楕円関数 sn の値を計算すれば複素数の極 t_j が求まる。この式で正の虚部を持つ極に対する添字は、前半のもの $j=1, 2, \dots, \lfloor \ell/2 \rfloor$ である。実数の極は ℓ が奇数

のときに限りただ 1 つ存在してその添字は $j = (\ell + 1)/2$ であり、値 $t_R = t_{(\ell+1)/2}$ は $t_R < -1$ を必ず満たす。

$$\left\{ \begin{aligned} \sin \varphi &= \sqrt{\frac{1 - \Gamma^{-2}}{1 - L^{-2}}} = \frac{2L\sqrt{\sigma(\sigma + 1)}}{(2\sigma + 1)L + 1}, \\ \cos \varphi &= \sqrt{\frac{\Gamma^{-2} - L^{-2}}{1 - L^{-2}}} = \frac{\sqrt{L^2 + 2(2\sigma + 1)L + 1}}{(2\sigma + 1)L + 1}, \\ y &= F(\varphi, (L^{-1})') = F'(\varphi, L^{-1}), \\ \omega_j &= \frac{1}{\ell} \left\{ 4j - 2 - \frac{y}{K(L^{-1})} \sqrt{-1} \right\}, \\ t_j &= -\text{sn}\{(\omega_j - 1)K(\xi^{-1}), \xi^{-1}\}. \end{aligned} \right. \quad (77)$$

6.4.3 極の係数 c_j の計算

楕円有理関数の明示的な有理関数表示 (56) とその零点 x_i と極 \tilde{x}_i の値の明示式 (57) から楕円有理関数 $R(\xi, t)$ の対数微分 $\Psi(t) \equiv \frac{d}{dt} \log R_\ell(\xi, t)$ (これは $\{\frac{d}{dt} R_\ell(\xi, t)\} / R_\ell(\xi, t)$ に等しい) を計算するのに便利な式 (78) が導かれる。それを用いて複素数 t に対する $\Psi(t)$ の値が計算できる。

$$\Psi(t) = \frac{\text{mod}(\ell, 2)}{t} + 2t \sum_{i=1}^{\lfloor \ell/2 \rfloor} \left(\frac{1}{t^2 - x_i^2} - \frac{1}{t^2 - \tilde{x}_i^2} \right). \quad (78)$$

いま式 (59) を用いると、合成された有理関数 $\hat{x}(t)$ は式 (79) のように変形できる。

$$\hat{x}(t) = \frac{\mu + \sigma}{h(t) + \sigma} = \frac{2(\mu + \sigma)}{L + (2\sigma + 1)} \times \frac{L + R_\ell(\xi, t)}{R_\ell(\xi, t) + \Gamma}. \quad (79)$$

これを用いて部分分数分解の式 (14) における極の係数 c_j を式 (80) のように計算する。途中の式変形では、 $R_\ell(\xi, t)$ の t についての導関数を $R'_\ell(\xi, t)$ と表しており、また $R_\ell(\xi, t_j) = -\Gamma$ であることや、 $\frac{L - \Gamma}{\Gamma} = \frac{L^2 - 1}{(2\sigma + 1)L + 1}$ となることを用いている。

$$\begin{aligned} c_j &= \lim_{t \rightarrow t_j} (t - t_j) \hat{x}(t) \\ &= \frac{2(\mu + \sigma)(L - \Gamma)}{L + (2\sigma + 1)} \times \lim_{t \rightarrow t_j} \frac{t - t_j}{R_\ell(\xi, t) + \Gamma} \\ &= \frac{2(\mu + \sigma)(L - \Gamma)}{L + (2\sigma + 1)} \times \frac{1}{R'_\ell(\xi, t_j)} \\ &= -\frac{2(\mu + \sigma)}{L + (2\sigma + 1)} \times \frac{L - \Gamma}{\Gamma} \times \frac{1}{R'_\ell(\xi, t_j) / R_\ell(\xi, t_j)} \\ &= -\frac{2(\mu + \sigma)}{L + (2\sigma + 1)} \times \frac{L^2 - 1}{(2\sigma + 1)L + 1} \times \frac{1}{\Psi(t_j)}. \end{aligned} \quad (80)$$

まとめると、E-合成の場合の $\hat{x}(t)$ の部分分数分解の式 (14) における極 t_j に対する係数 c_j , $j=1, 2, \dots, \ell$ は式 (81) を計算することで求められる。

$$c_j \leftarrow -\frac{2(\mu + \sigma)(L^2 - 1)}{\{L + (2\sigma + 1)\}\{(2\sigma + 1)L + 1\}} \times \frac{1}{\Psi(t_j)}. \quad (81)$$

7. 実験について

7.1 使用した計算機システム環境

実験に使用した計算機は、東京大学情報基盤センターに設置されている Oakbridge-CX の 1 ノード (富士通 PRIMERGY CX2550M5) である。そのノードは密結合の 2 つの CPU から構成されている。ノード内の主記憶は DDR4 で容量は 192 GiB (ユーザが利用できるのは 168 GiB) であり、メモリのバンド幅は最大 281.6 GiB/s である。CPU は Intel Xeon Platinum 8280 で、1 つの CPU は 28 コア、クロック 2.7 GHz、L2 キャッシュ 1 MB/コア、L3 キャッシュ 38.5 MiB/CPU、CPU のメモリチャネル数は 6、Intel AVX-512 拡張命令セットで、AVX-512 FMA ユニット数は 2。OS は CentOS 7 for x86_64 (64 bit 版) である。Turbo-Boost 機能 (最大周波数 4.0 GHz) は on になっており (% cpupower frequency-info の出力から確認できる)、また Hyper-Thread 機能は off になっている (% cat /proc/cpuinfo | grep "physical id" の出力で、各物理 CPU に対して仮想 cpu の数が 28 になっていることから確認できる)。

プログラムは Fortran 90 言語に OpenMP 指示行を入れてコーディングをし、コンパイラは Intel Fortran v19.0.4.243 for x86_64 で、コンパイラのオプションに "-fast -openmp -axCORE-AVX512" を指定して OpenMP により 2 CPU の合計 56 スレッド並列で実行した。

本論文の計算に用いた数値と演算はすべて IEEE 754 の倍精度浮動小数点数 (2 進 64 bit) であり、複素数も実部と虚部がそれぞれ倍精度浮動小数点数のものである。

7.2 計算の主要部

計算の主要部は、行列 X と Y を実あるいは複素の N 次の縦ベクトル m 個を列方向に並べた $N \times m$ 行列とすると、与えられた X に対してレゾルベント $\mathcal{R}(\rho_j)$ を適用した結果である $Y \leftarrow \mathcal{R}(\rho_j)X$ を求めるところである。今回のフィルタは $\hat{\mathcal{F}} = g_s T_n(\hat{Y})$ であり、 $\hat{Y} \equiv 2\hat{\mathcal{X}} - I$ は以下の式 (82) により表される実線形作用素 (実ベクトルに対して実ベクトルを与える線形作用素) である。

$$\hat{Y} = \begin{cases} (2c_\infty - 1)I + \sum_{j=1}^{\ell/2} \text{Re}\{4\gamma_j \mathcal{R}(\rho_j)\} & (\ell \text{ が偶数}), \\ (2c_\infty - 1)I + 2\gamma_R \mathcal{R}(\rho_R) + \sum_{j=1}^{(\ell-1)/2} \text{Re}\{4\gamma_j \mathcal{R}(\rho_j)\} & (\ell \text{ が奇数}). \end{cases} \quad (82)$$

A と B を N 次の実対称行列としてシフトが ρ_j のレゾルベントは $\mathcal{R}(\rho_j) \equiv (A - \rho_j B)^{-1} B$ なので、 $C^{(j)} \equiv A - \rho_j B$ とすると、 $Y^{(j)} \leftarrow \mathcal{R}(\rho_j)X$ の計算は、各 j について右辺

が共通な値 BX を持つ連立 1 次方程式 $C^{(j)}Y^{(j)} = BX$ を $Y^{(j)}$ について解くことに帰着する。行列 $C^{(j)}$ は対称で、シフト ρ_j が実数であるか虚数であるかによって実行列あるいは複素行列になるが、さらに以下の例題のように A と B が帯行列であれば $C^{(j)}$ も帯行列となる。

今回の実験に用いたプログラムは、簡単化のために係数が実あるいは複素の帯対称行列である連立 1 次方程式を帯の内部には実際には非常に疎であるが、帯の内部を密であるかのようにして扱い、領域分割法やブロック化手法などは用いずに、古典的な BLAS-2 レベルの帯用の改訂コレスキ法で解いている。それを Fortran90 で記述したものに OpenMP の指示を挿入している。これは現段階での研究の主な目的は、フィルタ対角化法の手法の数値的な性質の検証であって、まだ使用する計算機システムの性能を高度に絞り出す段階ではないからである。今回の各実験結果に対して掲げた経過時間の測定値はそのような性格のものである。各種の計算機システム上で帯対称行列の連立 1 次方程式を効率良く解く方法の研究はまた別の課題になる。

連立 1 次方程式 $C^{(j)}Y^{(j)} = BX$ を解くには、各 j について共通の右辺ベクトルの組 BX を作るために、まず X に帯行列 B を乗じる処理 $Z \leftarrow BX$ を行う。そうして対称帯行列用の改訂コレスキ法による行列分解 $C^{(j)} \Rightarrow L^{(j)}D^{(j)}(L^{(j)})^T$ の処理と、その分解結果を用いて前進代入と後退代入を行って $Y^{(j)}$ を求める処理 $Y^{(j)} \leftarrow (L^{(j)})^{-T}\{(D^{(j)})^{-1}(L^{(j)})^{-1}Z\}$ が要る。複素版の改訂コレスキ法は通常の実数版の算法中の数値と四則演算をそのまま複素数のものに置き換えることで得られる。改訂コレスキ分解の結果は残しておいて繰り返し利用する。

注 1

シフト ρ_j が実数で $C^{(j)}$ が実対称正定値の場合には、ピボット選択をしなくても (改訂) コレスキ法の計算は数値安定であるが、シフト ρ_j が複素数で $C^{(j)}$ が複素対称である場合にはそのような保証はない。そのため $C^{(j)}$ が複素対称である場合にはその対称性を利用しないで、片側ピボット選択を入れた帯行列用の LU 分解法を用いれば数値安定性は保証される。しかしそのようにすると LU 分解の記憶量はピボット選択なしの (改訂) コレスキ法を用いる場合に比べて (行選択により上帯幅は 2 倍にまで広がる可能性がある) 3 倍になり、計算に用いる演算量もかなり増えるので、今回の実験例では係数行列 $C^{(j)}$ が複素対称の場合にはピボット選択なしの帯行列用の (複素版の) 改訂コレスキ法を用いている。

注 2

今回の実験ではまず j について順次に行列 $C^{(j)}$ の (改訂) コレスキ分解を作っておく。それから j について順次に連立 1 次方程式 $C^{(j)}Y^{(j)} = Z$ を解くにはすでに作っておいた行列 $C^{(j)}$ の分解結果を利用して前進代入と後退代入を行って解 $Y^{(j)}$ を求める。前進代入と後退代入の処理

の中では、 m 個のベクトルの組をまとめて計算することにより、マルチコア計算機の主記憶（共有メモリ）に置かれた帯行列 $C^{(j)}$ の分解で得た帯行列 $L^{(j)}$ に対する記憶参照の量を節約している。

7.3 例題として用いた一般固有値問題について

実験の例題に用いた行列の実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ は、各辺が座標軸に沿った一辺の長さ π の 3 次元立方体 $[0, \pi] \times [0, \pi] \times [0, \pi]$ の内部を領域とし、その境界である表面において零-ディリクレ境界条件を課したときの（符号が逆の）3次元ラプラシアン $-\Delta$ の固有値問題、それを有限要素法により離散化近似して得られるものである。

用いた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、要素内での展開基底関数には各辺方向の 3 重線形関数を用いた。この有限要素法の離散化で得られる行列 A と B の次数は $N = N_1 N_2 N_3$ であり、($N_1 \leq N_2 \leq N_3$ であるとして) 行列の帯幅がなるべく小さくな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 $w_L = 1 + N_1 + N_1 N_2$ にできる。

今の場合に、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 ($1 \leq i_k \leq N_k, k = 1, 2, 3$)。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 $i_1 + N_1(i_2 - 1) + N_1 N_2(i_3 - 1)$ 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。

そうして、固有値 λ が区間 $[a, b]$ に含まれる固有対の近似をフィルタ対角化法を適用して求めた。このテスト例題の一般固有値問題の固有値は簡単な数式で表せるので、厳密な値を式計算で容易に求めることができる。またそのことを用いて、固有値の厳密値を列挙して値の順に並べることで、固有値が区間 $[a, b]$ にある固有対の正しい数も求められる。

7.3.1 例題の固有値の厳密値を与える式

1次元の区間 $[0, \pi]$ における零-ディリクレ境界条件のラプラシアン（ただし符号を逆にした $-\Delta$ ）を有限要素法で離散化して得られる「行列の一般固有値問題」の固有値の全体は、区間 $[0, \pi]$ を $N + 1$ 等分して得られる幅 $h = \pi / (N + 1)$ の小区間による分割の区分線形関数を基底とする場合には、 $\phi_k \equiv hk$ とおくとときに以下の式 (83) で与えられる。

$$E(N, k) = \frac{6k^2 \{\sin(\phi_k) / \phi_k\}^2}{(1 + \cos \phi_k)(2 + \cos \phi_k)}, \quad k=1, 2, \dots, N. \quad (83)$$

そうして 3 次元立方体領域 $[0, \pi] \times [0, \pi] \times [0, \pi]$ の各座標軸方向の区間をそれぞれ $N_1 + 1, N_2 + 1, N_3 + 1$ に等分割して各方向の 1 次元の基底の直積型の基底関数（3 重線形関数）を用いた場合の有限要素法では、 $k_i = 1, 2, \dots, N_i$ とするとき (k_1, k_2, k_3) で識別される 3 次元ラプラシアン ($-\Delta$) の固有値 $E((N_1, N_2, N_3), (k_1, k_2, k_3))$ の値は 1 次元の場合の固有値の和の形で、式 $E(N_1, k_1) + E(N_2, k_2) + E(N_3, k_3)$ により与えられる。

7.4 フィルタ対角化法の概要

フィルタ対角化法では、まず最初に m 個のランダムなベクトルの組 ($N \times m$ 行列) を乱数で生成して、それらを B -正規直交化して m 個のベクトルの組 X を作る。そうして区間 $[a, b]$ を通過域とするフィルタを X に作用させて m 個のベクトルの組 $Y \leftarrow \mathcal{F}X$ を作る。

この X と Y をもとにして、使用したフィルタの伝達特性も考慮に入れて、元の一般固有値問題のある「不変部分空間」を近似する空間の基底を Y の列の適切な線形結合として構成する [16]。その「不変部分空間」は、区間 $[a, b]$ のある適切な近傍に入る固有値を持つ固有ベクトルの全体で張られるものである。その「不変部分空間」の近似空間の基底に Rayleigh-Ritz 法を適用して得られた Ritz 対を元の一般固有値問題の近似固有対として採用する。得られた近似固有対については（近似固有対に改良を加える場合はその後）、近似固有値を検査して指定区間 $[a, b]$ にないものについては棄却することができる。

7.5 近似固有対の品質評価に用いた相対残差について

計算で求めた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 (λ, \mathbf{v}) に対する相対残差 Θ の定義を式 (84) とする。ただしベクトルのノルム $\|\cdot\|$ には 2-ノルムを用いた。

$$\Theta \equiv \frac{\|A\mathbf{v} - \lambda B\mathbf{v}\|}{\|\lambda B\mathbf{v}\|}. \quad (84)$$

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 A と B を共通の非零の値でスケールしても変わらない。そうしてベクトルのノルムに 2-ノルムを用いたことにより、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 (85) が成り立つことが示せる。

$$\sin \phi \leq \Theta. \quad (85)$$

複数の近似固有対のベクトルを列として集めた行列を V とし、それから行列 A と B の対称性や帯性や疎性を利用して行列積 AV と BV を効率良く計算することにより、複数の近似固有対の相対残差をまとめて一度に計算することができる。いま第 k 番目の近似固有対 $(\lambda^{(k)}, \mathbf{v}^{(k)})$ のべ

クトル $\mathbf{v}^{(k)}$ を第 k 列とする行列を V とする. そうしていま $(AV)^{[k]}$ と $(BV)^{[k]}$ はそれぞれ行列積の結果である行列 AV と BV の第 k 列のベクトルを表すものとする. そのとき, 第 k 番目の近似固有対の相対残差 $\Theta^{(k)}$ は式 (86) で与えられることは容易に分かる.

$$\Theta^{(k)} = \frac{\| (AV)^{[k]} - \lambda^{(k)}(BV)^{[k]} \|}{\| \lambda^{(k)}(BV)^{[k]} \|}. \quad (86)$$

このように複数の近似固有対の相対残差を一度にまとめて計算すると, 行列 A と B の記憶に対する参照は行列積を作るところでそれぞれ 1 回ずつにまとめられるので, 近似固有対の相対残差を 1 つずつ独立に計算するのに比べて行列 A と B への記憶参照の量が大幅に減るため, 効率良く計算ができる.

8. 実験例

以下の例題 1~5 はいずれも FEM の要素分割が $(N_1, N_2, N_3) = (50, 60, 70)$ の場合で, 行列の次数と下帯幅は $N = 210,000$, $w_L = 3,051$ である.

8.1 例題 1 (中間固有対)

例題 1 では区間 $[a, b] = [1020, 1025]$ に固有値がある中間固有対を求めてみる. 区間 $[a, b] = [1020, 1025]$ に固有値がある真の固有対の数は 73 個である.

フィルタの設定 (例題 1)

フィルタの伝達関数の 3 つの形状に対する条件として, 4 種類の合成 (B, C, I, E) のどの場合にも $g_p = 0.1$, $g_s \leq 10^{-16}$, $\xi = 1.1$ を課した.

中間固有対を求める場合には合成用の有理関数の次数 ℓ は偶数でなければならないので, 形状の要求を満たせる最小の偶数とした. なおチェビシェフ多項式の次数 n は 50 以下に制限している. そうして 4 種類の合成すべての場合について, 規格化座標 t の区間 $[-1, 1]$ を固有値の区間 $[a, b]$ に対応させた.

各合成によるフィルタのパラメータを表 2 に示す. E-合成によるフィルタを構成するレゾルベントはシフトが虚数のもの 3 つである.

各合成によるフィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ を図 11 に示す.

計算結果 (例題 1)

フィルタを適用するベクトルの数は $m = 100$ とした. 各

表 2 例題 1: フィルタの構成に用いたパラメータ

Table 2 Example-1: parameters used to construct filter.

| フィルタ | ℓ | n | 実現した g_s の値 |
|------|--------|-----|---------------|
| B-合成 | 24 | 36 | 9.18E-17 |
| C-合成 | 8 | 48 | 9.57E-17 |
| I-合成 | 8 | 48 | 9.57E-17 |
| E-合成 | 6 | 10 | 1.45E-17 |

合成によるフィルタを用いて得られた近似固有対の相対残差の最大値および近似固有対を得るまでの経過時間を表 3 に示す. 表から経過時間が最も短いのは E-合成の場合であることが分かる.

各合成によるフィルタを用いて得られた近似固有対の相対残差のグラフを図 12 に示す.

8.2 例題 2 (中間固有対)

例題 2 では区間 $[a, b] = [100, 200]$ に固有値がある中間固有

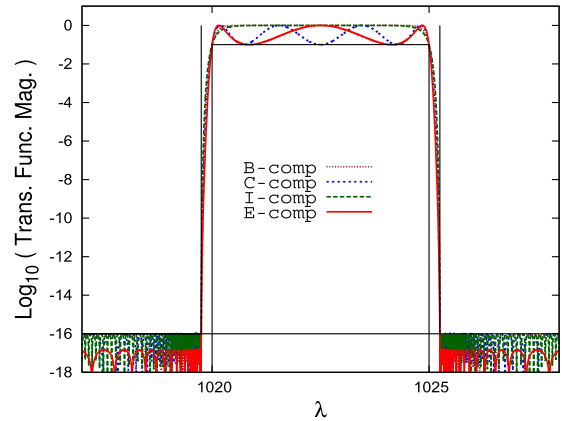


図 11 例題 1: フィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ の対数

Fig. 11 Example-1: log-plots of filter transfer function magnitudes $|\hat{f}(\lambda)|$ for compositions.

表 3 例題 1: 各合成による相対残差の最大値と経過時間

Table 3 Example-1: max relative residuals of approximate eigenpairs and elapsed times (for compositions).

| フィルタ | 相対残差の最大値 | 経過時間 (s) |
|------|----------|----------|
| B-合成 | 1.90E-13 | 3,127 |
| C-合成 | 4.72E-13 | 1,334 |
| I-合成 | 4.97E-13 | 1,325 |
| E-合成 | 4.48E-13 | 393 |

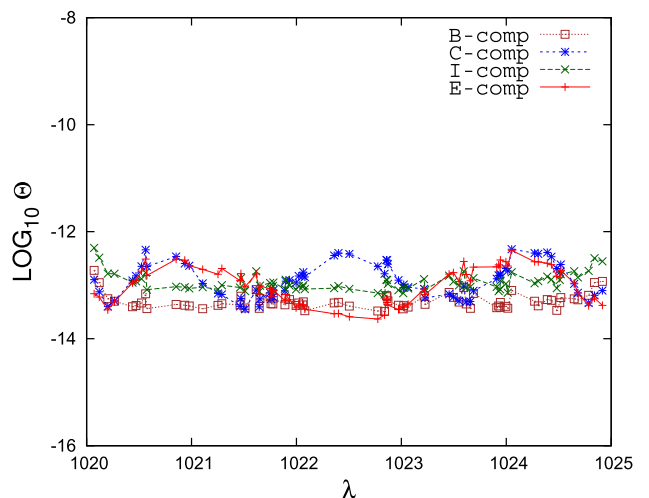


図 12 例題 1: 各合成による近似固有対の相対残差 Θ の対数

Fig. 12 Example-1: log-plots of relative residuals Θ of approximate eigenpairs (for compositions).

表 4 例題 2: フィルタの構成に用いたパラメータ

Table 4 Example-2: parameters used to construct filter for each composition.

| フィルタ | ℓ | n | 実現した g_p の値 |
|------|--------|-----|---------------|
| B-合成 | 24 | 36 | 0.1007 |
| C-合成 | 8 | 48 | 0.1003 |
| I-合成 | 8 | 48 | 0.1003 |
| E-合成 | 6 | 10 | 0.1444 |

表 5 例題 2: 各合成による相対残差の最大値と経過時間

Table 5 Example-2: max relative residuals of approximate eigenpairs and elapsed times (for compositions).

| フィルタ | 相対残差の最大値 | 経過時間 (s) |
|------|----------|----------|
| B-合成 | 4.55E-12 | 11,350 |
| C-合成 | 1.84E-12 | 5,204 |
| I-合成 | 1.39E-12 | 5,229 |
| E-合成 | 1.47E-11 | 1,060 |

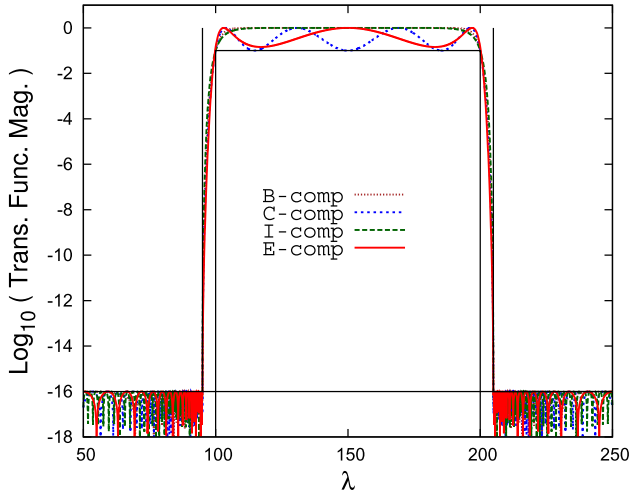


図 13 例題 2: フィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ の対数

Fig. 13 Example-2: log-plots of filter transfer function magnitudes $|\hat{f}(\lambda)|$ for compositions.

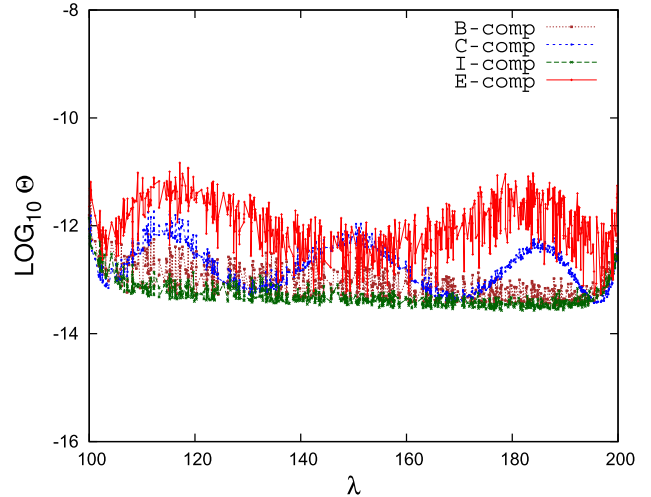


図 14 例題 2: 各合成による近似固有対の相対残差 Θ の対数

Fig. 14 Example-2: log-plots of relative residuals Θ of approximate eigenpairs (for compositions).

有対を求めてみる. 区間 $[a, b] = [100, 200]$ に固有値がある真の固有対の数は 801 個である.

フィルタの設定 (例題 2)

フィルタの伝達関数の 3 つの形状に対する条件として, 4 種類の合成 (B, C, I, E) のどの場合にも $g_p \geq 0.1$, $g_s = 10^{-16}$, $\xi = 1.1$ を課した.

中間固有対を求める場合には合成用の有理関数の次数 ℓ は偶数でなければならないので, 形状の要求を満たせる最小の偶数とした. なおチェビシェフ多項式の次数 n は 50 以下に制限している. そうして 4 種類の合成すべての場合について, 規格化座標 t の区間 $[-1, 1]$ を固有値の区間 $[a, b]$ に対応させた.

各合成によるフィルタのパラメータを表 4 に示す. E-合成によるフィルタを構成するレゾルベントはシフトが虚数のもの 3 つである.

各合成の場合のフィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ を図 13 に示す.

計算結果 (例題 2)

フィルタを適用するベクトルの数は $m = 1,000$ とした. 各合成によるフィルタを用いて得られた近似固有対の相対残差の最大値および近似固有対を得るまでの経過時間を表 5 に示す. 表から経過時間が最も短いのは E-合成の場合であることが分かる.

各合成によるフィルタを用いて得られた近似固有対の相対残差のグラフを図 14 に示す.

8.3 例題 3 (中間固有対)

例題 3 では区間 $[a, b] = [70, 80]$ に固有値がある中間固有対を求めてみる. 区間 $[a, b] = [70, 80]$ に固有値がある真の固有対の数は 54 個である.

フィルタの設定 (例題 3)

フィルタの伝達関数の 3 つの形状に対する条件として, 4 種類の合成 (B, C, I, E) のどの場合にも $g_p = 0.1$, $g_s \leq 10^{-16}$, $\xi = 1.3$ を課した.

中間固有対を求める場合には次数 ℓ は偶数でなければならないので, 形状の要求を満たせる最小の偶数を選んだ. なおチェビシェフ多項式の次数 n は 50 以下に制限している. そうして 4 種類の合成すべての場合について, 規格化座標 t の区間 $[-1, 1]$ を固有値の区間 $[a, b]$ に対応させた.

各合成によるフィルタのパラメータを表 6 に示す. E-合成によるフィルタを構成するレゾルベントはシフトが虚数のもの 2 つである.

各合成によるフィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ を図 15 に示す.

計算結果 (例題 3)

フィルタを適用するベクトルの数は $m = 100$ とした. 各

表 6 例題 3: フィルタの構成に用いたパラメータ

Table 6 Example-3: parameters used to construct filter for each composition.

| フィルタ | ℓ | n | 実現した g_s の値 |
|------|--------|-----|---------------|
| B-合成 | 10 | 20 | 6.97E-17 |
| C-合成 | 6 | 13 | 8.35E-17 |
| I-合成 | 6 | 13 | 8.35E-17 |
| E-合成 | 4 | 15 | 2.40E-17 |

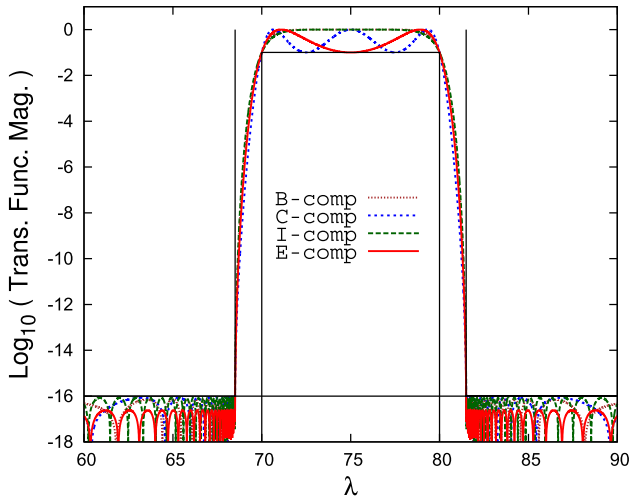


図 15 例題 3: フィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ の対数

Fig. 15 Example-3: log-plots of filter transfer function magnitudes $|\hat{f}(\lambda)|$ for compositions.

表 7 例題 3: 各合成による相対残差の最大値と経過時間

Table 7 Example-3: max relative residuals of approximate eigenpairs and elapsed times (for compositions).

| フィルタ | 相対残差の最大値 | 経過時間 (s) |
|------|----------|----------|
| B-合成 | 3.09E-13 | 901 |
| C-合成 | 1.10E-12 | 438 |
| I-合成 | 4.13E-13 | 439 |
| E-合成 | 3.30E-13 | 321 |

合成によるフィルタを用いて得られた近似固有対の相対残差の最大値および近似固有対を得るまでの経過時間を表 7 に示す. 表から経過時間が最も短いのは E-合成の場合であることが分かる.

各合成によるフィルタを用いて得られた近似固有対の相対残差のグラフを図 16 に示す.

8.4 例題 4 (下端固有対)

例題 4 では区間 $[a, b] = [0, 30]$ に固有値がある下端固有対を求めてみる. 区間 $[a, b] = [0, 30]$ に固有値がある真の固有対の数は 54 個である.

フィルタの設定 (例題 4)

フィルタの伝達関数の 3 つの形状に対する条件として, 4 種類の合成 (B, C, I, E) のどの場合にも $g_p \geq 0.1$, $g_s = 10^{-16}$, $\xi = 1.1$ を課した.

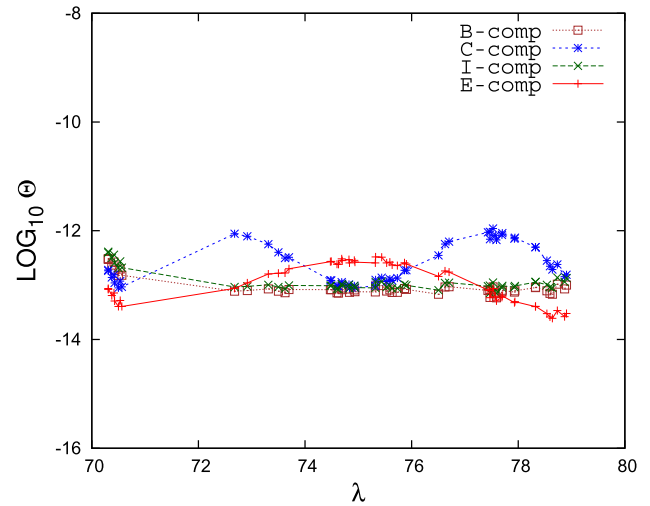


図 16 例題 3: 各合成による近似固有対の相対残差 Θ の対数

Fig. 16 Example-3: log-plots of relative residuals Θ of approximate eigenpairs (for compositions).

表 8 例題 4: フィルタの構成に用いたパラメータ

Table 8 Example-4: parameters used to construct filter for each composition.

| フィルタ | ℓ | n | 実現した g_p の値 |
|------|--------|-----|---------------|
| B-合成 | 24 | 36 | 0.1007 |
| C-合成 | 8 | 48 | 0.1003 |
| I-合成 | 8 | 48 | 0.1003 |
| E-合成 | 5 | 17 | 0.1131 |

合成用の有理関数の次数 ℓ は形状の要求を満たせる最小の整数とした. なおチェビシェフ多項式の次数 n は 50 以下に制限している. B-合成と I-合成は, どちらも次数 ℓ が偶数になったので, 中間固有値を求めることもできるフィルタであり, 規格化座標の区間 $t \in [-1, 1]$ を固有値の区間 $[a, b]$ に対応させた. C-合成と E-合成については, ℓ の偶奇によらずに規格化座標 t の区間 $[-1, 1]$ を固有値の区間 $[a, b]$ に対応させることが可能なのでそのようにした.

各合成のフィルタのパラメータを表 8 に示す. ここでは次数 ℓ が最小になるフィルタを構成する. 下端用のフィルタの次数 ℓ は奇数でも良いので, E-合成では実験例 2 の場合の $\ell = 6$ とは異なり $\ell = 5$ であり, それにともない n は 10 から 17 に増えている. E-合成のフィルタを構成するレゾルベントは, シフトが虚数のもの 2 つと実数のもの 1 つである.

各合成の場合のフィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ を図 17 に示す.

計算結果 (例題 4)

フィルタを適用するベクトルの数は $m = 80$ とした. 各合成によるフィルタを用いて得られた近似固有対の相対残差の最大値および近似固有対を得るまでの経過時間を表 9 に示す. 表から経過時間が最も短いのは E-合成の場合であることが分かる.

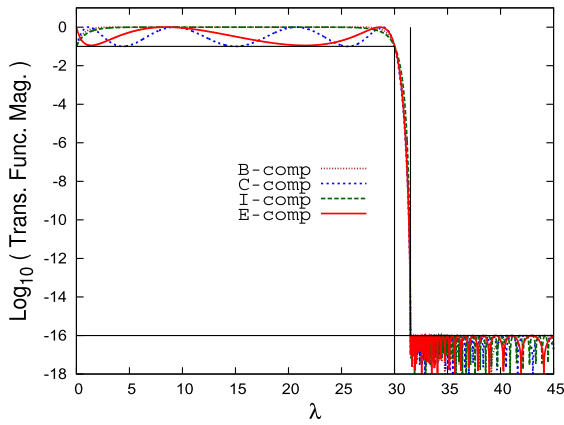


図 17 例題 4：フィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ の対数

Fig. 17 Example-4: log-plots of filter transfer function magnitudes $|\hat{f}(\lambda)|$ for compositions.

表 9 例題 4：各合成による相対残差の最大値と経過時間

Table 9 Example-4: max relative residuals of approximate eigenpairs and elapsed times (for compositions).

| フィルタ | 相対残差の最大値 | 経過時間 (s) |
|------|----------|----------|
| B-合成 | 5.14E-12 | 3,103 |
| C-合成 | 2.75E-11 | 1,308 |
| I-合成 | 4.35E-12 | 1,321 |
| E-合成 | 2.19E-11 | 406 |

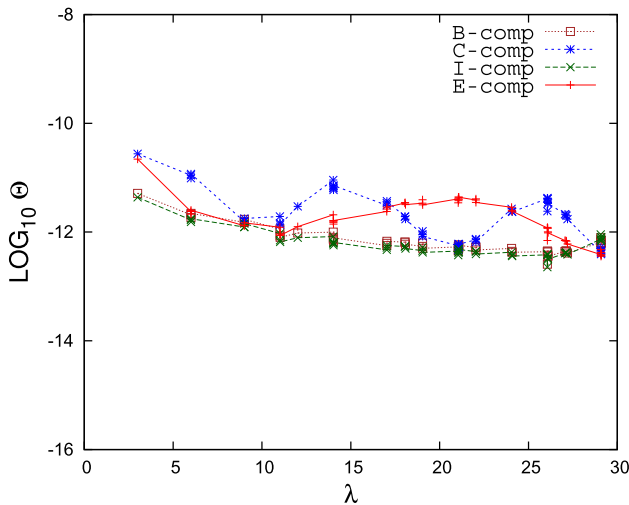


図 18 例題 4：各合成による近似固有対の相対残差 Θ の対数

Fig. 18 Example-4: log-plots of relative residuals Θ of approximate eigenpairs (for compositions).

各合成によるフィルタを用いて得られた近似固有対の相対残差のグラフを図 18 に示す。

8.5 例題 5 (下端固有対)

例題 5 では区間 $[a, b] = [0, 20]$ に固有値がある下端固有対を求めてみる。区間 $[a, b] = [0, 20]$ に固有値がある真の固有対の数は 26 個である。

フィルタの設定 (例題 5)

フィルタの伝達関数の 3 つの形状に対する条件として、

表 10 例題 5：フィルタの構成に用いたパラメータ

Table 10 Example-5: parameters used to construct filter for each composition.

| フィルタ | ℓ | ξ | n | 実現した g_s の値 |
|------|--------|-------|-----|---------------|
| B-合成 | 9 | 1.3 | 30 | 7.48E-17 |
| C-合成 | 4 | 1.6 | 17 | 3.51E-17 |
| I-合成 | 5 | 1.3 | 26 | 6.04E-17 |
| E-合成 | 3 | 1.6 | 24 | 6.71E-17 |

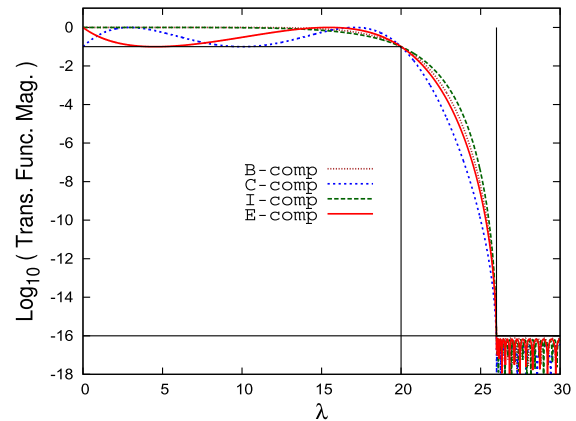


図 19 例題 5：フィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ の対数

Fig. 19 Example-5: log-plots of filter transfer function magnitudes $|\hat{f}(\lambda)|$ for compositions.

4 種類の合成 (B, C, I, E) のどの場合にも $g_p = 0.1$, $g_s \leq 10^{-16}$ を課した。

そうして B-合成と I-合成には次数 ℓ として奇数を許し、 ℓ が奇数になれば (この例の場合には実際そうである) $\xi = 1.3$ と指定し、規格化座標 t の区間 $[0, 1]$ を固有値の区間 $[a, b]$ に対応させることにする。他方で C-合成と E-合成では t の区間 $[-1, 1]$ を固有値の区間 $[a, b]$ に対応させることがつねに可能なのでそのようにして、そのかわり固有値の座標 λ で表したときのフィルタの伝達関数の阻止域の位置を B-合成や I-合成のものと一致させるために、C-合成と E-合成では $\xi = 1.6$ と指定した。

各合成によるフィルタのパラメータを表 10 に示す。E-合成によるフィルタを構成するレゾルベントはシフトが虚数のものと実数のもの 1 つずつである。E-合成で、実現可能な ℓ の最小値 3 を選ばずに $\ell = 4$ とした場合には、形状条件は $n = 9$ で満たせて、 $g_s = 2.29E-17$ になる。

各合成の場合のフィルタの伝達関数の大きさ $|\hat{f}(\lambda)|$ を図 19 に示す。

計算結果 (例題 5)

フィルタを適用するベクトルの数は $m = 50$ とした。各合成によるフィルタを用いて得られた近似固有対の相対残差の最大値および近似固有対を得るまでの経過時間を表 11 に示す。表から経過時間が最も短いのは E-合成の場合であることが分かるが、C-合成の場合とあまり大きな違いはない。

表 11 例題 5：各合成による相対残差の最大値と経過時間

Table 11 Example-5: max relative residuals of approximate eigenpairs and elapsed times (for compositions).

| フィルタ | 相対残差の最大値 | 経過時間 (s) |
|------|----------|----------|
| B-合成 | 3.80E-12 | 892 |
| C-合成 | 5.91E-12 | 299 |
| I-合成 | 3.30E-12 | 458 |
| E-合成 | 2.31E-11 | 264 |

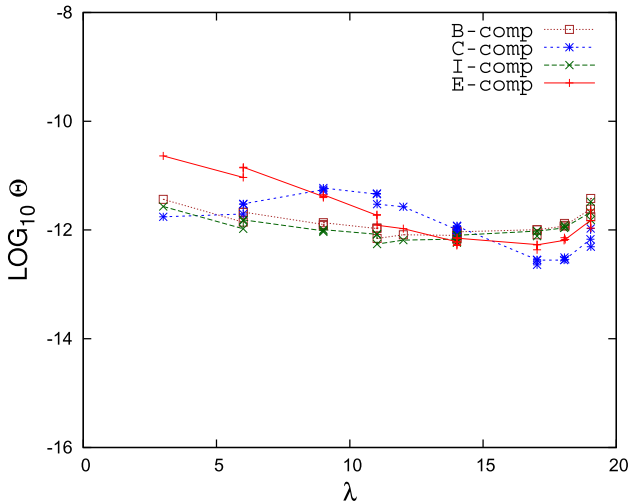


図 20 例題 5：各合成による近似固有対の相対残差 Θ の対数

Fig. 20 Example-5: log-plots of relative residuals Θ of approximate eigenpairs (for compositions).

各合成によるフィルタを用いて得られた近似固有対の相対残差のグラフを図 20 に示す。

なお、E-合成の場合に、実現可能な l の最小値 3 を選ばずに $l = 4$ とすると $n = 9$ にできて、その場合は結果の相対残差の最大値は $3.88E-12$ で経過時間は 231 秒となった。

8.6 各例題で用いた、E-合成の部分分数分解の数値表

例題 1 から例題 5 までの E-合成に用いた部分分数分解の式 (14) において虚部が非負である極 t_j とその係数 c_j の数値 (ただし次数 l が奇数の場合には 1 つだけ存在する実数の極を t_R , その極の係数である実数を c_R と記している), および定数項の係数 c_∞ (実数) の数値をそれぞれ表 12, 表 13, 表 14, 表 15, 表 16 に示す。

表中の数値は通常の IEEE-754 の倍精度浮動小数点 (2 進 64 bit) を用いた計算により求めたもので、本論文の実験に実際に用いた値である (もしも部分分数分解をより精度良く求めたければ、4 倍精度などの高精度の数値と演算を利用することができる)。

IEEE-754 の 64 bit 倍精度浮動小数点数の真の有効精度は 10 進で 15.95 桁であるのに、表の中で数値を 10 進 17 桁の精度で掲載している理由は、IEEE-754 の 2 進表現の倍精度の値はそれを 10 進表現の精度 17 桁の数値に変換してから再び IEEE-754 の 2 進表現の倍精度に変換すれば元の

表 12 例題 1 で用いた E-合成の部分分数分解 ($l = 6$)

Table 12 The complex partial fraction decomposition for the E-composition used in the example-1 ($l = 6$).

| 値 | 実部 (real-part) | 虚部 (imaginary-part) |
|------------|------------------------|------------------------|
| t_1 | 1.0183741988631465E+0 | 9.8314833085967862E-2 |
| t_2 | 0 | 8.3274449632028524E-1 |
| t_3 | -1.0183741988631465E+0 | 9.8314833085967862E-2 |
| c_1 | -5.7673926346438742E-1 | -1.7941921352872067E-1 |
| c_2 | 0 | -4.6337422188923432E+0 |
| c_3 | 5.7673926346438742E-1 | -1.7941921352872067E-1 |
| c_∞ | 0 | |

表 13 例題 2 で用いた E-合成の部分分数分解 ($l = 6$)

Table 13 The complex partial fraction decomposition for the E-composition used in the example-2 ($l = 6$).

| 値 | 実部 (real-part) | 虚部 (imaginary-part) |
|------------|------------------------|------------------------|
| t_1 | 1.0271978792948515E+0 | 1.0071437647109674E-1 |
| t_2 | 0 | 9.0488236260493948E-1 |
| t_3 | -1.0271978792948515E+0 | 1.0071437647109674E-1 |
| c_1 | -4.9959045495499710E-1 | -1.1673752735992586E-1 |
| c_2 | 0 | -4.1632752766212917E+0 |
| c_3 | 4.9959045495499710E-1 | -1.1673752735992586E-1 |
| c_∞ | 0 | |

表 14 例題 3 で用いた E-合成の部分分数分解 ($l = 4$)

Table 14 The complex partial fraction decomposition for the E-composition used in the example-3 ($l = 4$).

| 値 | 実部 (real-part) | 虚部 (imaginary-part) |
|------------|------------------------|------------------------|
| t_1 | 1.1552396197007031E+0 | 4.08977712721237828E-1 |
| t_2 | -1.1552396197007031E+0 | 4.08977712721237828E-1 |
| c_1 | -7.3704751451400419E-1 | -1.2273002275978494E-1 |
| c_2 | 7.3704751451400419E-1 | -1.2273002275978494E-1 |
| c_∞ | 1 | |

表 15 例題 4 で用いた E-合成の部分分数分解 ($l = 5$)

Table 15 The complex partial fraction decomposition for the E-composition used in the example-4 ($l = 5$).

| 値 | 実部 (real-part) | 虚部 (imaginary-part) |
|------------|------------------------|------------------------|
| t_1 | 1.0946512591586728E+0 | 1.3998460308860974E-1 |
| t_2 | -9.9242878792622491E-1 | 9.2097096311897797E-1 |
| t_R | -1.0666626241682344E+0 | |
| c_1 | -1.8267067418533467E-1 | 3.9500805513145824E-2 |
| c_2 | 1.0347753904109060E+0 | -5.4670156890649124E-1 |
| c_R | 5.6396101623216781E-2 | |
| c_∞ | 6.1585829119002800E-1 | |

値に戻るからである。ただし今回の実験では、数値を表から読み込むのではなく、計算して得られた内部 2 進表現の倍精度の数値を直接用いて計算を行った。

なお表の数値は次数 l が偶数である場合は $\hat{x}(t)$ が偶関数であり、符号が反対の極を持つなどの対称性も考慮している。

表 16 例題 5 で用いた E-合成の部分分数分解 ($\ell = 3$)Table 16 The complex partial fraction decomposition for the E-composition used in the example-5 ($\ell = 3$).

| 値 | 実部 (real-part) | 虚部 (imaginary-part) |
|------------|------------------------|------------------------|
| t_1 | 1.6081270689290255E+0 | 1.1428106752166010E+0 |
| t_R | -1.4570474685157588E+0 | |
| c_1 | -9.7349353519661719E-1 | -1.2680166120629213E-1 |
| c_R | 1.8156855388212986E-1 | |
| c_∞ | 7.0659077095890510E-1 | |

9. 単一のレゾルベントを用いるフィルタとの対比の考察

レゾルベントのベクトルへの作用に対応する連立 1 次方程式を直接法で解く場合には、係数行列の分解結果を保持できれば、その後は保持された分解結果を用いて前進代入と後退代入により比較的少ない手間で連立 1 次方程式を解くことができる。レゾルベントを複数のベクトルの組に適用する場合も、係数行列は共通で右辺が複数ある連立 1 次方程式の解の組を求めればよい。ただし、レゾルベントをベクトル m 個の組に対して適用する場合には、行列分解にかかる演算量は m にはよらないが、前進代入と後退代入にかかる演算量はベクトルの数 m に比例するし、ベクトル m 個の組を持つための記憶量も m に比例する。

フィルタとして単一のレゾルベントの n 次多項式、あるいは少数のレゾルベントの線形結合 (の実部) の n 次多項式を用いる定式化は、使用するレゾルベントそれぞれに対応する行列分解を最初に 1 度ずつ行って結果を保持しておけば、それらを用いることによりその後の各レゾルベントの作用を求める計算が容易なることを利用するものである。ただし用いるレゾルベントと同じ数の行列分解を行ってその結果を記憶上に保持できることが前提である。そうして各係数行列の分解結果を利用しながら複数のベクトル m 個に対応して右辺が m 通りの連立 1 次方程式の組を前進代入と後退代入を行って解く処理を、多項式の次数 n に等しい回数ずつ逐次に行うことになる。

文献 [21] では単一のレゾルベントの作用 (シフトが虚数の場合は作用の虚部) のチェビシェフ多項式をフィルタとして用いる方法を取りあげた。この方法は用いるレゾルベントが 1 つだけなので、行列分解の手間や分解結果の保持にかかる記憶量は最も少ないことが最大の利点である。ただし、そのような構成のフィルタの伝達特性はあまり良いものにはできない。たとえば通過域内での伝達率の均質性を良くすることと遷移域の幅を狭くすることは両立しない。

そこで文献 [23] では、フィルタには文献 [21] と同様の単一のレゾルベントによる多項式型のものを採用して、その伝達特性があまり良くないという欠点を補うための一種の反復改良法を取りあげた。指定区間近傍の固有値すべてに対応する不変部分空間を考えて、ベクトルの組に対して正

規直交化に続くフィルタ適用の組合せを数回繰り返す一種の直交化反復法によりベクトルの組で張られる不変部分空間の近似を逐次改良し、最終的に得られたベクトルの組から必要な近似固有対を Rayleigh-Ritz 法で求める方法を示して、数値実験や考察を行った。必要な反復回数もフィルタの特性からおおむね見積もれることを示した。直交化の操作を間に入れてフィルタを反復する過程の全体は、出力が入力に対する線形関係ではないので、その挙動を伝達関数を用いて簡単に分析することはできなくなる。

正規直交化 (あるいは SVD) を施す理由は、フィルタの適用により、精度の限られた数値で表されたベクトルの組の中の各ベクトルが含む固有ベクトルが、相互の伝達率の不均一さから、伝達率が相対的に大きいものが拡大成長して他の数値を覆うために伝達率の小さいものの情報の有効精度が低下する、そのような傾向が進むのを直交化により抑制するためである。

有理関数合成から導かれる少数のレゾルベントの多項式型のフィルタでは、合成用の有理関数 $h(t)$ の次数 ℓ が 2 のときには、4 種類の合成 (B, C, I, E) はすべて一致して、それにより得られるフィルタは虚数のシフトを持つ単一のレゾルベントのチェビシェフ多項式になり、文献 [21], [23] でとりあげたシフトが虚数である場合のものと同じしてしまう。一般には合成用の次数 ℓ は 2 より大きく、したがって必要なレゾルベントの数は 1 つより多くなるから、行列分解の演算量や分解結果を保持する記憶量は単一のレゾルベントからなるフィルタに比べて多くなる。そのため行列分解を保持する記憶量の制約から複数の行列分解は保持できないが単一の行列分解ならば保持できるという場合には、単一のレゾルベントからなるフィルタだけが適用できる。そうして固有対の近似精度を要求する用途であれば、そのフィルタを文献 [23] のように直交化と組み合わせて数回適用することで必要な精度を達成するようにするであろう。

レゾルベントを 1 つだけ使う方法と複数使う方法の優劣は、様々な要因が関係するので簡単ではないと思われるが、たとえば E-合成によるフィルタは比較的少数のレゾルベントを用いても遷移域の幅の狭い伝達特性が実現できるので、固有値を求めたい区間の近くに多数の不要な固有値が分布しているような場合は、必要な固有値の数に比べてフィルタを適用するベクトルの数 m をあまり多くせずに計算ができるので、E-合成によるフィルタにとって有利な状況であろう。

10. おわりに

単一のレゾルベントの多項式で、多項式に n 次の Chebyshev 多項式を用いた場合の簡易型のフィルタは、次数 n を増すことにより阻止域における伝達率の大きさの上限值 g_s を容易に小さくできる。しかしそれと引き換えに、通過域における伝達率の最大最小比 $1/g_p$ が増大する。この比の

値が大きいと（たとえば大きさが4~6桁）、得られる近似解の精度がそれだけ不均等になる可能性があり好ましくない。特に簡易型の伝達関数を用いる場合には、 g_s を増やさずに $1/g_p$ を減らそうとすると、遷移域の幅を決めるパラメータ μ の値を大きくする必要がある。しかし遷移域の幅 $\mu-1$ が広がると、固有値が遷移域にある不要な固有対の数が（通常は）増す。不変部分空間の近似空間の基底をうまく構成するためには、通過域と遷移域を合併した区間に固有値が入る固有ベクトルの数よりも多くのベクトルに対してフィルタを適用する必要があるので、 μ が大きいとそれだけ多くのベクトルにフィルタを適用する必要性が生じて計算の手間が増えるので実用性が低下する。

今回の関数合成を用いる方法で遷移域の幅を縮小できるが、それと引き替えにフィルタを構成に必要なレゾルベントの数は元の1つから複数に増える。

本論文でとりあげたB-合成あるいは特にC-合成やI-合成、さらにE-合成では、たとえば $l=4$ であればシフトが虚数のレゾルベントを2つ用いて、それらの線形結合の実部のChebyshev多項式であるフィルタを構成すれば、遷移域の幅が合成前の元のものよりもかなり狭くなり特性が改善される。 $l=6$ の場合にはシフトが虚数のレゾルベントを3つ用いてさらに特性を改善できる。

そうして l が偶数の場合には、求めたい固有対の固有値の区間 $[a, b]$ は任意に指定することができて中間固有対を求めるのに使える。ただし、連立1次方程式の係数行列の分解にかかる演算量とその分解結果を保持するための記憶量は、フィルタを構成するレゾルベントの数に比例する。それでもシフトが複素数のレゾルベント8~16個程度の線形結合の実部でフィルタを構成する場合に比べれば、シフトが複素数のレゾルベントを少数2~4個用いてフィルタを構成できるので、行列分解にかかる演算量と特に行列の分解結果を保持する記憶量を減らす手段として有効である。

本論文で新しくとりあげたE-合成（楕円型合成）による伝達関数は、4種類の合成（B, C, I, E）のうちで合成の次数 $l (> 2)$ が同じ場合には遷移域の幅が最も狭い急峻な遮断特性を実現することができる。いい換えればE-合成は急峻な遮断特性を最も小さい次数 l により実現できる。そのことは4種類の合成のうちで要求された特性を持つフィルタを最も少ない数のレゾルベントを用いて構成できるのはE-合成であることを意味する。すると、レゾルベントの作用を与える連立1次方程式を直接法で解く前提の下では、4種類の合成から導かれるフィルタのうちで行列分解を行う演算量や分解結果を保持するための記憶量が最も少ないのはE-合成によるものであるということになる。

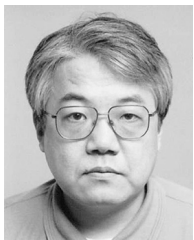
実験例（8章）の中で、実対称定値一般固有値問題の近似固有対を求める実験をE-合成（楕円型合成）によるフィルタを用いて行った例を示した（例題1~5）。大規模な問題では全体の計算の手間の中で行列分解の占める割合が多

くなるが、そのような場合にE-合成は他の合成（B, C, I）に比べて有利であり、経過時間が短いことを確認した。

参考文献

- [1] Abramowitz, M. and Stegun, I.A. (Eds.): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York (1972).
- [2] Daniels, R.W.: *Approximation Methods for Electronic Filter Design*, McGraw-Hill (1974).
- [3] Lutovac, M.D., Tošić, D.Y. and Evans, B.L.: *Filter Design for Signal Processing*, §12.8, Prentice Hall (2001).
- [4] Thompson, W.J.: *Atlas for Computing Mathematical Functions*, Wiley-Interscience (1997).
- [5] Toledo, S. and Rabani, E.: Very Large Electronic Structure Calculations Using an Out-of-Core Filter-Diagonalization Method, *J. Comput. Phys.*, Vol.180, No.1, pp.256-269 (2002).
- [6] Sakurai, T. and Sugiura, H.: A Projection Method for Generalized Eigenvalue Problems Using Numerical Integration, *J. Comput. Appl. Math.*, Vol.159, pp.119-128 (2003).
- [7] Sakurai, T. and Tadano, H.: CIRR: A Rayleigh-Ritz Type Method with Contour Integral for Generalized Eigenvalue Problems, *Hokkaido Math. J.*, Vol.36, No.4, pp.745-757 (2007).
- [8] Polizzi, E.: A Density Matrix-Based Algorithm for Solving Eigenvalue Problems, *Phys. Rev. B*, Vol.79, No.1, pp.115112-115117 (2009).
- [9] Ikegami, T., Sakurai, T. and Nagashima, U.: A Filter Diagonalization for Generalized Eigenvalue Problems Based on the Sakurai-Sugiura Projection Method, *J. Comput. Appl. Math.*, Vol.233, No.8, pp.1927-1936 (2010).
- [10] Olver, F.W.J., Lozier, D.W., Boisvert, R.F. and Clark, C.W. (Eds.): *NIST Handbook of Mathematical Functions*, Cambridge University Press (2010).
- [11] Galgon, M., Krämer, L. and Lang, B.: The FEAST Algorithm for Large Eigenvalue Problems, *PAMM*, Vol.11, pp.747-748 (2011).
- [12] Austin, A.P. and Trefethen, L.N.: Computing Eigenvalues of Real Symmetric Matrices with Rational Filters in Real Arithmetic, *SIAM J. Sci. Comput.*, Vol.37, No.3, pp.A1365-A1387 (2015).
- [13] Güttel, 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).
- [14] Murakami, H.: Filters Consist of a Few Resolvents to Solve Real Symmetric-Definite Generalized Eigenproblems, *Japan J. Indust. Appl. Math.*, Vol.36, No.2, pp.579-618 (online), DOI: 10.1007/s13160-019-00355-5 (2019).
- [15] 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 (ACS31), Vol.3, No.3, pp.1-21 (2010).
- [16] 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 (ACS35), Vol.4, No.4, pp.1-14 (2011).
- [17] 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).

- [18] 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).
- [19] 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 (ACS45), Vol.7, No.1, pp.57-72 (2014).
- [20] 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).
- [21] 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 (ACS64), Vol.12, No.2, pp.1-26 (2019).
- [22] 村上 弘：楕円型合成によるフィルタの伝達特性の改良，第 48 回数値解析シンポジウム予稿集, pp.57-60 (2019).
- [23] 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 (ACS65), Vol.12, No.3, pp.14-33 (2019).
- [24] 村上 弘：有理関数の合成による伝達関数の改良とそれに対応するフィルタの構成について，日本応用数学会 2019 年度年会講演予稿集, pp.566-567 (2019).
- [25] 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).



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，首都大学東京数理科学専攻の准教授。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。

日本応用数学会，日本数式処理学会，日本コンピュータ学会，SIAM，AMS，ACM，IEEE 各会員。