行政院國家科學委員會專題研究計畫 期末報告

對稱多盤上之頻譜Caratheodory-Fejer插值函數

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 100-2115-M-029-003-
執 行 期 間 ：100年08月01日至101年09月30日
執 行 單 位 ：東海大學數學系

計 畫 主 持 人：黃皇男

計畫參與人員：大專生-兼任助理人員：洪永章

報 告 附 件 ：出席國際會議研究心得報告及發表論文

公 開 資 訊 ：本計畫涉及專利或其他智慧財產權，2年後可公開查詢

中 華 民 國 102年01月01日

中 文 摘 要 ： 頻譜 Nevanlinna-Pick(SNP)插值理論是為了提供強健控制器設計之「贡合成理論」確切的數學理論而 發展，其目的在求從單位圓盤到單位頻球的解析(矩陣)函數，並滿足特定的函數值插值條件。利用矩 陣的特徵多項式的性質，可將這類問題轉會成在對稱雙盤上的 NP 插值問題來求解。所謂的頻譜 Caratheodory-Fejer(SCF) 插值問題便是在 SNP 的插值條件中加入導數的要求而得。因此在求解 SCF 問題，必須解決在對稱雙盤上的 CF 插值問題才行。本研究計畫之目的便是探討這類問題的解之存在 性與解的算法，著重在如何將已有的兩點對稱雙盤上的 SCF 插值問題的解法擴大為可求解對稱雙盤 上的 CF 插值問題。本計畫為延續以往計畫而提出的後續研究，基於已完成的研究成果為基礎，結合 SNP 插值理論，探討 3x3 矩陣 SCF 問題的解所形成之空間(稱為插值體)，刻畫出插值體條件以推導 更確切的判斷方法。以探討 2 階導數問題對應的插值體與解為主。同時，對 mu 合成設計的數學理論進行回顧整理。

中文關鍵詞： Nevanlinna-Pick 插值，Caratheodory-Fejer 插值，對稱雙盤，單位頻譜球。

英 文 摘 要 ： The aim of this two years project is to find the Caratheodory-Fejer matricial interpolating function which is analytic from the open unit disc into the open spectral unit ball such that satisfies certain interpolation conditions on its values and its derivatives. It is obvious that this problem is called the spectral Caratheodory-Fejer (SCF) problem. Our approach is to transfer the SCF problem into a classical Caratheodory-Fejer (CF) problem such that a more efficient condition based on the given interpolation data is obtained for the existence of the SCF solution. Based on the result of our previous year’s research project together with the known SNP and SCF theory, direct solvability condition of 3x3 SCF problems can be characterized as an interpolation body. The properties of the interpolation body corresponding to the SCF problem up to second derivatives is analyzed in the first year of this two-year project. Furthermore, the construction of interpolating function is also considered. Extension to higher derivatives problem will be studies in the consecutive year. Meanwhile, a solvable instances of

mu-synthesis is totally reviewed for the present mathematical conclusion.

英文關鍵詞： Nevanlinna-Pick interpolation, Caratheodory-Fejer interpolation, symmetrized polydisc, spectral unit ball.

行政院國家科學委員會補助專題研究計畫 ■成 果 報 告
□期中進度報告

對稱多盤上之頻譜 Caratheordory-Fejer 插值函數

Spectral Caratheodory-Fejer Interpolation Functions in Symmetrise Polydisc

計畫類別: ■ 個別型計畫 □ 整合型計畫
計畫編號: NSC 100 – 2115 – M – 029 – 003 –
執行期間: 100 年 8 月 1 日 至 101 年 9 月 30 日

計畫主持人: 黃皇男
共同主持人:
計畫參與人員: 洪永章、張天財

成果報告類型(依經費核定清單規定繳交): ■精簡報告□ 完整報告

本成果報告包括以下應繳交之附件:
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
■出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式: 除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外, 得立即公開查詢
□涉及專利或其他智慧財產權, □一年■二年後可公開查詢

執行單位: 東海大學數學系

中 華 民 國 一百零一 年 十二 月 三十一 日

Contents

中文摘要

關鍵詞：Nevanlinna-Pick插值，Caratheodory-Fejer插值，對稱雙盤，單位頻譜球。

頻譜 Nevanlinna-Pick(SNP)插值理論是爲了提供強健控制器設計之「 □合成理論 」確切的數學理論而 發展，其目的在求從單位圓盤到單位頻球的解析(矩陣)函數，並滿足特定的函數值插值條件。利用矩 陣的特徵多項式的性質，可將這類問題轉會成在對稱雙盤上的 NP 插值問題來求解。所謂的頻譜 Caratheodory-Fejer(SCF) 插值問題便是在 SNP 的插值條件中加入導數的要求而得。因此在求解 SCF 問題，必須解決在對稱雙盤上的 CF 插值問題才行。本研究計畫之目的便是探討這類問題的解之存在 性與解的算法，著重在如何將已有的兩點對稱雙盤上的 SCF 插值問題的解法擴大爲可求解對稱雙盤 上的 CF 插值問題。本計畫爲延續以往計畫而提出的後續研究，基於已完成的研究成果爲基礎，結合 SNP 插值理論，探討 3x3 矩陣 SCF 問題的解所形成之空間(稱爲插值體)，刻畫出插值體條件以推導 更確切的判斷方法。以探討 2 階導數問題對應的插值體與解爲主。同時，對μ合成設計的數學理論進行回顧整理。

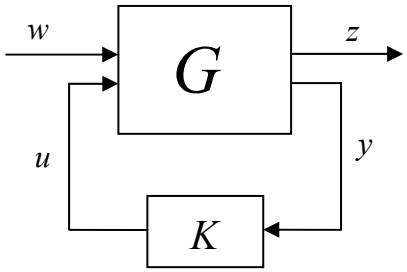
英文摘要

Keywords: Nevanlinna-Pick interpolation, Caratheodory-Fejer interpolation, symmetrized polydisc, spectral unit ball.

The aim of this two years project is to find the Caratheodory-Fejer matricial interpolating function which is analytic from the open unit disc into the open spectral unit ball such that satisfies certain interpolation conditions on its values and its derivatives. It is obvious that this problem is called the spectral Caratheodory-Fejer (SCF) problem. Our approach is to transfer the SCF problem into a classical Caratheodory-Fejer (CF) problem such that a more efficient condition based on the given interpolation data is obtained for the existence of the SCF solution. Based on the result of our previous year ' s research project together with the known SNP and SCF theory, direct solvability condition of 3x3 SCF problems can be characterized as an interpolation body. The properties of the interpolation body corresponding to the SCF problem up to second derivatives is analyzed in the first year of this two-year project. Furthermore, the construction of interpolating function is also considered. Extension to higher derivatives problem will be studies in the consecutive year. Meanwhile, a solvable instances of mu-synthesis is totally reviewed for the present mathematical conclusion.

1 前言

考慮下列廣義的控制系統



其轉移函數可表成

$$\begin{bmatrix} z \\ y \end{bmatrix} = G \begin{bmatrix} w \\ u \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix}$$

這裡 w 稱爲外部輸入，y 爲量測輸出， u 爲控制輸入， z 爲控制輸出，且 K 爲控制器，以對系統進行回授控制，即

$$u = Ky.$$

故全系統從到之轉移函數則變成是

$$\begin{aligned} z &= G_{zw} w \\ &= \left[G_{11} + G_{12} K (I - G_{22} K)^{-1} G_{21} \right] w \end{aligned}$$

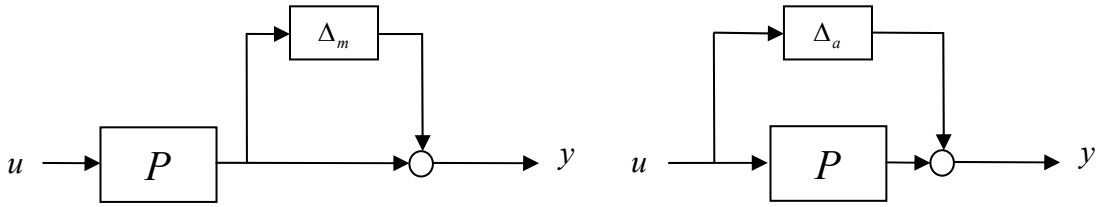
H^∞ 控制的目的便是在給定的控制誤差 γ， 針對所有可能的外部輸入 w ，設計控制器 K，以儘量抑制控制量 z，即使轉移含數滿足條件

$$\|G_{zw}\| < \gamma.$$

此問題自 1980 年代研究至今，並已廣泛應用在實際問題的控制器設計。對應的數學理論可由 Navanlinna-Pick (NP) 插值理論等做一完全的解決，並導出所有控制器存在之條件與解的形式。雖然 N H^∞控制也可以處理當系統 G 內具有下列之非結構化不確定性。

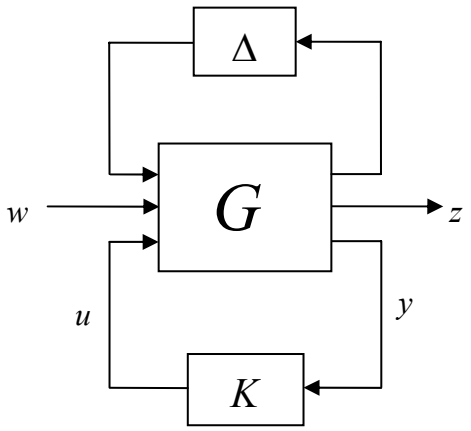
$$\bar{\sigma}(\Delta_m(j\omega)) < |w_m(j\omega)|, \qquad \bar{\sigma}(\Delta_a(j\omega)) < |w_a(j\omega)|, \quad \forall \omega$$

即 G 之長相分別爲



其中 $\bar{\sigma}$ 表最大奇異值， P 表實際受控體。此種設計的缺點是設計出的 K 過於保守，往往使性能表現受阻。

爲了分析具有結構化擾動的系統，提升設計性能則導入結構化特徵值 μ 之控制器合成。設 Δ 表 構化的擾動，在擾動的作用下，廣義控制系統爲



對於擾動 Δ，定義複數矩陣 $M \in \mathbb{C}^{n\times n}$ 之結構化特徵值 $\mu\Delta(M)$ 爲

$$\mu_\Delta(M) = \frac{1}{\min\{\bar{\sigma}(\delta) \ : \ \delta \in \Delta, \ \det(I - M\delta) = 0\}}.$$

如此一來，上述系統爲內部穩定且從 d 到 e 之 H^∞ 範數在 1 以下的充要條件爲

$$\mu\Delta(G_{zw}(j\omega)) < 1, \quad \forall\omega$$

此一設計，自 1990 年後即廣泛應用在飛機、光碟機、汽車等控制器設計。唯一的缺點是雖有許多數 值方法如 D—K 疊代法[M] 可用於估算 μ 之值，但缺乏確切之數學理論，以致於諸如所有控制器存在 的條件至今尚未建立。

爲了建立 μ 合成設計提供確切的數學理論, Newcastle大學的 Prof. N. J. Young 自 1985 年起便開始進 行研究, 至一直到 1999 年左右才與加州大學 San Diego 分校數學系的 Prof. J. Alger 提出以由 spectral Nevanlinna-Pick (SNP) 插值問題來下手解決。此一 SNP 插值問題與 μ 控制器合成之相關性， 可參見 Bercovici, Foias 及 Tannenbaum 等學者之論文[BFT1, BFT2, BFT3]。Alger 及 Young 兩人並一 起合作進行 SNP 插值問題的求解[AY1, AY2, AY3, AY4, AY5, AY6], 至於解的實現 (realization) 則由本系葉芳柏教授和他們兩位一起合作完成[AYY]。到目前爲止, 研究結果只限於 2×2 矩陣的二點插值問題, 個人目前也在研究一般的情形的解。此外, Costara [Cost1, Cost2, Cost3] 以及其他學者 等也都開始研究與此相關的有關代數與幾何問題，但離眞正解出這一問題尚遠。

2 研究方法

本計畫的主要目的便是以先前計畫有關頻譜 Nevanlinna-Pick (SNP) 插值問題的成果爲基礎，進行頻譜 Carathéodory-Fejér (SCF)插值問題的求解計算，以期將來能建立 μ 控制器合成之基礎理論。

設 $\mathbb{D} = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$ 爲複數平面上的單位圓盤，$M_m(\mathbb{C})$ 爲 $m \times m$ 複數矩陣所形成之集合， SNP 插值問題可敘述如下:

(SNP) 已知 $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{D}$ 且 $W_1, W_2, \ldots, W_n \in M_m(\mathbb{C})$，求正則函數 $F : \mathbb{D} \to M_m(\mathbb{C})$ 滿足

$$F(\lambda_j) = W_j, \ 1 \le j \le n$$

且

$$r(F(\lambda)) \le 1, \quad \forall\lambda \in \mathbb{D}.$$

式中 $r(A)$ 表任意方陣 A 之頻譜半徑。

而SCF插值問題可敘述如下:

(SCF) 已知 $\lambda_0 \neq 0$ 且 $V_0, V_1, \ldots, V_n \in M_m(\mathbb{C})$，求正則函數 $F : \mathbb{D} \to M_m(\mathbb{C})$ 滿足

$$F(\lambda_0) = V_0, \quad F^{(j)}(\lambda_0) = V_j, \ 1 \le j \le n$$

且

$$r(F(\lambda)) \le 1, \quad \forall\lambda \in \mathbb{D}.$$

在不失其一般性下，我們可以假設 $\lambda_0 = 0$ 。此一問題的存在性已經由計畫主持人與兩位外國學者共同完成[HMY1], 但對給定的數據資料 $V_0, V_1, \ldots, V_n$，如何驗證正則函數 F 仍是相當困難。同時，此函數之計算也需建立在特徵多項式的插值問題的解上。因此本計畫嘗試運用在求解 SNP 插值問題所發展的出的 magic function 來將 SCF 插值問題轉換成傳統 CF 插值問題來求解。

SCF 插值問題所針對的是一般的矩陣大小，相當困難；爲了能掌握這類問題解的技巧，本計畫先針對 (2×2矩陣)來討論，首先考慮 的特殊情形來討論。如此一來，此問題可改敘如下:

(Simplest SCF) 設 $V_0 = \begin{bmatrix} 0 & -p_0 \\ 1 & s_0 \end{bmatrix}$, $V_1 \in M_m(\mathbb{C})$ 爲 2×2 之複數矩陣，求正則函數 $F : \mathbb{D} \to M_m(\mathbb{C})$ 滿足

$$F(\lambda_0) = V_0, \quad F'(\lambda_0) = V_1$$

且

$$r(F(\lambda)) \leq 1, \quad \forall \lambda \in \mathbb{D}.$$

此一問題有兩個子問題要先處理:

1. 函數 F 存在之充要條件爲何?
2. 若函數 F 確實存在，試問如何計算此 F?

針對一般 SCF 插值問題, 由於矩陣 $F(\lambda)$ 的特徵方程式可表爲

$$\begin{aligned} f(z,\lambda) &= \det(zI - F(\lambda)) \\ &= z^m - c_1(F(\lambda))z^{m-1} + c_2(F(\lambda))z^{m-2} + \cdots + (-1)^m c_m(F(\lambda)) \\ &= z^m - h_1(\lambda)z^{m-1} + h_2(\lambda)z^{m-2} + \cdots + (-1)^m h_m(\lambda) \end{aligned}$$

且令

$$h(\lambda) = (h_1(\lambda), h_2(\lambda), \ldots, h_m(\lambda)) = \mathbf{c} \circ F(\lambda)$$

其中 $\mathbf{c} = (c_1, c_2, \ldots, c_m)$ 以及

$$I^n(0; G_m) = \left\{ (h(0), h'(0), \ldots, h^{(n)}(0)) : h : \mathbb{D} \to G_m \text{ is analytic} \right\},$$

函數 $h(\lambda)$ 以及其導數和矩陣 $V_0, V_1, \ldots, V_n$ 之關係, 可利用[HMY2]的方法來計算，即將其導數視爲函數與 F 的合成函數之導數:

$$h^{(k)}(\lambda) = \Delta^k \mathbf{c}(V_0, V_1, \ldots, V_k),\ 0 \leqslant k \leqslant n$$

其中

$$\begin{aligned} \Delta^0 \mathbf{c}(F(\lambda)) &= \mathbf{c}(F(\lambda)) \\ \Delta^k \mathbf{c}(F(\lambda), F'(\lambda), \ldots, F^{(k+1)}(\lambda)) &= \frac{d}{dt} \Delta^{k-1} \mathbf{c}(F(\lambda) + tF'(\lambda), F'(\lambda) + tF''(\lambda), \\ &\qquad \ldots, F^{(k)}(\lambda) + tF^{(k+1)}(\lambda))\big|_{t=0} \end{aligned}$$

其中 $k = 1, 2, \ldots, n$. G_m 稱爲對稱多重圓盤，爲 m 階多項式

$$(x + z_1)(x + z_2)\cdots(x + z_m) = x^m + c_1^m(z)x^{m-1} + c_2^m(z)x^{m-2} + \cdots + c_{m-1}^m(z)x + c_m^m(z),$$

係數所形成的集合:

$$G_m = \{(c_1^m(z), c_2^m(z), \ldots, c_m^m(z)) : z = (z_1, z_2, \ldots, z_m), |z_j| < 1, 1 \leqslant j \leqslant m\}$$

如此 SCF 插值問題的可解性表示如下:

定理 [HMY1]. 下列敘述同意義:
(1) SCF 插值問題的正則解存在,
(2) $\Delta^n \mathbf{c}(V_0, V_1, \ldots, V_n) \in I^n(0, G_m)$.

同時一旦我們找到滿足

$$\begin{aligned} h^{(k)}(\lambda) &= \Delta^k \mathbf{c}(V_0, V_1, \ldots, V_k),\ 0 \leqslant k \leqslant n, \\ h(\lambda) \in G_m, &\qquad \forall \lambda \in \mathbb{D}. \end{aligned}$$

的函數 $h(\lambda)$ 則利用[HMY1]的方法可以計算出 $F(\lambda)$，此一問題稱之爲特徵函數插值問題。

3 結果與討論

計畫主要成果，分成下面三節說明之。對μ合成數學理論的回顧，與N.J. Young的討論結果另行放在最後作爲附錄。

3.1 Volume of the Symmetrized Polydisc

Let $\mathbb{D}$ denote the unit disk, and $\mathbb{T}$ denote the unit circle. The space G_2, which is the interior of Γ_2, i.e., $\Gamma_2 = \overline{G}_2$, can be characterized by the following theorem:

Proposition 5.1 Let $s, p \in \mathbb{C}$. The following statements are equivalent:

1. $(s,p) \in G_2$,
   (a) $|s - \bar{s}p| < 1 - |p|^2$,
   (b) $2|s - \bar{s}p| + |s^2 - 4p| < 4 - |s|^2$,
   (c) $|s| < 2$, and $\left|\frac{2zp-s}{2-zs}\right| < 1$ for $z \in \mathbb{D}$,
   (d) $|p| < 2$, and there is a $\beta \in \mathbb{D}$ such that $s = \beta\, p + \bar{\beta}$, where $\beta = \frac{\bar{s}-s\bar{p}}{1-|p|^2}$.

Since Γ_2 is a subspace of $\mathbb{C}^2$ and suppose $(s,p) \in \Gamma_2$ then its volume can be computed by

$$\mathrm{Vol}(\Gamma_2) = \int_{\Gamma_2} d\Gamma_2 = \int_{\Gamma_2} ds d\bar{s} dp d\bar{p} \tag{1}$$

where s and p must satisfy the condition 2 in Proposition 3.1, i.e.,

$$|s - \bar{s}p| \leq 1 - |p|^2 . \tag{2}$$

We can't integrate equation (1) directly, thus it must be tranferred it into other form for integration. Consider an equivalent definition for the space of Γ_2:

$$\Gamma_2 = \left\{ (\lambda_1 + \lambda_2, \lambda_1 \cdot \lambda_2) \in \mathbb{C}^2 \middle| \, \lambda_1, \lambda_2 \in \overline{\mathbb{D}} \right\} \tag{3}$$

with λ_1 and λ_2 are two independent variables, i.e.,

$$(s,p) \in \Gamma_2 \Rightarrow s = \lambda_1 + \lambda_2, p = \lambda_1\lambda_2, \quad \lambda_1 \text{and} \, \lambda_2 \in \overline{\mathbb{D}}. \tag{4}$$

The only draw back is the relationship between (s,p) and (λ_1, λ_2) are not one-one. Since the quadratic equation $z^2 - sz + p = (z - \lambda_1)(z - \lambda_2) = 0$ remains the same when two roots, λ_1 and λ_2, are switched (this is why Γ_2 is called the symmetrized bidisc) this concludes that (1) can be expressed by

$$\int_{\Gamma_2} ds d\bar{s} dp d\bar{p} = \frac{1}{2} \int_{\bar{\mathbb{D}}} \int_{\bar{\mathbb{D}}} \left| \det \frac{\partial(s, \bar{s}, p, \bar{p})}{\partial(\lambda_1, \bar{\lambda}_1, \lambda_2, \bar{\lambda}_2)} \right| d\lambda_1 d\bar{\lambda}_1 \, d\lambda_2 d\bar{\lambda}_2 \tag{5}$$

One can compute the associated Jacobin of coordinate transformation as

$$\begin{aligned}
\det \frac{\partial(s, \bar{s}, p, \bar{p})}{\partial(\lambda_1, \bar{\lambda}_1, \lambda_2, \bar{\lambda}_2)} &= \begin{vmatrix} \frac{\partial s}{\partial \lambda_1} & \frac{\partial s}{\partial \bar{\lambda}_1} & \frac{\partial s}{\partial \lambda_2} & \frac{\partial s}{\partial \bar{\lambda}_2} \\ \frac{\partial \bar{s}}{\partial \lambda_1} & \frac{\partial \bar{s}}{\partial \bar{\lambda}_1} & \frac{\partial \bar{s}}{\partial \lambda_2} & \frac{\partial \bar{s}}{\partial \bar{\lambda}_2} \\ \frac{\partial p}{\partial \lambda_1} & \frac{\partial p}{\partial \bar{\lambda}_1} & \frac{\partial p}{\partial \lambda_2} & \frac{\partial p}{\partial \bar{\lambda}_2} \\ \frac{\partial \bar{p}}{\partial \lambda_1} & \frac{\partial \bar{p}}{\partial \bar{\lambda}_1} & \frac{\partial \bar{p}}{\partial \lambda_2} & \frac{\partial \bar{p}}{\partial \bar{\lambda}_2} \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \lambda_2 & 0 & \lambda_1 & 0 \\ 0 & \bar{\lambda}_2 & 0 & \bar{\lambda}_1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & \lambda_1 - \lambda_2 & 0 \\ 0 & 0 & 0 & \bar{\lambda}_1 - \bar{\lambda}_2 \end{vmatrix} \\
&= |\lambda_2 - \lambda_1|^2
\end{aligned} \tag{6}$$

Hence

$$\mathrm{Vol}(\Gamma_2) = \frac{1}{2}\int_{\overline{\mathbb{D}}}\int_{\overline{\mathbb{D}}} |\lambda_2 - \lambda_1|^2 \, d\lambda_1 d\bar{\lambda}_1 \, d\lambda_2 d\bar{\lambda}_2 \tag{7}$$

By using the polar coordinate, i.e., $\lambda_1 = r_1 e^{i\theta_1},\ \lambda_2 = r_2 e^{i\theta_2}$, the above integral becomes

$$\begin{aligned}
\mathrm{Vol}(\Gamma_2) &= \frac{1}{2}\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1 \left|r_2 - r_1 e^{i(\theta_1+\theta_2)}\right|^2 r_1 dr_1 d\theta_1 \, r_2 dr_2 d\theta_2 \\
&= \frac{1}{2}\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1 \left[r_1^2 + r_2^2 - 2r_1 r_2\cos(\theta_1-\theta_2)\right] r_1 dr_1 d\theta_1 \, r_2 dr_2 d\theta_2 \\
&= \pi^2/2.
\end{aligned}$$

Hence we arrive at our conclusion:

$$\int_{\Gamma_2} d\Gamma_2 = \frac{\pi^2}{2} \tag{8}$$

which is the same as the volume of unit ball in $\mathbb{R}^4$. Due to the permulation of the zeros in the polynomial

$$\lambda^2 - s\lambda + p = 0,$$

the volume should be half of the direct integral of (s,p) in the domain Γ_2. The following relationship holds

$$\frac{1}{2}\overline{\mathbb{D}} \times \overline{\mathbb{D}} \subset \Gamma_2 \subset \frac{1}{2}(2\overline{\mathbb{D}}) \times \overline{\mathbb{D}},$$

i.e. the volume of Γ_2 must satisfy

$$\frac{1}{2}\mathrm{Vol}(\overline{\mathbb{D}} \times \overline{\mathbb{D}}) = \frac{\pi^2}{2} \le \mathrm{Vol}(\Gamma_2) \le \frac{1}{2}\mathrm{Vol}((2\overline{\mathbb{D}}) \times \overline{\mathbb{D}}) = 2\pi^2.$$

Define the following symmetrized mapping:

$$\pi_2 \ : \ \overline{\mathbb{D}} \times \overline{\mathbb{D}} \to \Gamma_2 \subset \mathbb{C}^2 : (\lambda_1, \lambda_2) \mapsto (s,p) = (\lambda_1 + \lambda_2, \lambda_1\lambda_2) \tag{9}$$

then the volume of the space is the measure on the range of the mapping π_2. It is obvious true that if $(s,p) = \pi_2(\lambda_1,\lambda_2)$ for two numbers in $\overline{\mathbb{D}}$, the for these two numbers the relation ship $\pi_2(\lambda_2,\lambda_1) = (s,p)$ is also holds.

We extend this relationship by the following recursive relation

$$\pi_n(\lambda_1,\lambda_2,\cdots,\lambda_{n-1},\lambda_n) = (\pi_{n-1}(\lambda_1,\lambda_2,\cdots,\lambda_{n-1}),0) + (1,\pi_{n-1}(\lambda_1,\lambda_2,\cdots,\lambda_{n-1}))\lambda_n \tag{10}$$

with $n = 3,4,\ldots$. When $n=3$,

$$\pi_3(\lambda_1,\lambda_2,\lambda_3) = (\pi_2(\lambda_1,\lambda_2),0) + (1,\pi_2(\lambda_1,\lambda_2))\lambda_3 = (\lambda_1+\lambda_2+\lambda_3, \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3, \lambda_1\lambda_2\lambda_3),$$

i.e.,

$$\pi_3 \ : \ \overline{\mathbb{D}}^3 \to \Gamma_3 \subset \mathbb{C}^3 : (\lambda_1,\lambda_2,\lambda_3) \mapsto (s_1,s_2,s_3) = (\lambda_1+\lambda_2+\lambda_3, \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3, \lambda_1\lambda_2\lambda_3)$$

where Γ_3 is the coefficient space of the polynomial

$$\lambda^3 - s_1\lambda^2 + s_2\lambda - s_3 = 0$$

with all its zeros lies within the unit disk $\mathbb{D}$. And then

$$\int_{\Gamma_3} ds_1 d\bar{s}_1 ds_2 d\bar{s}_2 ds_3 d\bar{s}_3 = \frac{1}{3!}\int_{\overline{\mathbb{D}}}\int_{\overline{\mathbb{D}}}\int_{\overline{\mathbb{D}}} \left|\det \frac{\partial(s_1,\bar{s}_1,s_2,\bar{s}_2,s_3,\bar{s}_3)}{\partial(\lambda_1,\bar{\lambda}_1,\lambda_2,\bar{\lambda}_2,\lambda_3,\bar{\lambda}_3)}\right| d\lambda_1 d\bar{\lambda}_1 \, d\lambda_2 d\bar{\lambda}_2 \, d\lambda_3 d\bar{\lambda}_3 \tag{11}$$

where the assoicated Jacobian is computed as follows:

$$
\begin{aligned}
\det \frac{\partial(s_1,\bar{s}_1,s_2,\bar{s}_2,s_3,\bar{s}_3)}{\partial(\lambda_1,\bar{\lambda}_1,\lambda_2,\bar{\lambda}_2,\lambda_3,\bar{\lambda}_3)} &= \begin{vmatrix} \frac{\partial s_1}{\partial\lambda_1} & \frac{\partial s_1}{\partial\bar{\lambda}_1} & \frac{\partial s_1}{\partial\lambda_2} & \frac{\partial s_1}{\partial\bar{\lambda}_2} & \frac{\partial s_1}{\partial\lambda_3} & \frac{\partial s_1}{\partial\bar{\lambda}_3} \\ \frac{\partial\bar{s}_1}{\partial\lambda_1} & \frac{\partial\bar{s}_1}{\partial\bar{\lambda}_1} & \frac{\partial\bar{s}_1}{\partial\lambda_2} & \frac{\partial\bar{s}_1}{\partial\bar{\lambda}_2} & \frac{\partial\bar{s}_1}{\partial\lambda_3} & \frac{\partial\bar{s}_1}{\partial\bar{\lambda}_3} \\ \frac{\partial s_2}{\partial\lambda_1} & \frac{\partial s_2}{\partial\bar{\lambda}_1} & \frac{\partial s_2}{\partial\lambda_2} & \frac{\partial s_2}{\partial\bar{\lambda}_2} & \frac{\partial s_2}{\partial\lambda_3} & \frac{\partial s_2}{\partial\bar{\lambda}_3} \\ \frac{\partial\bar{s}_2}{\partial\lambda_1} & \frac{\partial\bar{s}_2}{\partial\bar{\lambda}_1} & \frac{\partial\bar{s}_2}{\partial\lambda_2} & \frac{\partial\bar{s}_2}{\partial\bar{\lambda}_2} & \frac{\partial\bar{s}_2}{\partial\lambda_3} & \frac{\partial\bar{s}_2}{\partial\bar{\lambda}_3} \\ \frac{\partial s_3}{\partial\lambda_1} & \frac{\partial s_3}{\partial\bar{\lambda}_1} & \frac{\partial s_3}{\partial\lambda_2} & \frac{\partial s_3}{\partial\bar{\lambda}_2} & \frac{\partial s_3}{\partial\lambda_3} & \frac{\partial s_3}{\partial\bar{\lambda}_3} \\ \frac{\partial\bar{s}_3}{\partial\lambda_1} & \frac{\partial\bar{s}_3}{\partial\bar{\lambda}_1} & \frac{\partial\bar{s}_3}{\partial\lambda_2} & \frac{\partial\bar{s}_3}{\partial\bar{\lambda}_2} & \frac{\partial\bar{s}_3}{\partial\lambda_3} & \frac{\partial\bar{s}_3}{\partial\bar{\lambda}_3} \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ \lambda_2+\lambda_3 & 0 & \lambda_1+\lambda_3 & 0 & \lambda_1+\lambda_2 & 0 \\ 0 & \bar{\lambda}_2+\bar{\lambda}_3 & 0 & \bar{\lambda}_1+\bar{\lambda}_3 & 0 & \bar{\lambda}_1+\bar{\lambda}_2 \\ \lambda_2\lambda_3 & 0 & \lambda_1\lambda_3 & 0 & \lambda_1\lambda_2 & 0 \\ 0 & \bar{\lambda}_2\bar{\lambda}_3 & 0 & \bar{\lambda}_1\bar{\lambda}_3 & 0 & \bar{\lambda}_1\bar{\lambda}_2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & \lambda_1-\lambda_2 & 0 & \lambda_1-\lambda_3 & 0 \\ 0 & 0 & 0 & \bar{\lambda}_1-\bar{\lambda}_2 & 0 & \bar{\lambda}_1+\bar{\lambda}_2 \\ 0 & 0 & (\lambda_1-\lambda_2)\lambda_3 & 0 & (\lambda_1-\lambda_3)\lambda_2 & 0 \\ 0 & 0 & 0 & (\bar{\lambda}_1-\bar{\lambda}_2)\bar{\lambda}_3 & 0 & (\bar{\lambda}_1-\bar{\lambda}_3)\bar{\lambda}_2 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & \lambda_1-\lambda_2 & 0 & \lambda_1-\lambda_3 & 0 \\ 0 & 0 & 0 & \bar{\lambda}_1-\bar{\lambda}_2 & 0 & \bar{\lambda}_1-\bar{\lambda}_3 \\ 0 & 0 & 0 & 0 & (\lambda_1-\lambda_3)(\lambda_2-\lambda_3) & 0 \\ 0 & 0 & 0 & 0 & 0 & (\bar{\lambda}_1-\bar{\lambda}_3)(\bar{\lambda}_2-\bar{\lambda}_3) \end{vmatrix} \\
&= |\lambda_2-\lambda_1|^2|\lambda_3-\lambda_1|^2|\lambda_3-\lambda_2|^2, \qquad (12)
\end{aligned}
$$

Hence

$$
\mathrm{Vol}(\Gamma_3) = \frac{1}{3!}\int_{\overline{\mathbb{D}}}\int_{\overline{\mathbb{D}}}\int_{\overline{\mathbb{D}}} |\lambda_2-\lambda_1|^2|\lambda_3-\lambda_1|^2|\lambda_3-\lambda_2|^2 d\lambda_1 d\bar{\lambda}_1\, d\lambda_2 d\bar{\lambda}_2\, d\lambda_3 d\bar{\lambda}_3 \qquad (13)
$$

By using the polar coordinate, i.e., $\lambda_j = r_j e^{i\theta_j},\ j=1,2,3$, the above integral becomes

$$
\begin{aligned}
\mathrm{Vol}(\Gamma_3) &= \frac{1}{3!}\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1 \left|r_2 - r_1 e^{i(\theta_1+\theta_2)}\right|^2 \left|r_3 - r_1 e^{i(\theta_1+\theta_2)}\right|^2 \left|r_3 - r_2 e^{i(\theta_1+\theta_2)}\right|^2 \\
&\quad r_1 dr_1 d\theta_1\, r_2 dr_2 d\theta_2\, r_3 dr_3 d\theta_3 \\
&= \frac{1}{3}\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1\int_0^{2\pi}\int_0^1 \left[r_1^2 + r_2^2 - 2r_1 r_2\cos(\theta_1-\theta_2)\right] r_1 dr_1 d\theta_1\, r_2 dr_2 d\theta_2 \\
&= \pi^3/3!.
\end{aligned}
$$

Hence we arrive at our conclusion:

$$
\int_{\Gamma_3} d\Gamma_3 = \frac{\pi^2}{3!} \qquad (14)
$$

which is the same as the volume of the unit ball in $\mathbb{R}^4$.

In general, I conject on the following results:

Proposition 5.2: The volume of polydisk Γ_n is equal to the volume of the unit ball in $\mathbb{R}^{2n}$.

By using the Computor Algebra System - Maple 11, we have verified the above proposition upto $n=7$ which means the conjecture is true, however the analytical way to construct the proof is still under development.

3.2 Surface Area of the Symmetrized Polydisc

In mathematics, the Minkowski-Steiner formula is a formula relating the surface area and volume of compact subsets of Euclidean space. More precisely, it defines the surface area as the "derivative" of enclosed volume in an appropriate sense.

The Minkowski-Steiner formula can be states as follow:

Let $n \geq 2$, and let $A \subsetneq \mathbb{R}^n$ be a compact set. Let $\mu(A)$ denote the Lebesgue measure (volume) of A. Define the quantity $\lambda(\partial A)$ by the ``Minkowski-Steiner formula''

$$\lambda(\partial A) \triangleq \liminf_{\delta \to 0} \frac{\mu\left(A + \overline{B_\delta}\right) - \mu(A)}{\delta},$$

where

$$\overline{B_\delta} = \left\{ x = (x_1, \ldots, x_n) \in \mathbb{R}^n \,\middle|\, |x| := \sqrt{x_1^2 + \cdots + x_n^2} \leq \delta \right\}$$

denotes the closed ball of radius $\delta > 0$, and$A + \overline{B_\delta} \triangleq \left\{ a + b \in \mathbb{R}^n \,\middle|\, a \in A, b \in \overline{B_\delta} \right\}$ is the Minkowski sum of A and $\overline{B_\delta}$, so that $A + \overline{B_\delta} = \{x \in \mathbb{R}^n \,||x - a| \leq \delta \text{ for some } a \in A\}$.

For "sufficiently regular" sets A, the quantity $\lambda(\partial A)$ does indeed correspond with the $(n-1)$-dimensional measure (surface) of the boundary ∂A of A.

The volume of the unit ball on $\mathbb{R}^2$ is πr^2, by taking the derivative of r we get the arc-length: $2\pi r$, and the volume of the unit ball on $\mathbb{R}^3$ is $\frac{4}{3}\pi r^3$, we get the surface: $4\pi r^2$.

Now we can apply the Minkowski-Steiner formula to calculate surface area, we have

$$\begin{aligned} S(\Gamma_2) &= 3\pi^2 \\ S(\Gamma_3) &= 2\pi^3 \end{aligned}$$

Till now we only have the gereral form to evalute the volume of Γ_n, we calculate $V(\Gamma_2)$ and $V(\Gamma_3)$ step by step, but still can't find a easy to solve the gereral case. Since the volume of $V(\Gamma_2)$ and $V(\Gamma_3)$ happen to be the volume of the unit ball on $\mathbb{R}^4$ and $\mathbb{R}^6$. The general result is given by

Proposition 5.3: The surface area of polydisk Γ_n is given by $S(\Gamma_n^r) = n(n+1)\frac{\pi^n}{n!}$ which is much larger than the surface area of the unit ball in $\mathbb{R}^{2n}$.

3.3 The CF interpolation function in the symmetrized bidisc

Problem: To seek an analytic function $h : \mathbb{D} \to \mathbb{G}$ such that $h(\lambda_0) = (s_0, p_0) \triangleq z_0 \in \mathbb{G}$ and $h'(\lambda_0) = (s_1, p_1) \triangleq z_1$, i.e., find functions $s(\lambda)$ and $p(\lambda)$ such that $s(\lambda_0) = s_0$, $s'(\lambda_0) = s_1$, $p(\lambda_0) = p_0$, and $p'(\lambda_0) = p_1$, and $(s(\lambda), p(\lambda)) \in \mathbb{G}$, for all $\lambda \in \mathbb{D}$.

First of all we must check the existence of the solution by checking the infinitesimal Carathéodory distance between these two points[HMY1, Theorem 1]. Suppose for this data set we obtain

$$c_{\mathbb{G}}(z_0, z_1) = \sup_{|\omega|=1} \left| \frac{s_1(1 - \omega^2 p_0) - \omega p_1(2 - \omega s_0)}{\omega^2(s_0 - \bar{s}_0 p_0) - 2\omega(1 - |p_0|^2) + \bar{s}_0 - s_0 \bar{p}_0} \right|.$$

with the extremal argument denoted by ω_0. When $c_{\mathbb{G}}(z_0, z_1) \leq 1$, then there does exist such an analytic function.

The derivatives of magic function $\Phi_\omega(s, p) = \dfrac{2\omega p - s}{2 - \omega s}$ are computed as below:

$$\begin{aligned} \frac{\partial \Phi_\omega(s,p)}{\partial s} &= \frac{-(2 - \omega s) - (2\omega p - s)(-\omega)}{(2 - \omega s)^2} = -2\frac{1 - \omega^2 p}{(2 - \omega s)^2}, \\ \frac{\partial \Phi_\omega(s,p)}{\partial p} &= \frac{2\omega}{2 - \omega s}. \end{aligned}$$

Let $f : \mathbb{D} \to \mathbb{D}$ which maps $\lambda \mapsto f(\lambda) = \Phi_{\omega_0} \circ h(\lambda)$, then the interpolation problem from the unit disk to the symmetrized bidisc, i.e., $h : \mathbb{D} \to \mathbb{G}$, is then transformed by the magic function Φ_{ω_0} and becomes an interpolation problem becomes from the unit disk to unit disk with the interpolation conditions imposed on f as:

$$\begin{aligned} f(\lambda_0) &= \Phi_{\omega_0} \circ h(\lambda_0) \\ &= \frac{2\omega_0 p(\lambda_0) - s(\lambda_0)}{2 - \omega_0 s(\lambda_0)} \\ &= \frac{2\omega_0 p_0 - s_0}{2 - \omega_0 s_0} \triangleq c_0, \end{aligned} \tag{15}$$

$$\begin{aligned} f'(\lambda_0) &= \left[\begin{array}{cc} \frac{\partial \Phi_{\omega_0}}{\partial s}(h(\lambda_0)) & \frac{\partial \Phi_{\omega_0}}{\partial p}(h(\lambda_0)) \end{array} \right] \cdot h'(\lambda_0) \\ &= -2\frac{1 - \omega_0^2 p(\lambda_0)}{(2 - \omega_0 s(\lambda_0))^2} s'(\lambda_0) + \frac{2\omega_0}{2 - \omega_0 s(\lambda_0)} p'(\lambda_0) \\ &= -2\frac{1 - \omega_0^2 p_0}{(2 - \omega_0 s_0)^2} s_1 + \frac{2\omega_0 p_1}{2 - \omega_0 s_0} \\ &= 2\frac{\omega_0 p_1 (2 - \omega_0 s_0) - s_1 (1 - \omega_0^2 p_0)}{(2 - \omega_0 s_0)^2} \triangleq c_1. \end{aligned} \tag{16}$$

Let B denote the Blaschke product

$$B_\alpha(\lambda) = \frac{\lambda - \alpha}{1 - \bar{\alpha}\lambda}$$

and its derivative is

$$B'_\alpha(\lambda) = \frac{1 - |\alpha|^2}{(1 - \bar{\alpha}\lambda)^2}.$$

We know that f must be a Blaschke product of degree less than 2. When $c_{\mathbb{G}}(z_0, z_1) = 1$, f is unique up to Mobius transforms, but when $c_{\mathbb{G}}(z_0, z_1) < 1$, it is not. To find the unique or any function f , let

$$q(\lambda) = \frac{B_{c_0}(f(\lambda))}{B_{\lambda_0}(\lambda)} = \frac{f(\lambda) - c_0}{1 - \overline{c_0} f(\lambda)} \frac{1 - \bar{\lambda}_0 \lambda}{\lambda - \lambda_0}$$

then

$$q(\lambda_0) = \lim_{\lambda \to \lambda_0} \frac{f(\lambda) - c_0}{1 - \bar{c}_0 f(\lambda)} \frac{1 - \bar{\lambda}_0 \lambda}{\lambda - \lambda_0} = \frac{f'(\lambda_0)}{1 - |c_0|^2}(1 - |\lambda_0|^2) = \frac{c_1}{1 - |c_0|^2}(1 - |\lambda_0|^2) \triangleq \upsilon.$$

Since

$$\left| \frac{c_1}{1 - |c_0|^2} \right| = \left| \frac{s_1(1 - \omega_0^2 p_0) - \omega_0 p_1 (2 - \omega_0 s_0)}{\omega_0(s_0 - \bar{s}_0 p_0) - 2(1 - |p_0|^2) + \bar{\omega}_0(\bar{s}_0 - s_0 \bar{p}_0)} \right| = c_{\mathbb{G}}(z_0, z_1) \leq 1,$$

i.e., the value of υ is less than or equal to 1 and the function $q(\lambda)$ is solvable. Since $f(\lambda)$ is unique up to Möbius transform, we choose $q(\lambda) = \upsilon$ and then

$$f(\lambda) = B_{-c_0}\left(B_{\lambda_0}(\lambda)\upsilon\right) = \frac{B_{\lambda_0}(\lambda)\upsilon + c_0}{1 + \bar{c}_0 B_{\lambda_0}(\lambda)\upsilon} = \frac{(\upsilon - \bar{\lambda}_0 c_0)\lambda + (c_0 - \lambda_0 \upsilon)}{(1 - \lambda_0 \bar{c}_0 \upsilon) + (\bar{c}_0 \upsilon - \bar{\lambda}_0)\lambda}. \tag{17}$$

Solving the equation $\Phi_{\omega_0} \circ h = f$ for $s(\lambda)$ gives us

$$s(\lambda) = 2\frac{\omega_0 p(\lambda) - f(\lambda)}{1 - \omega_0 f(\lambda)}. \tag{18}$$

In order to guarantee the analyticity of $s(\lambda)$ we must compute the pole of $s(\lambda)$ by solving

$$f(\lambda) = \bar{\omega}_0,$$

here we denoted by $\lambda_* = B_{-\lambda_0}(B_{c_0}(\bar{\omega}_0)/\upsilon)$. When λ_* is outside the unit disc, then the function p to satisfy the interpolation condition $p(\lambda_0) = p_0$ and $p'(\lambda_0) = p_1$ is given by

$$p(\lambda) = B_{-p_0}\left(B_{\lambda_0}(\lambda) B_\zeta\left[B_{\lambda_0}(\lambda) p_2(\lambda)\right]\right), \quad p_2 \in \mathcal{RH}^\infty.$$

where

$$\zeta = \frac{p_1}{1-|p_0|^2}(1-|\lambda_0|^2). \tag{19}$$

For simplicity choose $p_2(\lambda) \equiv 0$ and then we obtain

$$p(\lambda) = B_{-p_0}\left(B_{\lambda_0}(\lambda)\zeta\right) = \frac{B_{\lambda_0}(\lambda)\zeta + p_0}{1+\bar{p}_0 B_{\lambda_0}(\lambda)\zeta} = \frac{(\zeta - \bar{\lambda}_0 p_0)\lambda + (p_0 - \lambda_0\zeta)}{(1-\lambda_0\bar{p}_0\zeta) + (\bar{p}_0\zeta - \bar{\lambda}_0)\lambda}, \tag{20}$$

and $s(\lambda)$ is then given by

$$s(\lambda) = 2\frac{\omega_0 B_{-p_0}\left(B_{\lambda_0}(\lambda)\zeta\right) - B_{-c_0}\left(B_{\lambda_0}(\lambda)\upsilon\right)}{1-\omega_0 B_{-c_0}\left(B_{\lambda_0}(\lambda)\upsilon\right)}. \tag{21}$$

Alternative, when λ_* inside the unit disc (when $c_{\mathbb{G}}(z_0, z_1) = 1$, this always happens) then one more interpolation condition for p, i.e.,

$$p(\lambda_0) = p_0, \quad p'(\lambda_0) = p_1, \quad p(\lambda_*) = \bar{\omega}_0^2, \quad \lambda_* = B_{-\lambda_0}(B_{c_0}(\bar{\omega}_0)/\upsilon).$$

The simplest one is given by

$$p(\lambda) = B_{-p_0}\left(B_{\lambda_0}(\lambda) B_{-\zeta}\left[\frac{B_{\lambda_0}(\lambda)}{B_{\lambda_0}(\lambda_*)} B_\zeta\left(\frac{B_{p_0}(\bar{\omega}_0^2)}{B_{\lambda_0}(\lambda_*)}\right)\right]\right), \tag{22}$$

and associated $s(\lambda)$ is also determined by (18) with $f(\lambda)$ from (17) and $p(\lambda)$ from (22).

3.3.1 Domain extension

When $c_{\mathbb{G}}(z_0, z_1) < 1$, the associated $f(\lambda)$ satisfying the interpolation conditions (15) and (16) is not unique up to Möbius tranforms. We now extend the domain for finding extremal value of $c_{\mathbb{G}}(z_0, z_1)$ from the unit disk to a large disk with radius r, denoted by $r\mathbb{D}$, such that the extremal value is $c_{\mathbb{G}}(z_0, z_1) = r$. And then determine the unique f in this new domain. Suppose $(s_1, p_1) \neq (0, 0)$, the value of r should satisfy the condition

$$\begin{aligned} c_{\mathbb{G}}(z_0, z_1) &= \sup_{|\omega|=r}\left|\frac{s_1(1-\omega^2 p_0) - \omega p_1(2-\omega s_0)}{\omega^2(s_0 - \bar{s}_0 p_0) - 2\omega(1-|p_0|^2) + \bar{s}_0 - s_0\bar{p}_0}\right| \\ &= \left|\frac{s_1(1-\omega_0^2 p_0) - \omega_0 p_1(2-\omega_0 s_0)}{\omega_0^2(s_0 - \bar{s}_0 p_0) - 2\omega_0(1-|p_0|^2) + \bar{s}_0 - s_0\bar{p}_0}\right| = r, \end{aligned}$$

which gives us $c_{\mathbb{G}}(z_0, z_1) = r$ with corresponding extremal argument ω_0. The interpolation condition for the new f is

$$\begin{aligned} f(\lambda_0) &= \Phi_{\omega_0} \circ h(\lambda_0) = \frac{2\omega_0 p_0 - s_0}{2-\omega_0 s_0} = c_0, \\ f'(\lambda_0) &= \left[\begin{array}{cc} \frac{\partial \Phi_{\omega_0}}{\partial s}(h(\lambda_0)) & \frac{\partial \Phi_{\omega_0}}{\partial p}(h(\lambda_0)) \end{array}\right] \cdot h'(\lambda_0) = 2\frac{\omega_0 p_1(2-\omega_0 s_0) - s_1(1-\omega_0^2 p_0)}{(2-\omega_0 s_0)^2} = c_1. \end{aligned}$$

We seek for a Möbius transform $M(\lambda)$ such that

$$M(f(\lambda_0)) = M(c_0) = \lambda_0, \quad \frac{d}{d\lambda} M(f(\lambda))\Big|_{\lambda=\lambda_0} = 1, \quad M(r\mathbb{D}) = r\mathbb{D}.$$

Assume the function $M(\lambda)$ is of the form

$$M(\lambda) = rB_{-\frac{\lambda_0}{r}} \circ B_\alpha(\frac{\lambda}{r} g(\lambda)), \quad \alpha = \frac{c_0}{r} g(c_0)$$

and select an analytic function $g(\lambda)$ such that $\frac{d}{d\lambda} M(f(\lambda))\big|_{\lambda=\lambda_0} = 1$ holds. Direct differentiation on $M(f(\lambda))$ gives us

$$\begin{aligned}
&\frac{d}{d\lambda} M(f(\lambda))\bigg|_{\lambda=\lambda_0} = M'(f(\lambda_0)) f'(\lambda_0) \\
&= rB'_{-\frac{\lambda_0}{r}} \left(B_\alpha \left(\frac{f(\lambda_0) g(f(\lambda_0))}{r} \right) \right) B'_\alpha \left(\frac{f(\lambda_0) g(f(\lambda_0))}{r} \right) \frac{f(\lambda_0) g'(f(\lambda_0)) + g(f(\lambda_0))}{r} f'(\lambda_0) \\
&= rB'_{-\frac{\lambda_0}{r}} \left(B_\alpha \left(\frac{c_0 g(c_0)}{r} \right) \right) B'_\alpha \left(\frac{c_0 g(c_0)}{r} \right) \frac{c_0 g'(c_0) + g(c_0)}{r} c_1 \\
&= B'_{-\frac{\lambda_0}{r}} \left(B_\alpha(\alpha) \right) B'_\alpha(\alpha) [c g'(c_0) + g(c_0)] c_1 \\
&= \left(1 - \left| \frac{\lambda_0}{r} \right|^2 \right) \frac{1 - |\alpha|^2}{(1 - \bar{\alpha}\alpha)^2} [c_0 g'(c_0) + g(c_0)] c_1 \\
&= \frac{1 - \left| \frac{\lambda_0}{r} \right|^2}{1 - |\alpha|^2} [c_0 g'(c_0) + g(c_0)] c_1 \\
&= 1,
\end{aligned}$$

and we want to find the function g to satisfy the differential equation

$$c_0 g'(c_0) + g(c_0) = \frac{1 - |\alpha|^2}{1 - \left| \frac{\lambda_0}{r} \right|^2} \frac{1}{c_1}.$$

Simplest solution is a constant function for g, i.e.,

$$g(\lambda) = k = \frac{1}{c_1} \frac{1 - |\alpha|^2}{1 - \left| \frac{\lambda_0}{r} \right|^2} = \frac{r^2}{c_1} \frac{1 - |\alpha|^2}{r^2 - |\lambda_0|^2},$$

then α must satisfy $\alpha = \frac{r c_0}{c_1} \frac{1 - |\alpha|^2}{r^2 - |\lambda_0|^2}$, i.e., the quantity $\frac{c_1}{c_0}\alpha = r \frac{1 - |\alpha|^2}{r^2 - |\lambda_0|^2}$ is a real number which is the solution of $\left| \frac{c_0}{c_1} \right|^2 \left(\frac{c_1}{c_0} \alpha \right)^2 + \frac{r^2 - |\lambda_0|^2}{r} \left(\frac{c_1}{c_0} \alpha \right) - 1 = 0$. We obtain the parameter α as following

$$\begin{aligned}
\alpha &= \frac{1}{2} \frac{c_0}{c_1} \left| \frac{c_1}{c_0} \right|^2 \left(-\frac{r^2 - |\lambda_0|^2}{r} \pm \sqrt{\left(\frac{r^2 - |\lambda_0|^2}{r} \right)^2 + 4 \left| \frac{c_0}{c_1} \right|^2} \right) \\
&= \frac{1}{2} \frac{\bar{c}_1}{\bar{c}_0} \left(-\frac{r^2 - |\lambda_0|^2}{r} \pm \sqrt{\left(\frac{r^2 - |\lambda_0|^2}{r} \right)^2 + 4 \left| \frac{c_0}{c_1} \right|^2} \right) \\
&= \frac{1}{2} \frac{\bar{c}_1}{r \bar{c}_0} \left(-(r^2 - |\lambda_0|^2) \pm \sqrt{(r^2 - |\lambda_0|^2)^2 + 4r^2 \left| \frac{c_0}{c_1} \right|^2} \right)
\end{aligned}$$

and the parameter k can also be expressed as $k = \frac{r}{c_0}\alpha$. Thus

$$\begin{aligned}
M(\lambda) &= rB_{-\frac{\lambda_0}{r}} \circ B_\alpha(\frac{k}{r}\lambda) = rB_{-\frac{\lambda_0}{r}} \circ B_\alpha(\frac{\alpha}{c_0}\lambda) \\
&= rB_{-\frac{\lambda_0}{r}} \left(\frac{\frac{\alpha}{c_0}\lambda - \alpha}{1 - \bar{\alpha}\frac{\alpha}{c_0}\lambda} \right) = rB_{-\frac{\lambda_0}{r}} \left(\alpha \frac{\lambda - c_0}{c_0 - |\alpha|^2 \lambda} \right) \\
&= r \frac{(\lambda_0 - r\alpha)c_0 + (r\alpha - \lambda_0|\alpha|^2)\lambda}{(r - \bar{\lambda}_0\alpha)c_0 + (\bar{\lambda}_0\alpha - r|\alpha|^2)\lambda}.
\end{aligned}$$

To double check, we see

$$\begin{aligned} M(f(\lambda_0)) &= M(c_0) = \lambda_0, \\ \frac{d}{d\lambda} M(f(\lambda))\Big|_{\lambda=\lambda_0} &= M'(c_0)c_1 \\ &= r\alpha \frac{c_0(r^2-|\lambda_0|^2)(1-|\alpha|^2)}{[(r-\bar{\lambda}_0\alpha)c_0 + (\bar{\lambda}_0\alpha - r|\alpha|^2)\lambda]^2}\Big|_{\lambda=c_0} c_1 \\ &= r\alpha \frac{c_0(r^2-|\lambda_0|^2)(1-|\alpha|^2)}{[r(1-|\alpha|^2)c_0]^2} c_1 \\ &= \frac{1}{r}\frac{c_1}{c_0}\alpha \frac{(r^2-|\lambda_0|^2)}{(1-|\alpha|^2)} = 1. \end{aligned}$$

Also, since $M(\lambda) = rB_{-\frac{\lambda_0}{r}} \circ B_\alpha(\frac{k}{r}\lambda)$, the function $B_\alpha(\frac{k}{r}\lambda)$ maps $r\mathbb{D}$ into $\mathbb{D}$ and the function $B_{-\frac{\lambda_0}{r}}(\lambda)$ maps $\mathbb{D}$ to $\mathbb{D}$, thus $M(\lambda)$ maps $r\mathbb{D}$ into $r\mathbb{D}$. Hence $M(\lambda)$ is the Mobius tranform what we want.

To solve for $h(\lambda)$, i.e., $s(\lambda)$ and $p(\lambda)$, we let $M(f(\lambda)) = \lambda$ to compute the unique $f(t)$ as following

$$\begin{aligned} f(\lambda) &= \Phi_{\omega_0} \circ h(\lambda) = M^{-1}(\lambda) \\ &= \frac{r}{k} B_{-\alpha}\left(B_{\frac{\lambda_0}{r}}(\frac{\lambda}{r})\right) = \frac{c_0}{\alpha} B_{-\alpha}\left(r\frac{\lambda-\lambda_0}{r^2-\bar{\lambda}_0\lambda}\right) \\ &= \frac{c_0}{\alpha}\frac{B_{\frac{\lambda_0}{r}}(\frac{\lambda}{r})+\alpha}{1+\bar{\alpha}B_{\frac{\lambda_0}{r}}(\frac{\lambda}{r})} = \frac{c_0}{\alpha}\frac{r(\lambda-\lambda_0)+\alpha(r^2-\bar{\lambda}_0\lambda)}{r^2-\bar{\lambda}_0\lambda+\bar{\alpha}r(\lambda-\lambda_0)} \\ &= \frac{c_0}{\alpha}\frac{(r-\bar{\lambda}_0\alpha)\lambda + r^2\alpha - r\lambda_0}{r^2 - r\bar{\alpha}\lambda_0 + (r\bar{\alpha}-\bar{\lambda}_0)\lambda}. \end{aligned} \tag{23}$$

Note that

$$f(\lambda_0) = \frac{c_0}{\alpha}\frac{\alpha(r^2-|\lambda_0|^2)}{r^2-|\lambda_0|^2} = c_0$$

and

$$f'(\lambda_0) = \frac{c_0}{\alpha}\left.\frac{r(r^2-|\lambda_0|^2)(1-|\alpha|^2)}{[r^2-r\bar{\alpha}\lambda_0+(r\bar{\alpha}-\bar{\lambda}_0)\lambda]^2}\right|_{\lambda=\lambda_0} = \frac{c_0}{\alpha} r\frac{(r^2-|\lambda_0|^2)(1-|\alpha|^2)}{(r^2-|\lambda_0|^2)^2} = \frac{c_0}{\alpha} r\frac{1-|\alpha|^2}{r^2-|\lambda_0|^2} = c_1.$$

Once the function $f(\lambda)$ is constructed then we can express $s(\lambda)$ as a function of $f(\lambda)$ and $p(\lambda)$, i.e.,

$$s(\lambda) = 2\frac{\omega_0 p(\lambda) - f(\lambda)}{1-\omega_0 f(\lambda)}.$$

To ensure that $s(\lambda)$ is analytic, $p(\lambda)$ satifies the original interpolation condition but also for those λ_* inside $r\mathbb{D}$ such that $f(\lambda_*) = 1/\omega_0 = \bar{\omega}_0/r^2$. It follows that

$$\begin{aligned} \lambda_* &= rB_{-\frac{\lambda_0}{r}} \circ B_\alpha(\frac{\alpha}{\omega_0 c_0}) = rB_{-\frac{\lambda_0}{r}}\left(\alpha\frac{1-\omega_0 c_0}{\omega_0 c_0 - |\alpha|^2}\right) \\ &= r\frac{B_\alpha(\frac{\alpha}{\omega_0 c_0}) + \frac{\lambda_0}{r}}{1+\frac{\bar{\lambda}_0}{r}B_\alpha(\frac{\alpha}{\omega_0 c_0})} = r\frac{r\alpha(1-\omega_0 c_0)+\lambda_0(\omega_0 c_0 - |\alpha|^2)}{r(\omega_0 c_0 - |\alpha|^2)+\alpha\bar{\lambda}_0(1-\omega_0 c_0)}. \end{aligned}$$

Hence $p(\lambda)$ is construct to satify

$$p(\lambda_0) = p_0, \quad p'(\lambda_0) = p_1, \quad p(\lambda_*) = \frac{1}{\omega_0^2} = \frac{\bar{\omega}_0^2}{r^4}.$$

which is given below:

$$p(\lambda) = B_{-p_0}\left(B_{\lambda_0}(\lambda)B_\zeta\left[\frac{B_{\lambda_0}(\lambda)}{B_{\lambda_0}(\lambda_*)}B_{-\zeta}\left(\frac{B_{p_0}(\bar{\omega}_0^2/r^4)}{B_{\lambda_0}(\lambda_*)}\right)\right]\right), \tag{24}$$

where

$$\zeta = \frac{p_1}{1-|p_0|^2}(1-|\lambda_0|^2).$$

3.3.2 Example

An example is presented for illustrative purpose.

To find an analytic function $h : \mathbb{D} \to \mathbb{G}$ such that $h(0) = z_0 = (1, \frac{1}{4})$ and $h'(0) = z_1 = (0, -\frac{1}{4})$, i.e., find functions $s(\lambda)$ and $p(\lambda)$ such that $s(0) = 1$, $s'(0) = 0$, $p(0) = \frac{1}{4}$, and $p'(0) = -\frac{1}{4}$, and $(s(\lambda), p(\lambda)) \in G$, for all $\lambda \in \mathbb{D}$.

Here $\lambda_0 = 0$, $s_0 = 1$, $s_1 = 0$, $p_0 = \frac{1}{4}$, and $p'(0) = -\frac{1}{4}$. One of the solution is given by $h(\lambda) = (1, \frac{1}{4}(1-\lambda))$.

First of all we need to compute $c_{\mathbb{G}}(z_1, z_2)$:

$$\begin{aligned}
& \sup_{|\omega|=1}\left|\frac{s_1(1-\omega^2 p_0)-\omega p_1(2-\omega s_0)}{\omega^2(s_0-\bar{s}_0 p_0)-2\omega(1-|p_0|^2)+\bar{s}_0-s_0\bar{p}_0}\right| \\
= & \sup_{|\omega|=1}\left|\frac{\omega\frac{1}{4}(2-\omega)}{\omega^2(1-\frac{1}{4})-2\omega(1-\frac{1}{16})+1-\frac{1}{4}}\right| \\
= & \sup_{|\omega|=1}\left|\frac{\omega\frac{1}{4}(2-\omega)}{(1+\omega^2)\frac{3}{4}-\frac{15}{8}\omega}\right| = \sup_{|\omega|=1} 2\left|\frac{\omega(2-\omega)}{6-15\omega+6\omega^2}\right| \\
= & \sup_{|\omega|=1}\frac{2}{3}\left|\frac{\omega(2-\omega)}{2-5\omega+2\omega^2}\right| = \sup_{|\omega|=1}\frac{2}{3}\left|\frac{\omega}{1-2\omega}\right| \\
= & \sup_{|\omega|=1}\frac{2}{3}\left|\frac{\omega}{2-\bar{\omega}}\right| = \sup_{|\omega|=1}\frac{2}{3}\left|\frac{\omega}{2-\omega}\right| = \frac{2}{3} \le 1,
\end{aligned}$$

hence there exists an analytic function which satisfy interpolation conditions corresponding to the given data set. The argument for this extremun is given by $\omega_0 = 1$.

Three different methods, direct method, domain extension method, and Schur method, are presented here for comparison.

Direct method: Let $f(\lambda) = \Phi_\omega \circ h(\lambda)$, then the interpolation condition on f is given by

$$f(0) = \frac{2\omega p_0 - s_0}{2-\omega s_0} = \frac{2\omega\frac{1}{4}-1}{2-\omega} = -\frac{1}{2}, \tag{25}$$

$$f'(0) = -2\frac{1+\omega p_0}{(2-\omega s_0)}s_1 + \frac{2\omega}{2-\omega s_0}p_1 = \frac{2\omega}{2-\omega s_0}p_1 = -\frac{1}{2}\frac{\omega}{2-\omega}. \tag{26}$$

At $\omega = \omega_0 = 1$, $c_0 = f(0) = -\frac{1}{2}$ and $c_1 = f'(0) = -\frac{1}{2}$, and then $\upsilon = c_1(1-|\lambda_0|^2)/(1-|c_0|^2) = -\frac{2}{3}$ and $\zeta = p_1(1-|\lambda_0|^2)/(1-|p_0|^2) = -\frac{4}{15}$. Substituting those values into (17) leeds to

$$f(\lambda) = -\frac{1}{2}\frac{\frac{4}{3}\lambda+1}{1+\frac{1}{3}\lambda} = -\frac{1}{2}\frac{4\lambda+3}{3+\lambda}. \tag{27}$$

Solving $f(\lambda) = 1$ gives us $\lambda_* = -\frac{3}{2}$ which is outside the unit disc. Then the function p to satisfy the interpolation condition $p(0) = \frac{1}{4}$ and $p'(0) = -\frac{1}{4}$ is given by (20)

$$p(\lambda) = -\frac{1}{4}\frac{16\lambda-15}{15-\lambda},$$

and from the equation $\Phi_{\omega_0} \circ h = f$ we obtain

$$s(\lambda) = \frac{8\lambda^2-27\lambda-45}{-45-27\lambda+2\lambda^2}.$$

By consider the extremal value of the function

$$\sup_{\lambda\in\mathbb{D}}|\beta| = \sup_{\lambda\in\mathbb{D}}\left|\frac{s(\lambda)-\bar{s}(\lambda)p(\lambda)}{1-|p(\lambda)|^2}\right| = \sup_{\lambda\in\mathbb{T}}\left|\frac{s(\lambda)-\bar{s}(\lambda)p(\lambda)}{1-|p(\lambda)|^2}\right| \simeq 0.96 < 1$$

we know the function

$$h(\lambda) = (s(\lambda), p(\lambda)) = \left(\frac{8\lambda^2-27\lambda-45}{-45-27\lambda+2\lambda^2}, \frac{1}{4}\frac{16\lambda-15}{-15+\lambda}\right)$$

is the required interpolation function into $\mathbb{G}$.

Domain extension method: Since $c_{\mathbb{G}}(z_1,z_2) < 1$, we extend the domain from $\mathbb{D}$ to denoted by $r\mathbb{D}$such that the extremal value is equal to r, i.e.,ld satisfy the condition

$$\sup_{|\omega|=r}\left|\frac{s_1(1-\omega^2p_0)-\omega p_1(2-\omega s_0)}{\omega^2(s_0-\bar{s}_0p_0)-2\omega(1-|p_0|^2)+\bar{s}_0-s_0\bar{p}_0}\right| = \sup_{|\omega|=r}\frac{2}{3}\left|\frac{\omega}{2-\omega}\right| = \sup_{|\omega|=r}\frac{2}{3}\frac{r}{|2-\omega|} = r,$$

which gives us $r=\frac{4}{3}$ with corresponding $\omega_0 = r$. The interpolation condition for the new f is

$$\begin{aligned} c_0 = f(0) &= \frac{2\omega_0\frac{1}{4}-1}{2-\omega_0\cdot 1} = -\frac{1}{2}, \\ c_1 = f'(0) &= \frac{2\omega_0 p_1}{2-\omega_0 s_0} = \frac{-2\omega_0\frac{1}{4}}{2-\omega_0\cdot 1} = -1. \end{aligned}$$

And then $\upsilon = c_1(1-|\lambda_0|^2)/(1-|c_0|^2) = -\frac{4}{3}$, $\zeta = p_1(1-|\lambda_0|^2)/(1-|p_0|^2) = -\frac{4}{15}$, and

$$\begin{aligned} \alpha &= \frac{1}{2}\frac{c_1}{c_0}(-r\pm\sqrt{r^2+4|\frac{c_0}{c_1}|^2}) = -\frac{4}{3}\pm\frac{5}{3} = \frac{1}{3}, -3, \\ \lambda_* &= rB_\alpha(\frac{\alpha}{\omega_0 c_0}) = r\alpha\frac{1-\omega_0 c_0}{\omega_0 c_0-|\alpha|^2} = -\frac{20}{21}, \frac{15}{29}. \end{aligned}$$

Here we choose $\alpha = 1/3$ and $\lambda_* = -20/21$ (inside the disk of $r\mathbb{D}$). The function f is computed by (23), i.e.,

$$f(\lambda) = \frac{c_0}{\alpha}B_{-\alpha}(\frac{\lambda}{r}) = -\frac{3}{2}B_{-\frac{1}{3}}(\frac{3}{4}\lambda) = -\frac{3}{2}\frac{\frac{3}{4}\lambda+\frac{1}{3}}{1+\frac{1}{3}\frac{3}{4}\lambda} = -\frac{1}{2}\frac{9\lambda+4}{4+\lambda},$$

and the function p to satisfy the interpolation condition $p(0)=\frac{1}{4}$, $p'(0)=-\frac{1}{4}$, and $p(-\frac{20}{21}) = \frac{9}{16}$ is

$$p(\lambda) = B_{-\frac{1}{4}}\left(\lambda B_{\frac{4}{15}}\left(-\frac{21}{20}\lambda B_{-\frac{4}{15}}\left(-\frac{21}{20}B_{\frac{1}{4}}(\frac{9}{16})\right)\right)\right) = 4\frac{21\lambda^2-43\lambda+39}{624-64\lambda+21\lambda^2}.$$

Thus the associated $s(\lambda)$ is then given by

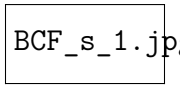
$$s(\lambda) = s(\lambda) = 2\frac{\omega_0 p(\lambda)-f(\lambda)}{1-\omega_0 f(\lambda)} = \frac{59\lambda^2-64\lambda+624}{624-64\lambda+21\lambda^2}.$$

Hence the requested interpolation function is

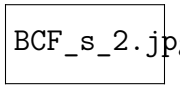
$$h(\lambda) = (s(\lambda), p(\lambda)) = \left(\frac{59\lambda^2-64\lambda+624}{624-64\lambda+21\lambda^2}, 4\frac{21\lambda^2-43\lambda+39}{624-64\lambda+21\lambda^2}\right).$$

Schur method [Schur]: Let $c_0 = f(0)$ and $c_j = f^{(j)}(0)$, $j\geq 1$. Now $c_0 = -\frac{1}{2}$, $c_1 = -\frac{1}{2}\frac{\omega}{2-\omega}$, and we can choose $c_2 = c_3 = \cdots = 0$. The corresponding Schur number is given by

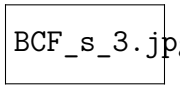
$$\begin{aligned} &\gamma_0 = c_0 = -\frac{1}{2}, \quad \gamma_1 = \frac{c_1}{1-|\gamma_0|^2} = \frac{-\frac{1}{2}\frac{\omega}{2-\omega}}{1-|\frac{1}{2}|^2} = -\frac{2}{3}\frac{\omega}{2-\omega} \\ &\gamma_2 = \frac{\frac{c_2}{1-|\gamma_0|^2}+\bar{\gamma}_0\gamma_1}{1-|\gamma_1|^2} = \frac{2-\frac{2}{3}\omega}{\frac{41}{9}-2(\omega+\bar{\omega})}, \ldots. \end{aligned}$$



(a) direct method



(b) domain extension method



(c) Schur method

Figure 1: The plot of $s(\lambda)$ with the height denoting the real part and the HSV color describing the image part.

We need check $|\gamma_0| \leq 1$, $|\gamma_1| \leq 1$, $|\gamma_2| \leq 1$, ...,$|\gamma_n| \leq 1$,.... Unless there exists some m such that $\gamma_{m+1} = \gamma_{m+2} = \cdots = 0$, it is not easy to apply. On the other hand, the analytic function f satisfy above interpolation condition iff the matrix

$$C = \begin{bmatrix} c_0 & c_1 \\ 0 & c_0 \end{bmatrix}$$

is contractive. This condition is equivalent to

$$1 - |c_0|^2 \geq 0 \Leftrightarrow |\gamma_0| \leq 1,$$
$$1 - |c_0|^2 - |c_1|^2 - \tfrac{|c_0\bar{c}_1|}{1-|c_0|^2} \geq 0 \Leftrightarrow |\gamma_1| \leq 1.$$

Since $(s_0, p_0) = (1, \frac{1}{4})$ is located inside G then $|\gamma_0| \leq 1$ is always true. And the condition from Theorem 1 in Ref. [?] is equivalent to the one $|\gamma_1| \leq 1$. The power series is given by:

$$f(\lambda) = c_0 + c_1\lambda = \frac{1}{2} - \frac{1}{2}\frac{\omega}{2-\omega}\lambda = -\frac{1}{2}\frac{2-\omega+\omega\lambda}{2-\omega} = \frac{2\omega\frac{1}{4}(1-\lambda)-1}{2-\omega} = \frac{2\omega p(\lambda) - s(\lambda)}{2-\omega s(\lambda)}$$

thus

$$s(\lambda) = 1, \quad p(\lambda) = \frac{1}{4}(1-\lambda).$$

Therefore the desired function is then defined by

$$h(\lambda) = (1, \frac{1}{4}(1-\lambda)).$$

Remarks: If we can choose other parameters $c_2, c_3, \ldots$ such that the associated matrix C is contractive, then we can find another function f, e.g. if we select $c_n = (-1)^n\frac{1}{2}(\frac{1}{3})^{n-1}$, then another function f is given by

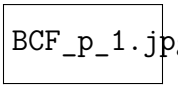
$$f(\lambda) = -\frac{1}{2}(1 + \lambda - \frac{1}{3}\lambda^2 + \frac{1}{3^2}\lambda^3 + \cdots) = -\frac{1}{2}(1 + \lambda\frac{1}{1+\frac{1}{3}\lambda}) = -\frac{1}{2}\frac{1+\frac{4}{3}\lambda}{1+\frac{1}{3}\lambda},$$

which is the same as the function $f(\lambda)$ from the direct method given in (27).

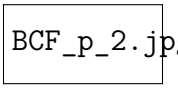
The funcitons $s(\lambda)$ and $p(\lambda)$ are plotted in the following figures. From the comparision between direct and domain extension method, it dipicts the domain extension prove a more smooth function.

3.4 On the Graph of Interpolating Functions

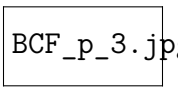
Given two 2×2 matrices W_1 and W_2, compute an analytic function such that $F(\lambda_1) = W_1$, $F(\lambda_2) = W_2$ and $r(F(\lambda)) < 1$, $\forall\lambda \in \mathbb{D}$. Since Σ_2 is a 4-dimensional space which is nonconvex,



(a) direct method



(b) domain extension method



(c) Schur method

Figure 2: The plot of $p(\lambda)$ with the height denoting the real part and the HSV color describing the image part.

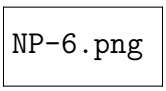


Figure 3: Transformation diagram from Σ_2 into Γ_2.

nosmooth and unbounded set, thus we transfer the domain from Σ_2 into Γ_2 which is also nonconvex and nonsmooth, but a compact set. Therefore we construct the interpolating function defined in Γ_2 first and then transfer it into the original domain Σ_2 whose relationship is shown in Fig3:

In this section, we find the first type of the analytic function $f : \mathbb{D} \to \Gamma_2$ such that

$$f(0) = (0,0), \quad f(\lambda_0) = (s_0, p_0);$$

and then consider the second type of the analytic function $f : \mathbb{D} \to \Gamma_2$ satisfying

$$f(\lambda_1) = (s_1, p_1), \quad f(\lambda_2) = (s_2, p_2).$$

Once the function f is obtained, the original interpolating function $F : \mathbb{D} \to \Sigma_2$ is then computerd.。

3.4.1 First type of interpolating funcitons

Example 3.4.1: To find an analytic function $f : \mathbb{D} \to \Gamma_2$ such that

$$f(0) = (0,0) \ \textit{and} \ f(\beta) = \left(\frac{2\beta}{1+\beta}, 0\right), \beta \in (0,1).$$

Since $f(\lambda) = (s(\lambda), p(\lambda))$, and $f(0) = (0,0)$, $f(\beta) = \left(\frac{2\beta}{1+\beta}, 0\right)$, one obtains

$$\begin{cases} s(0) = 0, \\ s(\beta) = \frac{2\beta}{1+\beta}, \end{cases} \qquad \begin{cases} p(0) = 0, \\ p(\beta) = 0. \end{cases}$$

By using Möbius property, we then arrive at

$$s(\lambda) \equiv \frac{2\lambda(1-\beta)}{1-\beta\lambda}, \qquad p(\lambda) \equiv \frac{\lambda(\lambda-\beta)}{1-\beta\lambda}.$$

That is, $f(\lambda) = \left(\frac{2\lambda(1-\beta)}{1-\beta\lambda}, \frac{\lambda(\lambda-\beta)}{1-\beta\lambda}\right)$ whose graph is shown below:

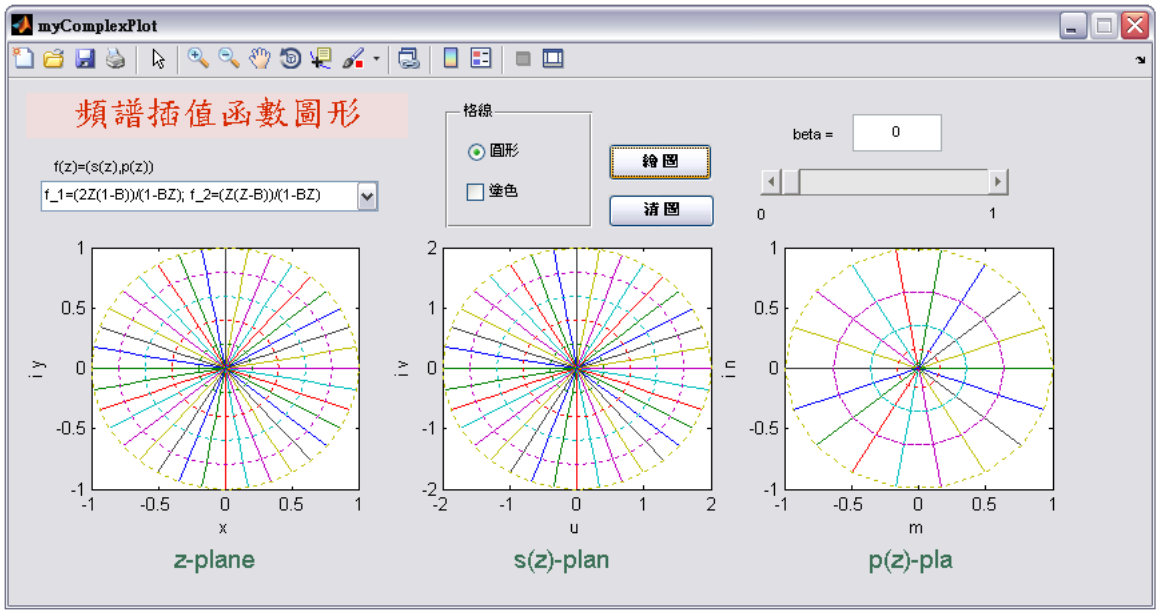


Figure 4: The graph of $f(z)$ when $\beta = 0$.

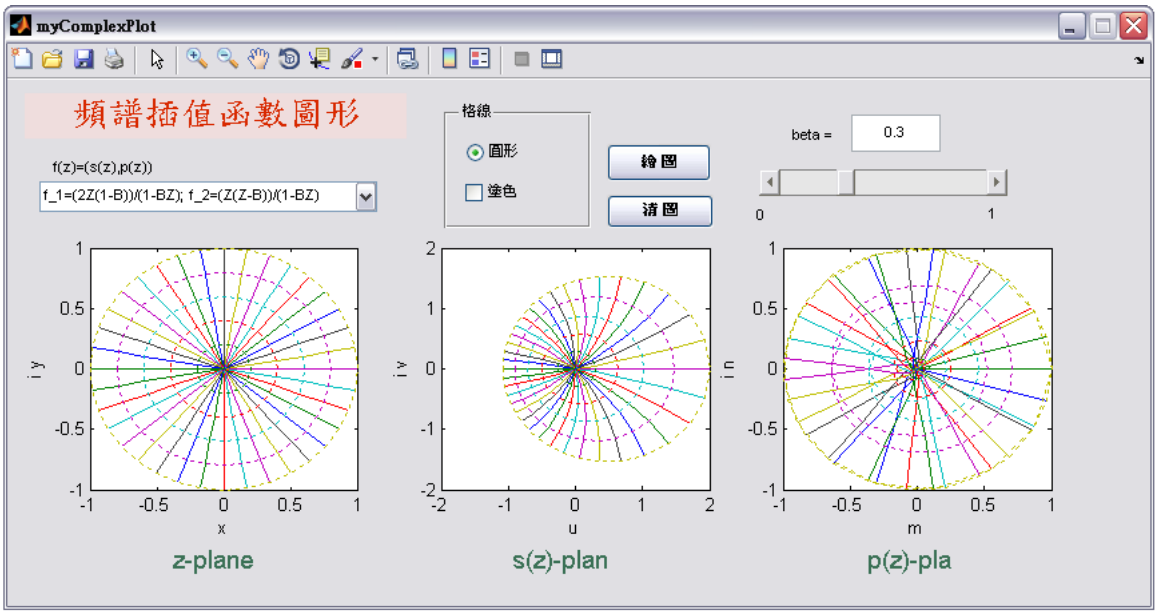


Figure 5: The graph of $f(z)$ when $\beta = 0.3$.

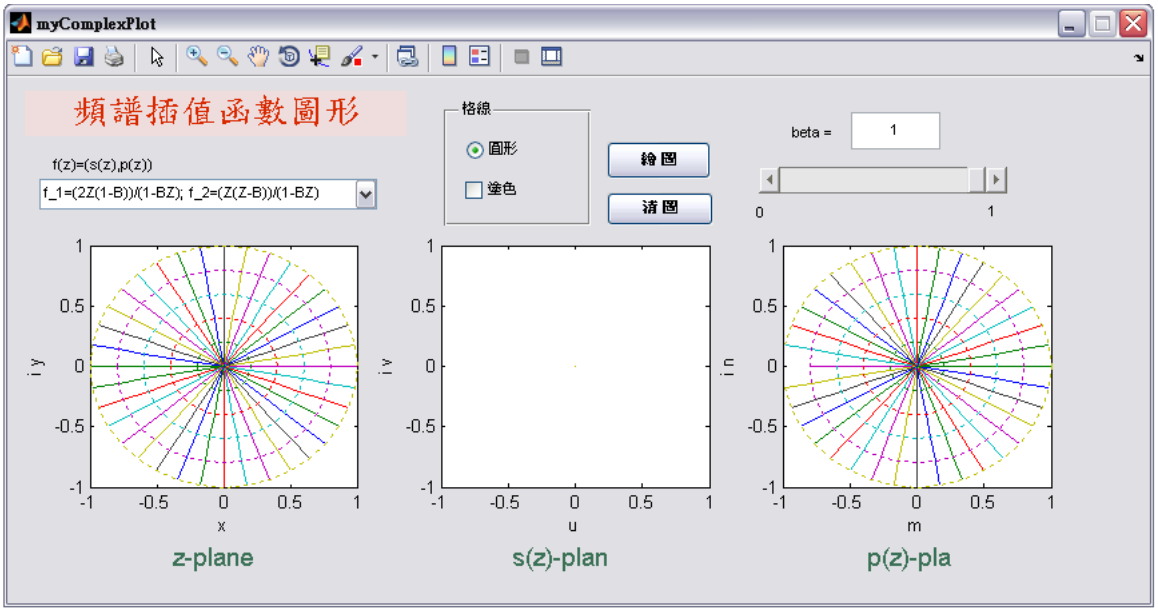


Figure 6: The graph of $f(z)$ when $\beta = 1$.

Example 3.4.2: Given $\lambda_1 = 0$, $\lambda_2 = \beta \in (0,1)$ and $\forall \alpha \in \mathbb{C}$, $F(0) = W_1(\alpha) = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$ and $F(\beta) = W_2 = \begin{bmatrix} 0 & 1 \\ 0 & \frac{2\beta}{1+\beta} \end{bmatrix}$, to find $F(\lambda)$ such that $r(F(\lambda)) \leq 1$, $\forall \lambda \in \mathbb{D}$.

Since $W_1(\alpha) = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}$ and $W_2 = \begin{bmatrix} 0 & 1 \\ 0 & \frac{2\beta}{1+\beta} \end{bmatrix}$ are scalar matrices, there exists a nonsingular matric $P(\lambda)$ such that

$$P(\lambda) F(\lambda) P^{-1}(\lambda) = \begin{bmatrix} 0 & 1 \\ -p(\lambda) & s(\lambda) \end{bmatrix}. \tag{28}$$

And it follows directly that $f(0) = (0,0)$、$f(\beta) = \left(\frac{2\beta}{1+\beta}, 0\right)$, that is

$$s(0) = 0、p(0) = 0;\ s(\beta) = \frac{2\beta}{1+\beta}、p(\beta) = 0$$

which gives us

$$P(0) \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} P(0), \tag{29}$$

$$P(\beta) \begin{bmatrix} 0 & 1 \\ 0 & \frac{2\beta}{1+\beta} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & \frac{2\beta}{1+\beta} \end{bmatrix} P(\beta). \tag{30}$$

Suppose let $P(0) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}$ in (29) and $P(\beta) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ in (30), thenby interpolation

$$P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & \alpha + \frac{\lambda}{\beta}(1-\alpha) \end{bmatrix} \text{ with } \alpha + \frac{\lambda}{\beta}(1-\alpha) \neq 0.$$

Substituting this $P(\lambda)$ into (28) leads to the following analytic function

$$F(\lambda) = \begin{bmatrix} 0 & \alpha + \frac{\lambda}{\beta}(1-\alpha) \\ -\frac{p(\lambda)}{\alpha + \frac{\lambda}{\beta}(1-\alpha)} & s(\lambda) \end{bmatrix}.$$

Alternatively by Example 3.4.1, we set $f(\lambda) = \left(\frac{2\lambda(1-\beta)}{1-\beta\lambda}, \frac{\lambda(\lambda-\beta)}{1-\beta\lambda}\right)$ which gives us the analytic function

$$F(\lambda) = \begin{bmatrix} 0 & \alpha + \frac{\lambda}{\beta}(1-\alpha) \\ -\frac{2\beta\lambda(1-\beta)}{(1-\beta\lambda)(\alpha\beta+\lambda-\alpha\lambda)} & \frac{\lambda(\lambda-\beta)}{1-\beta\lambda} \end{bmatrix}.$$

3.4.2 Second type of interpolating funcitons

Given $F(0) = W_1 = \begin{bmatrix} 1 & 2 \\ -\frac{3}{4} & -1 \end{bmatrix}$ and $F\left(\frac{4}{5}\right) = W_2 = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} & 1 \end{bmatrix}$, to find an analytic function $F(\lambda) \in \mathbb{C}^{2\times 2}$ such that

$$F(0) = W_1、 F\left(\frac{4}{5}\right) = W_2 \text{ and } r(F(\lambda)) < 1, \forall \lambda \in \mathbb{D}.$$

It follows that

$$\lambda_1 = 0,\ W_1 = \begin{bmatrix} 1 & 2 \\ -\frac{3}{4} & -1 \end{bmatrix},\ (s(0), p(0)) = (s_0, p_0) = \left(0, \tfrac{1}{2}\right) = z_0,$$
$$\lambda_2 = \tfrac{4}{5},\ W_2 = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} & 1 \end{bmatrix},\ \left(s\left(\tfrac{4}{5}\right), p\left(\tfrac{4}{5}\right)\right) = (s_1, p_1) = \left(1, \tfrac{1}{4}\right) = z_1.$$

Check the existence of the solution.

Theorem: *For any $W_1,\ W_2 \in \mathbb{C}^{2\times 2}$ and W_1 is nonderagotory , there exists an unique analytic funciton $F(\lambda)$ such that*

$$F(\lambda_1) = W_1 、 F(\lambda_2) = W_2 \text{ and } r(F(\lambda)) < 1, \forall \lambda \in \mathbb{D}$$

if and only if

$$C_{\mathbb{G}}(z_1, z_2) = \sup_{\omega \in \mathbb{T}} \left| \frac{(s_2 p_1 - s_1 p_2)\omega^2 + 2(p_2 - p_1)\omega + s_1 - s_2}{(s_1 - \overline{s_2} p_1)\omega^2 - 2(1 - p_1 \overline{p_2})\omega + \overline{s_2} - s_1 \overline{p_2}} \right| \leq d(\lambda_1, \lambda_2)$$

where

$$z_1 = (s_1, p_1), \quad z_2 = (s_2, p_2) \text{ and } s_i = trW_i, \quad p_i = \det W_i, i = 1, 2$$
$$\mathbb{G} \overset{def}{=} \{(\rho_1 + \rho_2, \rho_1 \rho_2) : |\rho_1| < 1, |\rho_2| < 1\} \subset \mathbb{C}^2$$

Since

$$d\left(0, \frac{4}{5}\right) = \left| \frac{0 - \frac{4}{5}}{1 - \frac{4}{5} \cdot 0} \right| = \frac{4}{5}$$

and

$$C_{\mathbb{G}}(z_1, z_2) = \sup_{\omega \in \mathbb{T}} \left| \frac{\frac{1}{2}\omega^2 - \frac{1}{2}\omega - 1}{-\frac{1}{2}\omega^2 - \frac{4}{7}\omega + 1} \right|$$

then $C_{\mathbb{G}}(z_1, z_2) = \frac{4}{5} = d\left(0, \frac{4}{5}\right)$ whose value is shwon in Figure 7. When $\omega_0 = 1$, the equality holds and thus the unique solution $F(\lambda)$ exists for the given daa set.

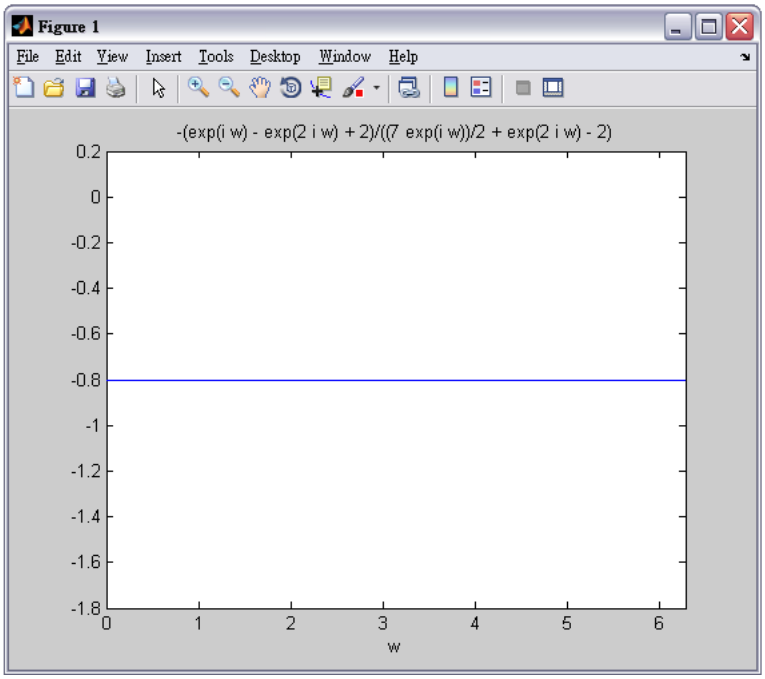


Figure 7: $C_{\mathbb{G}}(z_0, z_1) = \frac{4}{5}$

Suppose there exist functions $s(\lambda)$ and $p(\lambda)$ such that

$$\begin{cases} s(0) = s_1 = 0, \\ s\left(\frac{4}{5}\right) = s_2 = 1, \end{cases} \qquad \begin{cases} p(0) = p_1 = \frac{1}{2}, \\ p\left(\frac{4}{5}\right) = p_2 = \frac{1}{4}, \end{cases}$$

and

$$\begin{cases} s(\lambda) = \mathrm{tr}F(\lambda) \\ p(\lambda) = \det F(\lambda) \end{cases} \qquad (s,p) \in \mathbb{G}_2, \forall \lambda \in \mathbb{D}$$

that is, $f(0) = \left(0, \frac{1}{2}\right)$ 、 $f\left(\frac{4}{5}\right) = \left(1, \frac{1}{4}\right)$, $s(\lambda)$ and $p(\lambda)$ are computed as following.

Step 1: From $\Phi_\omega(s,p) = \dfrac{2\omega p - s}{2 - \omega s}$ we have

$$\Phi_{\omega_0}\left(0, \frac{1}{2}\right) = \frac{2\omega_0 \cdot \frac{1}{2} - 0}{2 - \omega_0 0} = \frac{1}{2}, \quad \Phi_{\omega_0}\left(1, \frac{1}{4}\right) = \frac{2\omega_0 \cdot \frac{1}{4} - 1}{2 - \omega_0 \cdot 1} = -\frac{1}{2}.$$

Step 2: Solve

$$p\left(\frac{1}{2}\right) = \frac{1}{2}, \quad p\left(-\frac{1}{2}\right) = \frac{1}{4}, \quad p(1) = 1.$$

For the condition

$$p\left(\frac{1}{2}\right) = \frac{1}{2}$$

we have

$$p(\alpha) = M_{-\frac{1}{2}} \circ \left(M_{\frac{1}{2}}(\alpha)\, \phi_1(\alpha)\right)$$

and then

$$p\left(-\frac{1}{2}\right) = M_{-\frac{1}{2}} \circ \left(M_{\frac{1}{2}}\left(-\frac{1}{2}\right) \phi_1\left(-\frac{1}{2}\right)\right) = \frac{1}{4}.$$

Thus the function ϕ_1 must satisfy

$$\phi_1\left(-\frac{1}{2}\right) = \frac{5}{14}.$$

Letting

$$\phi_1(\alpha) = M_{-\frac{5}{14}} \circ \left(M_{-\frac{1}{2}}(\alpha)\, \phi_2(\alpha)\right)$$

i.e.,

$$p(\alpha) = M_{-\frac{1}{2}} \circ \left(M_{\frac{1}{2}}(\alpha) M_{-\frac{5}{14}} \circ \left(M_{-\frac{1}{2}}(\alpha) \phi_2(\alpha) \right) \right)$$

and

$$p(1) = 1$$

gives us the requirement

$$\phi_2(1) = 1.$$

For simplicity, we select

$$\phi_2(\alpha) = 1$$

then

$$p(\alpha) = M_{-\frac{1}{2}} \circ \left(M_{\frac{1}{2}}(\alpha) M_{-\frac{5}{14}} \circ (M(\alpha)) \right)$$

i.e.,

$$p(\alpha) = \frac{12\alpha^2 + 5\alpha + 2}{2\alpha^2 + 5\alpha + 12}$$

and

$$s(\alpha) = \frac{2(\omega_0 p(\alpha) - \alpha)}{1 - \alpha\omega_0}$$

which leads to

$$s(\alpha) = \frac{4\alpha^2 - 10\alpha + 4}{2\alpha^2 + 5\alpha + 12}.$$

Step 3: Solve

$$\varphi(0) = \frac{1}{2}, \quad \varphi\left(\frac{4}{5}\right) = -\frac{1}{2}.$$

By

$$\varphi\left(\frac{4}{5}\right) = -\frac{1}{2}$$

we have

$$\varphi(\lambda) = M_{\frac{1}{2}} \circ \left(M_{\frac{4}{5}}(\lambda) \varphi_1(\lambda) \right)$$

and

$$\varphi(0) = M_{\frac{1}{2}} \circ \left(M_{\frac{4}{5}}(0) \varphi_1(0) \right) = \frac{1}{2}$$

which gives us

$$\varphi_1(0) = -1$$

By letting

$$\varphi_1(\lambda) = M_1 \circ (M_1(\lambda) \varphi_2(\lambda))$$

and choose

$$\varphi_2(\lambda) = 1$$

we arrive at

$$\varphi(\lambda) = M_{\frac{1}{2}} \circ \left(M_{\frac{4}{5}}(\lambda) \varphi_1(\lambda) \right) = \frac{2\lambda - 1}{\lambda - 2} = \alpha.$$

Step 4: Since $s(\alpha) = s(\varphi(\lambda)) \triangleq s(\lambda)$; $p(\alpha) = p(\varphi(\lambda)) \triangleq p(\lambda)$, i.e.,

$$s(\alpha) = s(\varphi(\lambda)) = \frac{6\lambda}{10\lambda^2 - 27\lambda + 20}, \quad p(\alpha) = p(\varphi(\lambda)) = \frac{20\lambda^2 - 27\lambda + 10}{10\lambda^2 - 27\lambda + 20},$$

and hence $s(\lambda) = \frac{6\lambda}{10\lambda^2 - 27\lambda + 20}$ and $p(\lambda) = \frac{20\lambda^2 - 27\lambda + 10}{10\lambda^2 - 27\lambda + 20}$), or equivalently $f(\lambda) = \left(\frac{6\lambda}{10\lambda^2 - 27\lambda + 20}, \frac{20\lambda^2 - 27\lambda + 10}{10\lambda^2 - 27\lambda + 20} \right)$.
The graph of $f(\lambda)$ is shown as below:

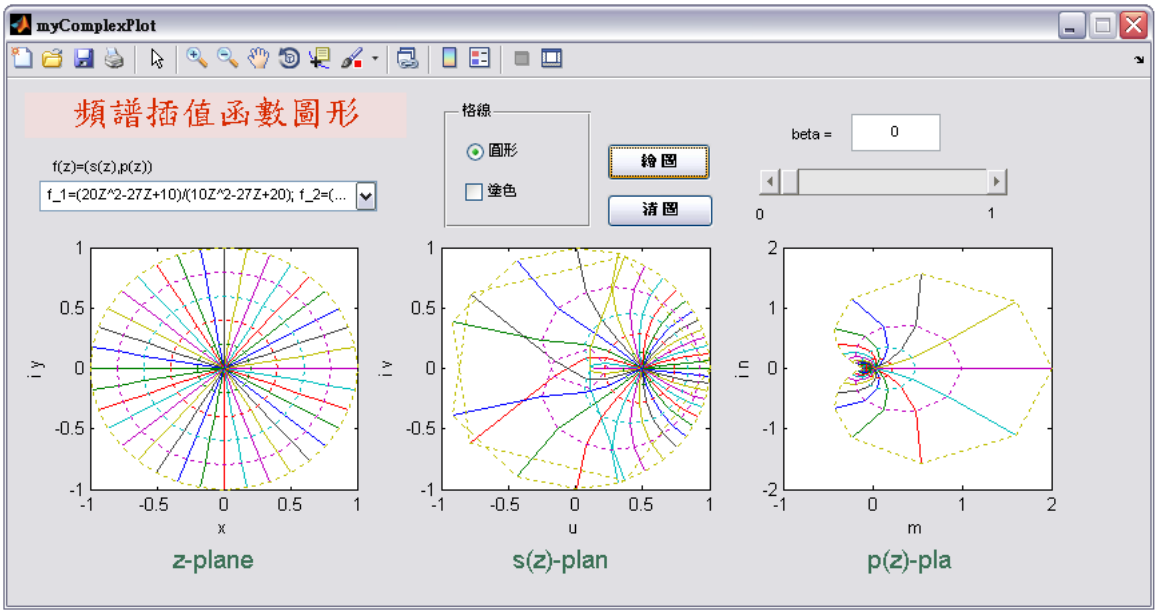


Figure 8: The graph of $f(\lambda) = (\frac{6\lambda}{10\lambda^2-27\lambda+20}, \frac{20\lambda^2-27\lambda+10}{10\lambda^2-27\lambda+20}))$

By the condition $s(\lambda) = \mathrm{tr}F(\lambda)$ and $p(\lambda) = \det F(\lambda)$, the characteristic polynomial of the matrix $F(\lambda)$ is

$$f(z,\lambda) = z^2 - s(\lambda)z + p(\lambda)$$

which must satisfy

$$f(z,0) = z^2 + \frac{1}{2}, \quad f(z,\frac{4}{5}) = z^2 - z + \frac{1}{4}.$$

Using L-shift's invariance to compute $F(\lambda)$.

Define

$$\tilde{f}(z,\lambda) = z^2 f\left(\frac{1}{z}\right) = 1 - s(\lambda)z + p(\lambda)z^2$$

and

$$\tilde{f}(0,\lambda) = 1.$$

$\left\{\frac{u_1(\cdot,\lambda)}{\tilde{f}(\cdot,\lambda)}, \frac{u_2(\cdot,\lambda)}{\tilde{f}(\cdot,\lambda)}\right\}$ is the base of the space of second order polynomial $\mathcal{P}_2$, hence

$$L\frac{u_j(z,\lambda)}{\tilde{f}(z,\lambda)} = \sum_{i=1}^{2} F_{ij}(\lambda)\frac{u_i(z,\lambda)}{\tilde{f}(z,\lambda)}$$

Letting $u(z,\lambda) = \begin{bmatrix} u_1(z,\lambda) \\ u_2(z,\lambda) \end{bmatrix}$ then

$$\frac{1}{z}\left[\frac{u(z,\lambda)}{\tilde{f}(z,\lambda)} - \frac{u(0,\lambda)}{\tilde{f}(z0\lambda)}\right] = F(\lambda)^T \frac{u(z,\lambda)}{\tilde{f}(z,\lambda)} \tag{31}$$

i.e.,

$$u(z,\lambda) - u(0,\lambda)\frac{\tilde{f}(z,\lambda)}{\tilde{f}(0,\lambda)} = zF(\lambda)^T u(z,\lambda)$$

Since $\widetilde{f}(z,0) = 1$,

$$u(z,\lambda) - u(0,\lambda)\tilde{f}(z,\lambda) = zF(\lambda)^T u(z,\lambda)$$

which leads to

$$u(z,\lambda) = \left[1 - zF(\lambda)^T\right]^{-1} u(0,\lambda)\tilde{f}(z,\lambda).$$

When $\lambda = 0$,

$$u(z,0) = \left[1 - zF(0)^T\right]^{-1} \tilde{f}(z,0)\, u(0,0)$$

choose $u(0,0) = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix}$ and it becomes

$$\begin{aligned} u(z,0) &= \begin{bmatrix} 1+z & -\frac{3}{4}z \\ 2z & 1-z \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 & \alpha_1 - \frac{3}{4}\beta_1 \\ \beta_1 & 2\alpha_1 - \beta_1 \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} \end{aligned}$$

where

$$\det \begin{bmatrix} \alpha_1 & \alpha_1 - \frac{3}{4}\beta_1 \\ \beta_1 & 2\alpha_1 - \beta_1 \end{bmatrix} \neq 0 \Rightarrow \alpha_1\beta_1 \neq 0.$$

Therefore, select $\alpha_1 = 1$ and $\beta = 0$ to give

$$u(z,0) = \begin{bmatrix} 1+z \\ 2z \end{bmatrix}$$

When $\lambda = \frac{4}{5}$,

$$u\left(z, \frac{4}{5}\right) = \left[1 - zF\left(\frac{4}{5}\right)^T\right]^{-1} \tilde{f}\left(z, \frac{4}{5}\right) u\left(0, \frac{4}{5}\right)$$

Setting $u\left(0, \frac{4}{5}\right) = \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix}$,

$$\begin{aligned} u\left(z, \frac{4}{5}\right) &= \begin{bmatrix} 1-z & -\frac{1}{4} \\ z & 1 \end{bmatrix} \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} \\ &= \begin{bmatrix} \alpha_2 & -\left(\alpha_2 + \frac{1}{4}\beta_2\right) \\ \beta_2 & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} \end{aligned}$$

where

$$\det \begin{bmatrix} \alpha_2 & -\left(\alpha_2 + \frac{1}{4}\beta_2\right) \\ \beta_2 & \alpha_2 \end{bmatrix} \neq 0 \Rightarrow \alpha_2 \neq -\frac{1}{2}\beta_2.$$

Select $\alpha_2 = 1$、$\beta_2 = 0$ which gives us

$$u\left(z, \frac{4}{5}\right) = \begin{bmatrix} 1-z \\ z \end{bmatrix}.$$

and $u(z,0) = \begin{bmatrix} 1+z \\ 2z \end{bmatrix}$ 、 $u\left(z, \frac{4}{5}\right) = \begin{bmatrix} 1-z \\ z \end{bmatrix}$, then $u(z,\lambda)$ is computed by linear interpolation, i.e.,

$$u(z,\lambda) = \begin{bmatrix} 1+z-\frac{5}{2}\lambda z \\ 2z - \frac{5}{4}\lambda z \end{bmatrix} = \begin{bmatrix} 1 & 1-\frac{5}{2}\lambda \\ 0 & 2-\frac{5}{4}\lambda \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}$$

where

$$\det \begin{bmatrix} 1 & 1-\frac{5}{2}\lambda \\ 0 & 2-\frac{5}{4}\lambda \end{bmatrix} = \frac{5}{4}\left(\frac{8}{5} - \lambda\right) \neq 0, \forall \lambda \in \mathbb{D}.$$

Substituting $u(z,\lambda)$ back into (31)

$$\begin{bmatrix} 1-\frac{5}{2}\lambda + s(\lambda) & -p(\lambda) \\ 2-\frac{5}{4}\lambda & 0 \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix} = F(\lambda)^T \begin{bmatrix} 1 & 1-\frac{5}{2}\lambda \\ 0 & 2-\frac{5}{4}\lambda \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}.$$

Hence

$$\begin{aligned} F(\lambda)^T &= \begin{bmatrix} 1-\frac{5}{2}\lambda+s(\lambda) & -p(\lambda) \\ 2-\frac{5}{4}\lambda & 0 \end{bmatrix} \begin{bmatrix} 1 & 1-\frac{5}{2}\lambda \\ 0 & 2-\frac{5}{4}\lambda \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1-\frac{5}{2}\lambda+s(\lambda) & -p(\lambda) \\ 2-\frac{5}{4}\lambda & 0 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1-\frac{5}{2}\lambda}{2-\frac{5}{4}\lambda} \\ 0 & \frac{1}{2-\frac{5}{4}\lambda} \end{bmatrix} \\ &= \begin{bmatrix} 1-\frac{5}{2}\lambda+s(\lambda) & -\frac{\left(1-\frac{5}{2}\lambda\right)\left[1-\frac{5}{2}\lambda+s(\lambda)\right]+p(\lambda)}{2-\frac{5}{2}\lambda} \\ 2-\frac{5}{4}\lambda & -\left(1-\frac{5}{2}\lambda\right) \end{bmatrix}, \det\left(2-\frac{5}{4}\lambda\right) \neq 0, \forall \lambda \in \mathbb{D} \end{aligned}$$

that is

$$F(\lambda) = \begin{bmatrix} 1-\frac{5}{2}\lambda+s(\lambda) & 2-\frac{5}{4}\lambda \\ -\frac{\left(1-\frac{5}{2}\lambda\right)\left[1-\frac{5}{2}\lambda+s(\lambda)\right]+p(\lambda)}{2-\frac{5}{4}\lambda} & -\left(1-\frac{5}{2}\lambda\right) \end{bmatrix}.$$

Direct substituting the data point into above formulat for verification:

$$F(0) = \begin{bmatrix} 1+s(0) & 2 \\ -\frac{1+s(0)+p(0)}{2} & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -\frac{3}{4} & 1 \end{bmatrix} = V_0$$

$$F\left(\frac{4}{5}\right) = \begin{bmatrix} -1+s\left(\frac{4}{5}\right) & 1 \\ -1+s\left(\frac{4}{5}\right)-p\left(\frac{4}{5}\right) & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{1}{4} & -1 \end{bmatrix} = V_1$$

$$\mathrm{tr}F(\lambda) = s(\lambda),$$

$$\det F(\lambda) = p(\lambda).$$

References

[AY1] J. Agler and N. J. Young, A commutant lifting theorem for a domain in and spectral interpolation, J. Funct. Anal. 161(1999), 452--477.

[AY2] J. Agler and N. J. Young, A Schwarz Lemma for the symmetrized bidisc, Bull. London Math. Soc. 33 (2001), 175--186.

[AY3] J. Agler and N. J. Young, A two-point spectral Nevanlinna-Pick Problem, Integral Equations Operator Theory 37 (2000) 375--385.

[AY4] J. Agler and N. J. Young, A model theory for Γ-contractions, J. Operator Theory 49 (1) (2003) 45--60.

[AY5] J. Agler and N. J. Young, The two-by-two spectral Nevanlinna-Pick Problem, Trans. Amer. Math. Soc. 335(2) (2003) 573--585.

[AY6] J. Agler and N. J. Young, The hyperbolic geometry of the symmetrized bidisc, accepted in J. Geometric Analysis (2004).

[AYY] J. Agler, F.-B. Yeh, and N. J. Young, Realization of functions into the symmetrized bidisc, in Reproducing Kernel Spaces and Applications, Operator Theory: Advances and Applications 143, 1--37, Birkhaüser (2003), edited. by Daniel Alpay.

[B] H. Bercovici, Spectral versus classical Nevanlinna-Pick problem in dimension 2, Electronic Journal of Linear Algebra 10 (2003), 60--64.

[BFT1] H. Bercovici, C. Foias and A. Tannenbaum, A spectral commutant lifting theorem, Trans. Amer. Math. Soc. 325 (1991), 741--763.

[BFT2] H. Bercovici, C. Foias and A. Tannenbaum, Spectral variants of the Nevanlinna-Pick interpolation problem, in Signal processing, scattering and operator theory, and numerical methods (Amsterdam, 1989), 23--45, Progr. Systems Control Theory, 5, Birkhuser Boston, Boston, MA, 1990.

[BFT3] H. Bercovici, C. Foias and A. Tannenbaum, The structured singular value for linear input/output operators, SIAM J. Control Optim. 34(1996).

[C] Ming-Li Chiang, μ-Synthesis via Spectral Interpolation Theory, Master Thesis, Advisor: Huang-Nan Huang, Tunghai University, June 2002.

[CaraFejer] C. Carathéodory and L. Fejér, Über den zusammenhang der extremen von harmonischen funktionen mit ihren Keoffizienten und über den Picard-Landauschen Satz, Rend. Circ. Mat. Palermo 32 (1911), 218--239.

[Cost1] Constantin Costara, Le Problèm De Nevanlinna-Pick Specral, Doctorial Thesis, Université Laval, January 2004.

[Cost2] Constantin Costara, The symmetrized bidisc and Lempert ' s theorem, Bulletin of the London Mathematics Society, 36(5) (2004) 565--662.

[Cost3] Constantin Costara, On the 2×2 spectral Nevanlinna-Pick problem, Preprint, 2004.

[D] J. C. Doyle, Structured uncertainty in control systems, IFAC Workshop on Model Error Concepts and Compensation, Boston, MA 1985.

[EdigZ] A. Edigarian and W. Zwonek, Geometry of symmetrized polydisc, Archiv. der Math. (2004), http://arxiv.org/pdf/math.CV/0402033.

[Hijab] Omar Hijab, The Volume of the Unit Ball in $\mathbb{C}^n$, *The American Mathematical Monthly*, Vol. 107, No. 3 (Mar., 2000), p. 259.

[HJ] R. A. Horn and C. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1990.

[HMY1] H.-N. Huang, S.A.M. Marcantognini, and N. J. Young, The spectral Carathéodory-Fejér problem, Integral Equations and Operator Theory , Published in On-Line First Form (2005).

[HMY2] H.-N. Huang, S.A.M. Marcantognini, and N. J. Young, Chain Rules for Higher Derivatives, The Mathematical Intelligencer (in press, 2005).

[JP1] Marek Jarnicki and Peter Pflug, On automorphisms of the symmetrized bidisc, Archiv der Mathematik 83(3) (2004), 264--266.

[JP2] Marek Jarnicki and Peter Pflug, Invariant Distances and Metrics in Complex Analysis, revisited (2004), http://www.math.uni-oldenburg.de/personen/pflug/new-pr.pdf.

[L1] Cheng-Tsai Lin, Schwarz Lemma On Symmetrized Bidisc, Master Thesis, Advisor: Fang-Bo Yeh, Tunghai University, July 2001.

[L2] Tien-De Lin, Spectral Nevalinna-Pick Interpolation On Symmetrized Bidisc, Master Thesis, Advisor: Fang-Bo Yeh, Tunghai University, July 2001.

[L3] Chun-Ming Lin, Realization of Spectral Nevalinna-Pick Interpolation On Symmetrized Bidisc, Master Thesis, Advisor: Fang-Bo Yeh, Tunghai University, July 2003.

[M] Matlab μ—Analysis and Synthesis Toolbox, The Math Works Inc., Natick, Massachusetts (http://www.mathworks.com/products/muanalysis/).

[P] T. W. Palmer, Banach algebras and the general theory of *-algebras, Cambridge University Press, 1994.

[PY] V. Pták and N. J. Young, A generalization of the zero location theorem of Schur and Cohn, IEEE Trans. Automatic Control 25(1980), 978--980.

[R] J. Rosttand, On the automorphisms of the spectral unit ball, Studia math. 155(3) (2004), 207--230.

[RW] T. J. Randsford and M. C. White, Holomorphic self-maps of the spectral unit ball, Preprint (2004), http://www.mas.ncl.ac.uk/~nmcw/papers/sp1ball2.pdf.

[Schur] I. Schur, Überdie Potenzreihen, die Innern des Einheitskreises beschränkt sind, I & II, J. Reine Angew. Math., 147 (1917), 205--232; 148 (1918), 122--145 (German); English transl.: Operator Theory: Adv. Appl., 18, Birkhäuser Verlag, Basel, 1986.

[T] Wan-Fang Tseng, Minimal Realization for Two-Point Spectral Nevalinna-Pick Problem, Master Thesis, Advisor: Fang-Bo Yeh, Tunghai University, July 2003.

Some analysable instances of mu-synthesis

Abstract. I describe a verifiable criterion for the solvability of the 2×2 spectral Nevanlinna-Pick problem with two interpolation points, and likewise for three other special cases of the μ-synthesis problem. The problem is to construct an analytic 2×2 matrix function F on the unit disc subject to a finite number of interpolation constraints and a bound on the cost function $\sup_{\lambda \in \mathbb{D}} \mu(F(\lambda))$, where μ is an instance of the structured singular value.

Mathematics Subject Classification (2010). Primary 93D21, 93B36; Secondary 32F45, 30E05, 93B50, 47A57.

Keywords. Robust control, stabilization, analytic interpolation, symmetrized bidisc, tetrablock, Carathéodory distance, Lempert function.

1. Introduction

It is a pleasure to be able to speak at a meeting in San Diego in honour of Bill Helton, through whose early papers (especially [31]) I first became interested in applications of operator theory to engineering. I shall discuss a problem of Heltonian character: a hard problem in pure analysis, with immediate applications in control engineering, which can be addressed by operator-theoretic methods. Furthermore, the main advances I shall describe are based on some highly original ideas of Jim Agler, so that San Diego is the ideal place for my talk.

The μ-synthesis problem is an interpolation problem for analytic matrix functions, a generalization of the classical problems of Nevanlinna-Pick, Carathéodory-Fejér and Nehari. The symbol μ denotes a type of cost function that generalizes the operator and H^∞ norms, and the μ-synthesis problem is to construct an analytic matrix function F on the unit disc satisfying a finite number of interpolation conditions and such that $\mu(F(\lambda)) \leq 1$ for $|\lambda| < 1$. The precise definition of μ is in Section 4 below, but for most of the paper we need only a familiar special case of μ – the spectral radius of a square matrix A, which we denote by $r(A)$.

The purpose of this lecture is to present some cases of the μ-synthesis problem that are amenable to analysis. I shall summarize some results that are scattered through a number of papers, mainly by Jim Agler and me but also several others of my collaborators, without attempting to survey all the literature on the topic. I shall also say a little about recent results of some specialists in several complex variables which bear on the matter and may lead to progress on other instances of μ-synthesis.

Although the cases to be described here are too special to have significant practical applications, they do throw some light on the μ-synthesis problem. More concretely, the results below could be used to provide test data for existing numerical methods and to illuminate the phenomenon (known to engineers) of the numerical instability of some μ-synthesis problems.

We are interested in citeria for μ-synthesis problems to be solvable. Here is an example. We denote by $\mathbb{D}$ and $\mathbb{T}$ the open unit disc and the unit circle respectively in the complex plane $\mathbb{C}$.

Theorem 1.1. *Let $\lambda_1, \lambda_2 \in \mathbb{D}$ be distinct points, let W_1, W_2 be nonscalar 2×2 matrices of spectral radius less than 1 and let $s_j = \operatorname{tr} W_j$, $p_j = \det W_j$ for $j = 1, 2$. The following three statements are equivalent:*

(1) *there exists an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that*

$$F(\lambda_1) = W_1, \qquad F(\lambda_2) = W_2$$

and

$$r(F(\lambda)) \leq 1 \quad \text{for all } \lambda \in \mathbb{D};$$

(2)

$$\max_{\omega\in\mathbb{T}} \left| \frac{(s_2p_1 - s_1p_2)\omega^2 + 2(p_2 - p_1)\omega + s_1 - s_2}{(s_1 - \bar{s}_2p_1)\omega^2 - 2(1 - p_1\bar{p}_2)\omega + \bar{s}_2 - s_1\bar{p}_2} \right| \leq \left| \frac{\lambda_1 - \lambda_2}{1 - \bar{\lambda}_2\lambda_1} \right|;$$

(3)

$$\left[\frac{\overline{(2 - \omega s_i)}(2 - \omega s_j) - \overline{(2\omega p_i - s_i)}(2\omega p_j - s_j)}{1 - \bar{\lambda}_i\lambda_j} \right]_{i,j=1}^{2} \geq 0$$

for all $\omega \in \mathbb{T}$.

The paper is organised as follows. Section 2 contains the definition of the spectral Nevanlinna-Pick problem, sketches the ideas that led to Theorem 1.1 – reduction to the complex geommetry of the symmetrized bidisc $\mathbb{G}$, the associated "magic functions" Φ_ω and the calculation of the Carathéodory distance on $\mathbb{G}$ – and fills in the final details of the proof of Theorem 1.1 using the results of [11]. It also discusses ill-conditioning and the possibility of generalization of Theorem 1.1. In Section 3 there is an analogous solvability criterion for a variant of the spectral Nevanlinna-Pick problem in which the two interpolation points coalesce (Theorem 3.1). In Section 4, besides the definition of μ and μ-synthesis, there is some motivation and history. Important work by H. Bercovici, C. Foiaş and A. Tannenbaum is briefly described, as is Bill Helton's alternative approach to robust stabilization problems. In Section 5 we consider an instance of μ-synthesis other than the spectral radius.

Here we can only obtain a solvability criterion in two very special circumstances (Theorems 5.1 and 5.2). The paper concludes with some speculations in Section 6.

We shall denote the closed unit disc in the complex plane by Δ.

2. The spectral Nevanlinna-Pick problem

A particularly appealing special case of the μ-synthesis problem is the *spectral Nevanlinna-Pick problem*:

Problem SNP *Given distinct points $\lambda_1, \dots, \lambda_n \in \mathbb{D}$ and $k \times k$ matrices $W_1, \dots, W_n$, construct an analytic $k \times k$ matrix function F on $\mathbb{D}$ such that*

$$F(\lambda_j) = W_j \quad \text{for } j = 1, \dots, n \tag{2.1}$$

and

$$r(F(\lambda)) \leq 1 \quad \text{for all } \lambda \in \mathbb{D}. \tag{2.2}$$

When $k = 1$ this is just the classical Nevanlinna-Pick problem, and it is well known that a suitable F exists if and only if a certain $n \times n$ matrix formed from the λ_j and W_j is positive (this is *Pick's Theorem*). We should very much like to have a similarly elegant solvability criterion for the case that $k > 1$, but strenuous efforts by numerous mathematicians over three decades have failed to find one.

About 15 years ago Jim Agler and I devised a new approach to the problem in the case $k = 2$ based on operator theory and a dash of several complex variables [5] to [13]. Since interpolation of the eigenvalues fails, how about interpolation of the coefficients of the characteristic polynomials of the W_j, or in other words of the elementary symmetric functions of the eigenvalues? This thought brought us to the study of the complex geometry of a certain set $\Gamma \subset \mathbb{C}^2$, defined below. By this route we were able to analyse quite fully the simplest then-unsolved case of the spectral Nevanlinna-Pick problem: the case $n = k = 2$. For the purpose of engineering application this is a modest achievement, but it nevertheless constituted progress. It had the merit of revealing some unsuspected intricacies of the problem, and may yet lead to further discoveries.

2.1. The symmetrized bidisc Γ

We introduce the notation

$$\begin{aligned} \Gamma &= \{(z + w, zw) : z, w \in \Delta\}, \\ \mathbb{G} &= \{(z + w, zw) : z, w \in \mathbb{D}\}. \end{aligned} \tag{2.3}$$

Γ and $\mathbb{G}$ are called the *closed* and *open symmetrized bidiscs* respectively. Their importance lies in their relation to the sets

$$\Sigma \stackrel{\text{def}}{=} \{A \in \mathbb{C}^{2\times 2} : r(A) \leq 1\},$$
$$\Sigma^o \stackrel{\text{def}}{=} \{A \in \mathbb{C}^{2\times 2} : r(A) < 1\}.$$

Σ and its interior Σ^o are sometines called "spectral unit balls", though the terminology is misleading since they are not remotely ball-like, being unbounded and non-convex. Observe that, for a 2×2 matrix A,

$$\begin{aligned} A \in \Sigma \Leftrightarrow & \text{ the zeros of the polynomial } \lambda^2 - \operatorname{tr} A\lambda + \det A \text{ lie in } \Delta \\ & \Leftrightarrow \operatorname{tr} A = z + w, \det A = zw \text{ for some } z, w \in \Delta. \end{aligned}$$

We thus have the following simple assertion.

Proposition 2.1. *For any* $A \in \mathbb{C}^{2\times 2}$

$$A \in \Sigma \text{ if and only if } (\operatorname{tr} A, \det A) \in \Gamma,$$
$$A \in \Sigma^o \text{ if and only if } (\operatorname{tr} A, \det A) \in \mathbb{G}.$$

Consequently, if $F : \mathbb{D} \to \Sigma$ is analytic and satisfies the equations (2.1) above, where $k = 2$, then $h \stackrel{\text{def}}{=} (\operatorname{tr} F, \det F)$ is an analytic map from $\mathbb{D}$ to Γ satisfying the interpolation conditions

$$h(\lambda_j) = (\operatorname{tr} W_j, \det W_j) \text{ for } j = 1, \ldots, n. \tag{2.4}$$

Let us assume that none of the target matrices W_j is a scalar multiple of the identity. On this hypothesis it is simple to show the converse [16] by similarity transformation of the W_j to companion form.

Proposition 2.2. *Let* $\lambda_1, \ldots, \lambda_n$ *be distinct points in* $\mathbb{D}$ *and let* $W_1, \ldots, W_n$ *be nonscalar* 2×2 *matrices. There exists an analytic map* $F : \mathbb{D} \to C^{2\times 2}$ *such that equations* (2.1) *and* (2.2) *hold if and only if there exists an analytic map* $h : \mathbb{D} \to \Gamma$ *that satisfies the conditions* (2.4)*.*

We have therefore (in the case $k = 2$) reduced the given analytic interpolation problem for Σ-valued functions to one for Γ-valued functions (the assumption on the W_j is harmless, since any constraint for which W_j is scalar may be removed by the standard process of Schur reduction).

Why is it an advance to replace Σ by Γ? For one thing, of the two sets, the geometry of Γ is considerably the less rebarbative. Σ is an unbounded, non-smooth 4-complex-dimensional set with spikes shooting off to infinity in many directions. Γ is somewhat better: it is compact and only 2-complex-dimensional, though Γ too is non-convex and not smoothly bounded. But the true reason that Γ is amenable to analysis is that there is a 1-parameter family of linear fractional functions, analytic on $\mathbb{G}$, that has special properties *vis-à-vis* Γ. For ω in the unit circle $\mathbb{T}$ we define

$$\Phi_\omega(s, p) = \frac{2\omega p - s}{2 - \omega s}. \tag{2.5}$$

We use the variables s and p to suggest "sum" and "product". The Φ_ω determine $\mathbb{G}$ in the following sense.

Proposition 2.3. *For every $\omega \in \mathbb{T}$, Φ_ω maps $\mathbb{G}$ analytically into $\mathbb{D}$. Conversely, if $(s,p) \in \mathbb{C}^2$ is such that $|\Phi_\omega(s,p)| < 1$ for all $\omega \in \mathbb{T}$, then $(s,p) \in \mathbb{G}$.*

Both statements can be derived from the identity

$$|2 - z - w|^2 - |2zw - z - w|^2 = 2(1 - |z|^2)|1 - w|^2 + 2(1 - |w|^2)|1 - z|^2.$$

See [11, Theorem 2.1] for details.

There is an analogous statement for Γ, but there are some subtleties. For one thing Φ_ω is undefined at $(2\bar{\omega}, \bar{\omega}^2) \in \Gamma$ when $\omega \in \mathbb{T}$.

Proposition 2.4. *For every $\omega \in \mathbb{T}$, Φ_ω maps $\Gamma \setminus \{(2\bar{\omega}, \bar{\omega}^2)\}$ analytically into Δ. Conversely, if $(s,p) \in \mathbb{C}^2$ is such that $|\Phi_\omega(rs, r^2p)| < 1$ for all $\omega \in \mathbb{T}$ and $0 < r < 1$ then $(s,p) \in \Gamma$.*

In the second statement of the proposition the parameter r is needed: it does not suffice that $|\Phi_\omega(s,p)| \leq 1$ for all $\omega \in \mathbb{T}$ (in the case that $p = 1$ the last statement is true if and only if $s \in \mathbb{R}$, whereas for $(s,p) \in \Gamma$, of course $|s| \leq 2$).

We found the functions Φ_ω by applying Agler's theory of families of operator tuples [5, 6]. We studied the family $\mathcal{F}$ of commuting pairs of operators for which Γ is a spectral set, and its dual cone $\mathcal{F}^\perp$ (that is, the collection of hereditary polynomials that are positive on $\mathcal{F}$). Agler had previously done the analogous analysis for the bidisc, and shown that the dual cone was generated by just two hereditary polynomials; this led to his celebrated realization theorem for bounded analytic functions on the bidisc. On incorporating symmetry into the analysis we found that the cone $\mathcal{F}^\perp$ had the 1-parameter family of generators $1 - \Phi_\omega^\vee \Phi_\omega$, $\omega \in \mathbb{T}$. From this fact many conclusions follow: see [13] for more on these ideas.

Operator theory played an essential role in our discovery of the functions Φ_ω. Once they are known, however, the geometry of $\mathbb{G}$ and Γ can be developed without the use of operator theory.

2.2. A necessary condition

Suppose that F is a solution of the spectral Nevanlinna-Pick problem (2.1), (2.2) with $k = 2$. For any $\omega \in \mathbb{T}$ and $0 < t < 1$ the composition

$$\mathbb{D} \xrightarrow{tF} \Sigma^o \xrightarrow{(\mathrm{tr},\det)} \mathbb{G} \xrightarrow{\Phi_\omega} \mathbb{D}$$

is an analytic self-map of $\mathbb{D}$ for which

$$\lambda_j \mapsto \Phi_\omega(t \operatorname{tr} W_j, t^2 \det W_j) = \frac{2\omega t^2 \det W_j - t \operatorname{tr} W_j}{2 - \omega t \operatorname{tr} W_j} \quad \text{for } j = 1, \ldots, n.$$

Thus, by Pick's Theorem,

$$\left[\frac{1 - \bar{\Phi}_\omega(t \operatorname{tr} W_i, t^2 \det W_i) \Phi_\omega(t \operatorname{tr} W_j, t^2 \det W_j)}{1 - \bar{\lambda}_i \lambda_j} \right]_{i,j=1}^n \geq 0. \tag{2.6}$$

On conjugating this matrix inequality by $\operatorname{diag}\{2 - \omega t \operatorname{tr} W_j\}$ and letting $\alpha = t\omega$, we obtain the following necessary condition for the solvability of a 2×2 spectral Nevanlinna-Pick condition [5, Theorem 5.2].

Theorem 2.5. *If there exists an analytic map $F : \mathbb{D} \to \Sigma$ satisfying the equations* (2.1) *and* (2.2) *then, for every $\alpha \in \Delta$,*

$$\left[\frac{\overline{(2-\alpha s_i)}(2-\alpha s_j) - |\alpha|^2\overline{(2\alpha p_i - s_i)}(2\alpha p_j - s_j)}{1-\bar\lambda_i\lambda_j}\right]_{i,j=1}^{n} \geq 0 \tag{2.7}$$

where

$$s_j = \operatorname{tr} W_j, \qquad p_j = \det W_j \quad \text{for } j = 1, \dots, n.$$

This condition is less simple than the classical Pick condition in that it comprises an infinite collection of algebraic inequalities, but it is nevertheless checkable in practice with the aid of standard numerical packages. Its major drawback is that it is *not* sufficient for solvability of the 2×2 spectral Nevanlinna-Pick problem.

Example 2.6. Let $0 < r < 1$ and let

$$\varphi(\lambda) = \left(2(1-r)\frac{\lambda^2}{1+r\lambda^3}, \frac{\lambda(\lambda^3+r)}{1+r\lambda^3}\right).$$

Pick any three distinct points $\mu_1, \mu_2, \mu_3 \in \mathbb{D}$ and let $\varphi(\mu_j) = (s_j, p_j)$ for $j = 1, 2, 3$. There exists $t > 1$ such that the inequality (2.7) holds for all $\alpha \in \Delta$ with $n = 3$ and $\lambda_j = t\mu_j$ but that there is no analytic map $h : \mathbb{D} \to \Gamma$ such that $h(\lambda_j) = (s_j, p_j)$ for $j = 1, 2, 3$.

Hence, if we choose nonscalar 2×2 matrices W_1, W_2, W_3 such that $(\operatorname{tr} W_j, \det W_j) = (s_j, p_j)$, then the spectral Nevanlinna-Pick problem with data $\lambda_j \mapsto W_j$ satisfies the necessary condition of Theorem 2.5 and yet has no solution.

The statement in the example will be proved in a future paper [3]; see also [22].

2.3. Two points and two-by-two matrices

When $n = k = 2$ the condition in Theorem 2.5 *is* sufficient for the solvability of the spectral Nevanlinna-Pick problem.

We shall now prove the main theorem from Section 1. Recall the statement:

Theorem 1.1. *Let $\lambda_1, \lambda_2 \in \mathbb{D}$ be distinct points, let W_1, W_2 be nonscalar 2×2 matrices of spectral radius less than 1 and let $s_j = \operatorname{tr} W_j$, $p_j = \det W_j$ for $j = 1, 2$. The following three statements are equivalent:*

(1) *there exists an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that*

$$F(\lambda_1) = W_1, \qquad F(\lambda_2) = W_2$$

and

$$r(F(\lambda)) \leq 1 \quad \text{for all } \lambda \in \mathbb{D};$$

(2)

$$\max_{\omega\in\mathbb{T}} \left|\frac{(s_2p_1 - s_1p_2)\omega^2 + 2(p_2 - p_1)\omega + s_1 - s_2}{(s_1 - \bar s_2 p_1)\omega^2 - 2(1 - p_1\bar p_2)\omega + \bar s_2 - s_1\bar p_2}\right| \leq \left|\frac{\lambda_1 - \lambda_2}{1 - \bar\lambda_2\lambda_1}\right|; \tag{2.8}$$

(3)

$$\left[\frac{\overline{(2-\omega s_i)}(2-\omega s_j)-\overline{(2\omega p_i-s_i)}(2\omega p_j-s_j)}{1-\bar\lambda_i\lambda_j}\right]_{i,j=1}^{2}\geq 0 \qquad (2.9)$$

for all $\omega\in\mathbb{T}$.

The proof depends on some elementary notions from the theory of invariant distances. A good source for the general theory is [35], but here we only need the following rudiments.

We denote by d the pseudohyperbolic distance on the unit disc $\mathbb{D}$:

$$d(\lambda_1,\lambda_2)=\left|\frac{\lambda_1-\lambda_2}{1-\bar\lambda_2\lambda_1}\right| \quad \text{for } \lambda_1,\lambda_2\in\mathbb{D}.$$

For any domain $\Omega\in\mathbb{C}^n$ we define the *Lempert function* $\delta_\Omega:\Omega\times\Omega\to\mathbb{R}^+$ by

$$\delta_\Omega(z_1,z_2)=\inf d(\lambda_1,\lambda_2) \qquad (2.10)$$

over all $\lambda_1,\lambda_2\in\mathbb{D}$ such that there exists an analytic map $h:\mathbb{D}\to\Omega$ such that $h(\lambda_1)=z_1$ and $h(\lambda_2)=z_2$. We define[1] the *Carathéodory distance* $C_\Omega:\Omega\times\Omega\to\mathbb{R}^+$ by

$$C_\Omega(z_1,z_2)=\sup d(f(z_1),f(z_2)) \qquad (2.11)$$

over all analytic maps $f:\Omega\to\mathbb{D}$. If Ω is bounded then C_Ω is a metric on Ω.

It is not hard to see (by the Schwarz-Pick Lemma) that $C_\Omega\leq\delta_\Omega$ for any domain Ω. The two quantities C_Ω, δ_Ω are not always equal – the punctured disc provides an example of inequality. The question of determining the domains Ω for which $C_\Omega=\delta_\Omega$ is one of the concerns of invariant distance theory.

Proof. Let $z_j=(s_j,p_j)\in\mathbb{G}$.
(1)⇔(2) In view of Proposition 2.2 we must show that the inequality (2.8) is equivalent to the existence of an analytic $h:\mathbb{D}\to\Gamma$ such that $h(\lambda_j)=z_j$ for $j=1,2$. By definition of the Lempert function $\delta_\mathbb{G}$, such an h exists if and only if

$$\delta_\mathbb{G}(z_1,z_2)\leq d(z_1,z_2).$$

By [11, Corollary 5.7] we have $\delta_\mathbb{G}=C_\mathbb{G}$, and by [11, Theorem 1.1 and Corollary 3.4],

$$\begin{aligned} C_\mathbb{G}(z_1,z_2)&=\max_{\omega\in\mathbb{T}} d(\Phi_\omega(z_1),\Phi_\omega(z_2)) \qquad (2.12)\\ &=\max_{\omega\in\mathbb{T}}\left|\frac{(s_2p_1-s_1p_2)\omega^2+2(p_2-p_1)\omega+s_1-s_2}{(s_1-\bar s_2p_1)\omega^2-2(1-p_1\bar p_2)\omega+\bar s_2-s_1\bar p_2}\right|. \end{aligned}$$

Thus the desired function h exists if and only if the inequality (2.8) holds.

(2)⇔(3) By equation (2.12), the inequality (2.8) is equivalent to

$$d(\Phi_\omega(z_1),\Phi_\omega(z_2))\leq d(\lambda_1,\lambda_2) \quad \text{for all } \omega\in\mathbb{T}.$$

[1] Conventionally the definition of the Carathéodory distance contains a $\tanh^{-1}$ on the right hand side of (2.11). For present purposes it is convenient to omit the $\tanh^{-1}$.

By the Schwarz-Pick Lemma, this inequality holds if and only if, for all $\omega \in \mathbb{T}$, there exists a function f_ω in the Schur class such that $f_\omega(\lambda_j) = \Phi_\omega(z_j)$ for $j = 1, 2$. By Pick's Theorem this in turn is equivalent to the relation

$$\left[\frac{1 - \bar{\Phi}_\omega(z_i)\Phi_\omega(z_j)}{1 - \bar{\lambda}_i\lambda_j}\right]_{i,j=1}^{2} \geq 0.$$

Conjugate by $\mathrm{diag}\{2 - \omega s_1, 2 - \omega s_2\}$ to obtain (2)⇔(3). □

Remark 2.7. If one removes the hypothesis that W_1, W_2 be nonscalar from Theorem 1.1 one can still give a solvability criterion. If both of the W_j are scalar matrices then the problem reduces to a scalar Nevanlinna-Pick problem. If $W_1 = cI$ and W_2 is nonscalar then the corresponding spectral Nevanlinna-Pick problem is solvable if and only if

$$r((W_2 - cI)(I - \bar{c}W_2)^{-1}) \leq d(\lambda_1, \lambda_2)$$

(see [7, Theorem 2.4]). This inequality can also be expressed as a somewhat cumbersome algebraic inequality in c, s_2, p_2 and $d(\lambda_1, \lambda_2)$ [7, Theorem 2.5(2)].

2.4. Ill-conditioned problems

The results of the preceding subsection suggest that solvability of spectral Nevanlinna-Pick problems depends on the derogatory structure of the target matrices – that is, in the case of 2×2 matrices, on whether or not they are scalar matrices. It is indeed so, and in consequence problems in which a target matrix is close to scalar can be very ill-conditioned.

Example 2.8. [7, Example 2.3] Let $\beta \in \mathbb{D} \setminus \{0\}$ and, for $\alpha \in \mathbb{C}$ let

$$W_1(\alpha) = \begin{bmatrix} 0 & \alpha \\ 0 & 0 \end{bmatrix}, \quad W_2 = \begin{bmatrix} 0 & \beta \\ 0 & \frac{2\beta}{1+\beta} \end{bmatrix}.$$

Consider the spectral Nevanlinna-Pick problem with data $0 \mapsto W_1(\alpha)$, $\beta \mapsto W_2$. If $\alpha = 0$ then the problem is not solvable. If $\alpha \neq 0$, however, by Proposition 2.2 the problem is solvable if and only if there exists an analytic function $f : \mathbb{D} \to \Gamma$ such that

$$f(0) = (0, 0) \text{ and } f(\beta) = \frac{2\beta}{1 + \beta}.$$

It may be checked [8] that

$$f(\lambda) = \left(\frac{2(1-\beta)\lambda}{1 - \beta\lambda}, \frac{\lambda(\lambda - \beta)}{1 - \beta\lambda}\right)$$

is such a function. Thus the problem has a solution F_α for any $\alpha \neq 0$. Consider a sequence (α_n) of nonzero complex numbers tending to zero: the functions F_{α_n} cannot be locally bounded, else they would have a cluster point, which would solve the problem for $\alpha = 0$. If α is, say, 10^{-100} then *any* numerical method for the spectral Nevanlinna-Pick problem is liable to run into difficulty in this example.

2.5. Uniqueness and the construction of interpolating functions

Problem SNP *never* has a unique solution. If F is a solution of Problem SNP then so is $P^{-1}FP$ for any analytic function $P : \mathbb{D} \to \mathbb{C}^{k\times k}$ such that $P(\lambda)$ is nonsingular for every $\lambda \in \mathbb{D}$ and $P(\lambda_j)$ is a scalar matrix for each interpolation point λ_j. There are always many such P that do not commute with F, save in the trivial case that F is scalar. Nevertheless, the solution of the corresponding interpolation problem for Γ *can* be unique. Consider again the case $n = k = 2$ with W_1, W_2 nonscalar. By Theorem 1.1, the problem is solvable if and only if inequality (2.8) holds. In fact it is solvable *uniquely* if and only if inequality (2.8) holds with equality. This amounts to saying that each pair of distinct points of $\mathbb{G}$ lies on a unique complex geodesic of $\mathbb{G}$, which is true by [12, Theorem 0.3]. (An analytic function $h : \mathbb{D} \to \mathbb{G}$ is a *complex geodesic* of $\mathbb{G}$ if h has an analytic left-inverse). Moreover, in this case the unique analytic function $h : \mathbb{D} \to \mathbb{G}$ such that $h(\lambda_j) = (s_j, p_j)$ for $j = 1, 2$ can be calculated explicitly as follows [11, Theorem 5.6].

Choose an $\omega_0 \in \mathbb{T}$ such that the maximum on the left hand side of (2.8) is attained at ω_0. Since equality holds in (2.8), we have

$$d(\Phi_{\omega_0}(z_1), \Phi_{\omega_0}(z_2)) = d(\lambda_1, \lambda_2),$$

where $z_j = (s_j, p_j)$. Thus Φ_{ω_0} is a Carathéodory extremal function for the pair of points z_1, z_2 in $\mathbb{G}$. It is easy (for example, by Schur reduction) to find the unique Blaschke product p of degree at most 2 such that

$$p(\lambda_1) = p_1, \quad p(\lambda_2) = p_2 \quad \text{and} \quad p(\bar{\omega}_0) = \bar{\omega}_0^2).$$

Define s by

$$s(\lambda) = 2\frac{\omega_0 p(\lambda) - \lambda}{1 - \omega_0\lambda} \text{ for } \lambda \in \mathbb{D}.$$

Then $h \stackrel{\text{def}}{=} (s, p)$ is the required complex geodesic.

Note that h is a rational function of degree at most 2. It can also be expressed in the form of a realization: $h(\lambda) = (\operatorname{tr} H(\lambda), \det H(\lambda))$ where H is a 2×2 function in the Schur class given by

$$H(\lambda) = D + C\lambda(1 - A\lambda)^{-1}B$$

for a suitable unitary 3×3 or 4×4 matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ given by explicit formulae (see [4], [12, Theorem 1.7]).

2.6. More points and bigger matrices

Our hope in addressing the case $n = k = 2$ of the spectral Nevanlinna-Pick problem was of course that we could progress to the general case. Alas, we have not so far managed to do so. We have some hope of giving a good solvability criterion for the case $k = 2$, $n = 3$, but even the case $n = 4$ appears to be too complicated for our present methods.

The case of two points and $k \times k$ matrices, for any k, looks at first sight more promising. There is an obvious way to generalize the symmetrized

bidisc: we define the *open symmetrized polydisc* $\mathbb{G}_k$ to be the domain

$$\mathbb{G}_k = \{(\sigma_1(z), \dots, \sigma_k(z)) : z \in \mathbb{D}^k\} \subset \mathbb{C}^k$$

where σ_m denotes the elementary symmetric polynomial in $z = (z^1, \dots, z^k)$ for $1 \leq m \leq k$. Similarly one defines the closed symmetrized polydisc Γ_k. As in the case $k = 2$, one can reduce Problem SNP to an interpolation problem for functions from $\mathbb{D}$ to Γ_k under mild hypotheses on the target matrices W_j (specifically, that they be nonderogatory). However, the connection between Problem SNP and the corresponding interpolation problems for Γ_k are more complicated for $k > 2$, because there are more possibilities for the rational canonical forms of the target matrices [37]. The analogues for Γ_k of the Φ_ω were described by D. J. Ogle [39] and subsequently other authors, e.g. [23, 29]. Ogle generalized to higher dimensions the operator-theoretic method of [6] and thereby obtained a necessary condition for solvability analogous to Theorem 2.5.

The solvability of Problem SNP when $n = 2$ is generically equivalent to the inequality

$$\delta_{\mathbb{G}_k}(z_1, z_2) \leq d(\lambda_1, \lambda_2)$$

where z_j is the $k$tuple of coefficients in the characteristic polynomial of W_j. All we need is an effective formula for $\delta_{\mathbb{G}_k}$. It turns out that this is a much harder problem for $k > 2$. In particular, it is *false* that $\delta_{\mathbb{G}_k} = C_{\mathbb{G}_k}$ when $k > 2$. This discovery [38] was disappointing, but not altogether surprising.

There is another type of solvability criterion for the 2×2 spectral Nevanlinna-Pick problem with general n [10, 14], but it involves a search over a nonconvex set, and so does not count for the purpose of this paper as an analytic solution of the problem. Another paper on the topic is [24].

It is heartening that the study of the complex geometry and analysis of the symmetrized polydisc has been taken up by a number of specialists in several complex variables, including G. Bharali, C. Costara, A. Edigarian, M. Jarnicki, L. Kosinski, N. Nikolov, P. Pflug, P. Thomas and W. Zwonek. Between them they have made many interesting discoveries about these and related domains. There is every hope that some of their results will throw further light on the spectral Nevanlinna-Pick problem.

3. The spectral Carathéodory-Fejér problem

This is the problem that arises from the spectral Nevanlinna-Pick problem when the interpolation points coalesce at 0.

Problem SCF *Given $k \times k$ matrices $V_0, V_1, \dots, V_n$, find an analytic function $F : \mathbb{D} \to \mathbb{C}^{k \times k}$ such that*

$$F^{(j)}(0) = V_j \quad \text{for } j = 1, \dots, n \tag{3.1}$$

and

$$r(F(\lambda)) \leq 1 \quad \text{for all } \lambda \in \mathbb{D}. \tag{3.2}$$

This problem also can be converted to an interpolation problem for analytic functions from $\mathbb{D}$ into Γ_k [34, Theorem 2.1], [37]. However, the resulting problem is again hard when $k \geq 2$, and the only truly explicit solution we have is in the case $k = 2, n = 1$ [34, Theorem 1.1].

Theorem 3.1. *Let*

$$V_m = [v_{ij}^m]_{i,j=1}^2 \quad \text{for } m = 0, 1$$

and suppose that V_0 is nonscalar. There exists an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that

$$F(0) = V_0, \quad F'(0) = V_1 \quad \text{and} \quad r(F(\lambda)) < 1 \text{ for all } \lambda \in \mathbb{D} \tag{3.3}$$

if and only if

$$\max_{|\omega|=1} \left| \frac{(s_1 p_0 - s_0 p_1)\omega^2 + 2\omega p_1 - s_1}{\omega^2(s_0 - \bar{s}_0 p_0) - 2\omega(1 - |p_0|^2) + \bar{s}_0 - s_0 \bar{p}_0} \right| \leq 1, \tag{3.4}$$

where

$$s_0 = \operatorname{tr} V_0, \quad p_0 = \det V_0,$$
$$s_1 = \operatorname{tr} V_1, \quad p_1 = \begin{vmatrix} v_{11}^0 & v_{12}^1 \\ v_{21}^0 & v_{22}^1 \end{vmatrix} + \begin{vmatrix} v_{11}^1 & v_{12}^0 \\ v_{21}^1 & v_{22}^0 \end{vmatrix}.$$

The proof of this theorem in [34] again depends on the calculation in [11] of the Carathéodory metric on $\mathbb{G}$, but this time on the infinitesimal version $c_{\mathbb{G}}$ of the metric: the left hand side of inequality (3.4) is the value of $c_{\mathbb{G}}$ at (s_0, p_0) in the direction (s_1, p_1). This fact is [11, Corollary 4.4], but unfortunately there is an ω missing in the statement of Corollary 4.4. The proof shows that the correct formula is as in (3.4). An important step is the proof that the infinitesimal Carathéodory and Kobayashi metrics on $\mathbb{G}$ coincide.

The ideas behind Theorem 3.1 can be used to find solutions of Problem SCF: see [34, Section 6]. The ideas can also be used to derive a necessary condition for the spectral Carathéodory-Fejér problem (3.1), (3.2) in the case that $n = 1$ and $k > 2$ [34, Theorem 4.1], but there is no reason to expect this condition to be sufficient.

4. The structured singular value

The *structured singular value* of a matrix relative to a space of matrices was introduced by J. C. Doyle and G. Stein in the early 1980s [25, 26] and was denoted by μ. It is a refinement of the usual operator norm of a matrix and is motivated by the problem of the robust stabilization of a plant that is subject to structured uncertainty. Initially, in the H^∞ approach to robustness, the uncertainty of a plant was modelled by a meromorphic matrix function (on a disc or half plane) that is subject to an L^∞ bound but is otherwise completely unknown. The problem of the simultaneous stabilization of the resulting collection of plant models could then be reduced to some classical

analysis and operator theory, notably to the far-reaching results of Adamyan, Arov and Krein from the 1970s [30].

In practice one may have some structural information about the uncertainty in a plant – for example, that certain entries are zero. By incorporating such structural information one should be able to achieve a less conservative stabilizing controller. The structured singular value was devised for this purpose. A good account of these notions is in [27, Chapter 8]. Unfortunately, the behaviour of μ differs radically from that of the operator norm – for one thing, μ is not in general a norm at all, and none of the relevant classical theorems (such as Pick's theorem) or methods appear to extend to the corresponding questions for μ. This provides a challenge for mathematicians: we should help out our colleagues in engineering by creating an AAK-type theory for μ.

For any $A \in \mathbb{C}^{k\times\ell}$ and any subspace E of $\mathbb{C}^{\ell\times k}$ we define the structured singular value $\mu_E(A)$ by

$$\frac{1}{\mu_E(A)} = \inf\{\|X\| : X \in E,\ 1 - AX \text{ is singular}\} \tag{4.1}$$

with the understanding that $\mu_E(A) = 0$ if $1 - AX$ is always nonsingular.

Two instances of the structured singular value are the operator norm $\|.\|$ (relative to the Euclidean norms on $\mathbb{C}^k$ and $\mathbb{C}^\ell$) and the spectral radius r. If we take $E = \mathbb{C}^{\ell\times k}$ then we find that $\mu_E(A) = \|A\|$. On the other hand, if $k = \ell$ and we choose E to be the space of scalar multiples of the identity matrix, then $\mu_E(A) = r(A)$. These two special μs are in a sense extremal: it is always the case, for any E, that $\mu_E(A) \le \|A\|$. If $k = \ell$ and E contains the identity matrix, then $\mu_E(A) \ge r(A)$. A comprehensive discussion of the properties of μ can be found in [40].

Here is a formulation of the *μ-synthesis problem* [26, 27].

Given positive integers k, ℓ, a subspace E of $\mathbb{C}^{\ell\times k}$ and analytic functions A, B, C on $\mathbb{D}$ of types $k\times\ell, k\times k$ and $\ell\times\ell$ respectively, construct an analytic function $F : \mathbb{D} \to \mathbb{C}^{k\times\ell}$ of the form

$$F = A + BQC \quad \text{for some analytic} \quad Q : \mathbb{D} \to \mathbb{C}^{k\times\ell} \tag{4.2}$$

such that

$$\mu_E(F(\lambda)) \le 1 \quad \text{for all} \quad \lambda \in \mathbb{D}. \tag{4.3}$$

The condition (4.2), that F be expressible in the form $A + BQC$ for some analytic Q, can be regarded as an interpolation condition on F. In the event that $k = \ell$, B is the scalar polynomial

$$B(\lambda) = (\lambda - \lambda_1)\dots(\lambda - \lambda_n)I$$

with distinct zeros $\lambda_j \in \mathbb{D}$ and C is constant and equal to the identity, then F is expressible in the form $A + BQC$ if and only if

$$F(\lambda_1) = A(\lambda_1), \dots, F(\lambda_n) = A(\lambda_n).$$

With this choice of B and C, if we take E to be the space of scalar matrices, we obtain precisely the spectral Nevanlinna-Pick problem. If we now replace B by the polynomial λ^n, we get the spectral Carathéodory-Fejér problem.

In engineering applications μ-synthesis problems arise after some analysis is carried out on the plant model to produce the A, B and C in condition (4.2), and the resulting B and C will not usually be scalar functions. Nevertheless, explicit pointwise interpolation conditions provide a class of easily-formulated test cases, and it is arguable that such problems are the *hardest* cases of μ-synthesis.

Conditions of the form (4.2) are said to be of *model matching type* [30].

The most sustained attempt to develop an AAK-type theory for the structured singular value in full generality is due to H. Bercovici, C. Foiaş and A. Tannenbaum ([15] to[21]). They have a far-reaching theory: *inter alia* they have constructed many illuminating examples, found properties of extremal solutions and obtained a type of solvability criterion for μ-synthesis problems. The criterion results from a combination of the Commutant Lifting Theorem with the application of similarity transformations. To apply the criterion to a concrete spectral Nevanlinna-Pick problem one must solve an optimization problem over a high-dimensional unbounded and non-convex set. We can certainly hope that this is not the last word on the subject of solvability. Despite the achievements of Bercovici, Foiaş and Tannenbaum, there is still plenty of room for further study of μ-synthesis.

One of their examples [18, Section 7, Example 5] exhibits an important fact about the spectral Nevanlinna-Pick problem: *diagonalization does not work*. It shows that diagonalization of the target matrices W_j in Problem SNP by similarity transformations, even when possible, does not help solve the problem. One could hope that if the W_j were diagonal one might be able to decouple the problem into a series of scalar interpolation problems, but they show that such a hope is vain.

Bill Helton himself, along with collaborators, has developed an alternative approach to the refinement of H^∞ control; his viewpoint is set out in [32]. His part in the introduction of the results of Adamyan, Arov, Krein and other operator-theorists into robust control theory in the early 1980s is well known. He subsequently worked extensively (with Orlando Merino, Trent Walker and others) during the 1990s on the more delicate optimization problems that arise from refinements of the basic H^∞ picture of modelling uncertainty. As in the μ approach, the aim is to incorporate more subtle specifications and robustness conditions into methods for controller design. He developed a very flexible formulation of such problems as optimization problems over spaces of vector-valued analytic functions on the disc, and devised an algorithm for their numerical solution – see [33] and several other papers. The authors proved convergence results and described numerical trials. However, the spectral Nevanlinna-Pick problem cannot be satisfactorily treated by the Helton scheme. Although it can be cast in the basic problem

formulation [32, Chapter 2], solution algorithms require smoothness properties (of the function "Γ") which the spectral radius does not possess.

5. The next case of μ

After the two extremes $\mu = \|.\|_{H^\infty}$ and $\mu = r$ the next natural case to consider is the one in which, in (4.1), $k = \ell$ and E is the space $\mathrm{Diag}(k)$ of diagonal matrices. For the rest of this section μ will denote $\mu_{\mathrm{Diag}(2)}$ and we shall study the following problem:

Given distinct points $\lambda_1, \dots, \lambda_n \in \mathbb{D}$ and 2×2 matrices $W_1, \dots, W_n$, construct an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that

$$F(\lambda_j) = W_j \quad \text{for } j = 1, \dots, n \tag{5.1}$$

and

$$\mu(F(\lambda)) \le 1 \quad \text{for all } \lambda \in \mathbb{D}. \tag{5.2}$$

For the 2×2 spectral Nevanlinna-Pick problem we had some modest success through reduction to an interpolation problem for Γ-valued functions. In the present case we tried an analogous approach, with still more modest success [1, 2, 41]. The following result is [2, Theorem 9.4 and Remark 9.5(iii)].

Theorem 5.1. *Let $\lambda_0 \in \mathbb{D}$, $\lambda \neq 0$, let $\zeta \in \mathbb{C}$ and let*

$$W_1 = \begin{bmatrix} 0 & \zeta \\ 0 & 0 \end{bmatrix}, \qquad W_2 = \begin{bmatrix} a & * \\ * & b \end{bmatrix}. \tag{5.3}$$

Suppose that $|b| \le |a|$ and let $p = \det W_2$. There exists an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that

$$F(\lambda_1) = W_1, \quad F(\lambda_2) = W_2 \quad \text{and} \quad \mu(F(\lambda)) \le 1 \quad \text{for all } \lambda \in \mathbb{D} \tag{5.4}$$

if and only if $|p| < 1$ and

$$\begin{cases} \dfrac{|a - \bar{b}p| + |ab - p|}{1 - |p|^2} \le |\lambda_0| & \text{if } \zeta \neq 0 \\ \\ |\lambda_0|^4 - (|a|^2 + |b|^2 + 2|ab - p|)|\lambda_0|^2 + |p|^2 \ge 0 & \text{if } \zeta = 0. \end{cases}$$

The stars in the formula for W_2 in (5.3) denote arbitrary complex numbers.

What is the analog of Γ for this case of μ? To determine whether a 2×2 matrix $A = [a_{ij}]$ satisfies $r(A) \le 1$ one needs to know only the two numbers $\mathrm{tr}\, A$ and $\det A$; this fact means that the spectral Nevanlinna-Pick problem can generically be reduced to an interpolation problem for Γ. To determine whether $\mu(A) \le 1$ one needs to know the three numbers $a_{11}, a_{22}, \det A$. This led us to introduce a domain $\mathbb{E}$ which we call the *tetrablock*:

$$\mathbb{E} = \{x \in \mathbb{C}^3 : 1 - x^1 z - x^2 w + x^3 z w \neq 0 \text{ whenever } |z| \le 1, |w| \le 1\}. \tag{5.5}$$

Its closure is denoted by $\bar{\mathbb{E}}$. The name reflects the fact that the intersection of $\mathbb{E}$ with $\mathbb{R}^3$ is a regular tetrahedron. The domain $\mathbb{E}$ is relevant because

$\mu(A) < 1$ if and only if $(a_{11}, a_{22}, \det A) \in \mathbb{E}$. There exists a solution of the 2-point μ-synthesis problem (5.4) if and only if the corresponding interpolation problem for analytic functions from $\mathbb{D}$ to $\mathbb{E}$ is solvable [2, Theorem 9.2], and accordingly the solvability problem for this μ-synthesis problem is equivalent to the calculation of the Lempert function $\delta_{\mathbb{E}}$. As far as I know no one has yet computed $\delta_{\mathbb{E}}$ for a general pair of points of $\mathbb{E}$, but we did calculate it in the case that one of the points is the origin in $\mathbb{C}^3$, that is, we proved a Schwarz lemma for $\mathbb{E}$. The result is Theorem 5.1.

Observe that ill-conditioning appears in this instance of μ-synthesis too [2, Remark 9.5(iv)]. If, in Theorem 5.1, $a = b = p = \frac{1}{2}$ then there exists a solution F_ζ of the problem if and only if

$$|\lambda_0| \geq \begin{cases} \frac{2}{3} & \text{if } \zeta \neq 0 \\ \\ \frac{1}{\sqrt{2}} & \text{if } \zeta = 0 \end{cases}$$

Thus if $\frac{2}{3} < |\lambda_0| < \frac{1}{\sqrt{2}}$, the F_ζ are not locally bounded as $\zeta \to 0$, and so are sensitive to small changes in ζ near 0.

The complex geometry of $\mathbb{E}$ has also proved to be of interest to researchers in several complex variables. To my surprise, it was recently shown [28] that the Lempert function and the Carathéodory distance on $\mathbb{E}$ coincide. This might be a step on the way to the derivation of a formula for $\delta_{\mathbb{E}}$. It would suffice to compute $\delta_{\mathbb{E}}$ in the case that one of the two points is of the form $(0, 0, \lambda)$ for some $\lambda \in [0, 1)$, since every point of $\mathbb{E}$ is the image of such a point under an automorphism of $\mathbb{E}$ [41, Theorem 5.2].

The fourth and final special case of μ-synthesis in this paper is the μ-analog of the 2×2 Carathéodory-Fejér problem:

Given 2×2 matrices $V_0, \ldots, V_n$, construct an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that

$$F^{(j)}(0) = V_j \text{ for } j = 0, \ldots, n \quad \text{and} \quad \mu(F(\lambda)) \leq 1 \quad \text{for all } \lambda \in \mathbb{D}.$$

Again the problem can be reduced to an interpolation problem for $\mathbb{E}$, but the resulting problem has only been solved in an exceedingly special case.

Theorem 5.2. *Let V_0, V_1 be 2×2 matrices such that*

$$V_0 = \begin{bmatrix} 0 & \zeta \\ 0 & 0 \end{bmatrix}$$

for some $\zeta \in \mathbb{C}$ and $V_1 = [v_{ij}]$ is nondiagonal. There exists an analytic function $F : \mathbb{D} \to \mathbb{C}^{2\times 2}$ such that

$$F(0) = V_0, \quad F'(0) = V_1 \quad \text{and} \quad \mu(F(\lambda)) \leq 1 \text{ for all } \lambda \in \mathbb{D}$$

if and only if

$$\max\{|v_{11}|, |v_{22}|\} + |\zeta v_{21}| \leq 1.$$

This result follows from [41, Theorem 2.1].

6. Conclusion

Although μ-analysis remains a useful tool, it is fair to say that μ-synthesis, as a major technique for robust control system design, has been something of a disappointment up to now. The trouble is that the μ-synthesis problem is difficult. It is a highly non-convex problem. There do exist heuristic numerical methods for addressing particular μ-synthesis problems, notably a Matlab toolbox [36] based on the "DK algorithm" [27, Section 9.3], but there is no practical solvability criterion, no fast algorithm nor any convergence theorem for any known algorithm. For these reasons engineers have largely turned to other approaches to robust stabilization over the past 20 years. If, however, a satisfactory analytic theory of the problem is developed, engineers' attention may well return to μ-synthesis as a promising design tool. We are still far from having such a theory, but perhaps these special cases and the interest of the several complex variables community may yet lead to one.

References

[1] A. A. Abouhajar, *Function theory related to H^∞ control,* Ph.D. thesis, Newcastle University, 2007.

[2] A. A. Abouhajar, M. C. White and N. J. Young, A Schwarz lemma for a domain related to mu-synthesis, *J. Geometric Analysis* **17** (2007) 717-750.

[3] J. Agler, Z. A. Lykova and N. J. Young, in preparation.

[4] J. Agler, F. B. Yeh and N. J. Young, Realization of functions into the symmetrized bidisc, in *Reproducing Kernel Spaces and Applications*, ed. D. Alpay, *Operator Theory: Advances and Applications* **143**, Birkhäuser Verlag (2003), 1-37.

[5] J. Agler and N. J. Young, A commutant lifting theorem for a domain in $\mathbb{C}^2$ and spectral interpolation, *J. Functional Analysis* **161** (1999) 452-477.

[6] J. Agler and N. J. Young, Operators having the symmetrized bidisc as a spectral set, *Proc. Edin. Math. Soc.* **43** (2000) 195-210.

[7] J. Agler and N. J. Young, The two-point spectral Nevanlinna-Pick problem, *Integral Equations Operator Theory* **37** (2000) 375–385.

[8] J. Agler and N. J. Young, A Schwarz lemma for the symmetrised bidisc, *Bull. London Math. Soc.* **33** (2001) 175–186.

[9] J. Agler and N. J. Young, A model theory for Γ-contractions, *J. Operator Theory* **49** (2003) 45-60.

[10] J. Agler and N. J. Young, The two-by-two spectral Nevanlinna-Pick problem, *Trans. Amer. Math. Soc.* **356** (2004) 573-585.

[11] J. Agler and N. J. Young, The hyperbolic geometry of the symmetrized bidisc, *J. Geom. Anal.* **14** (2004) 375-403.

[12] J. Agler and N. J. Young, The complex geodesics of the symmetrized bidisc, *International J. Math.* **17** (2006) 375-391.

[13] J. Agler and N. J. Young, The magic functions and automorphisms of a domain,*Complex Analysis and Operator Theory* **2** (2008) 383-404.

[14] H. Bercovici, Spectral versus classical Nevanlinna-Pick interpolation in dimension two, *Electronic Journal of Linear Algebra* **10** (2003), 60–64.

[15] H. Bercovici, C. Foiaş, P. P. Khargonekar and A. Tannenbaum, On a lifting theorem for the structured singular value, *J. Math. Analysis Appl.* **187** (1994) 617-627.

[16] H. Bercovici, C. Foiaş, and A. Tannenbaum, Spectral variants of the Nevanlinna-Pick interpolation problem, commutant lifting theorem, *Signal processing, scattering and operator theory, and numerical methods*, Progr. Systems Control Theory, Vol. 5, Birkhäuser, Boston, 1990, pp. 23–45.

[17] H. Bercovici, C. Foiaş and A. Tannenbaum, On the optimal solutions in spectral commutant lifting theory, *J. Functional Analysis* **101** (1991) 38-49.

[18] H. Bercovici, C. Foiaş and A. Tannenbaum, A spectral commutant lifting theorem, *Trans. Amer. Math. Soc.* **325** (1991) 741-763.

[19] H. Bercovici, C. Foiaş and A. Tannenbaum, On spectral tangential Nevanlinna-Pick interpolation, *J. Math. Analysis Appl.* **155** (1991) 156-176.

[20] H. Bercovici, C. Foiaş and A. Tannenbaum, Structured interpolation theory, *Operator Theory: Advances and Applications* **47** (1992) 195-220.

[21] H. Bercovici, C. Foiaş and A. Tannenbaum, The structured singular value for linear input-output operators, *SIAM J. Control Optimization* **34** (1996).

[22] G. Bharali, Some new observations on interpolation in the spectral unit ball, *Integral Eqns. Operator Theory* **59** (2007) 329-343.

[23] C. Costara, *Le problème de Nevanlinna-Pick spectrale*, Ph.D. thesis, Université Laval, Quebec City, Canada, 2004.

[24] C. Costara, The 2×2 spectral Nevanlinna–Pick problem, *J. London Math. Soc.* **71** (2005) 684-702.

[25] J. C. Doyle, Analysis of feedback systems with structured uncertainty, *IEE Proceedings* **129** (1982) 242-250.

[26] J. C. Doyle and G. Stein, Multivariable feedback design: concepts for a classical/modern synthesis, *IEEE Transactions on Automatic Control*, **26** (1981) 4-16.

[27] G. Dullerud and F. Paganini, *A course in robust control theory: a convex approach*, Texts in Applied Mathematics **36**, Springer (2000).

[28] A. Edigarian, L. Kosinski and W. Zwonek, The Lempert theorem and the tetrablock, arXiv:1006.4883 .

[29] A. Edigarian and W. Zwonek, Geometry of the symmetrised polydisc, *Archiv Math.*, **84** (2005) 364-374.

[30] B. A. Francis, *A Course in H_∞ Control Theory*, Lecture Notes in Control and Information Sciences No. 88, Springer Verlag, Heidelberg, 1987.

[31] J. W. Helton, Orbit structure of the Möbius transformation semigroup action on H^∞ (broadband matching), *Adv. in Math. Suppl. Stud.* **3**, Academic Press, New York (1978), 129 – 197.

[32] J. W. Helton, *Operator theory, analytic functions, matrices, and electrical engineering* CBMS Regional Conference Series in Mathematics No. 68, AMS, Providence RI, 1987.

[33] J. W. Helton, O. Merino and T. Walker, Algorithms for optimizing over analytic functions, *Indiana Univ. Math. J.* **42** (1993) 839-874.

[34] H-N. Huang, S. Marcantognini and N. J. Young, The spectral Carathéodory-Fejér problem, *Integral Equations and Operator Theory* **56** (2006) 229-256.

[35] M. Jarnicki and P. Pflug, Invariant distances and metrics in complex analysis revisited, *Dissertationes Math. (Rozprawy Mat.)* **430** (2005) 1-192.

[36] *Matlab μ-Analysis and Synthesis Toolbox*, The Math Works Inc., Natick, Massachusetts, http://www.mathworks.com/products/muanalysis/ .

[37] N. Nikolov, P. Pflug and P. J. Thomas, Spectral Nevanlinna-Pick and Carathéodory-Fejér problems, to appear in *Indiana Univ. Math. J.*, arXiv:1002.1706 .

[38] N. Nikolov, P. Pflug and W. Zwonek, The Lempert function of the symmetrized polydisc in higher dimensions is not a distance, *Proc. Amer. Math. Soc.* **135** (2007) 2921–2928.

[39] D. Ogle, Operator and Function Theory of the Symmetrized Polydisc, Ph. D. thesis, Newcastle University (1999), http://www.maths.leeds.ac.uk/ nicholas/abstracts/ogle.html .

[40] A. Packard and J. C. Doyle, The complex structured singular value, *Automatica* **29** (1993) 71-109.

[41] N. J. Young, The automorphism group of the tetrablock, *J. London Math. Soc.* **77** (2008) 757-770.

國科會補助專題研究計畫項下出席國際學術會議心得報告

日期：101年10月30日

計畫編號	NSC 100－2115－M－029－003－		
計畫名稱	對稱多盤上之頻譜 Caratheodory-Fejer 插值函數		
出國人員姓名	黃皇男	服務機構及職稱	東海大學數學系教授
會議時間	100 年 10 月 14 日 至 100 年 10 月 16 日	會議地點	Dalian, China
會議名稱	(中文)第七屆國際智慧訊息隱藏與多媒體訊號處理研討會 (英文) The Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP-2011)		
發表論文題目	(中文)數位水印之小波域熵 (英文) Wavelet-Based Entropy for Digital Audio Watermarking		

會議時間	100 年 7 月 7 日至 100 年 7 月 9 日	會議地點	Taiyuan, China
會議名稱	(中文)2012 國際計算、測量、控制與感測網路研討會 (英文) 2012 International Conference on Computing, Measurement, Control and Sensor Network (CMCNS 2012)		
發表論文題目	1. (中文)連續脈壓波之無失真特徵壓縮法 (英文) A Lossless Characteristic Compression Method for Continuous Arterial Pulse Waveforms 2. (中文) 統計 ECG 心電訊號識別小波域之最佳化影像浮水印設計 (英文) ECG Human Identification with Statistical Support Vector Machines 3. (中文)ECG 心電訊號之小波量水印設計 (英文) Wavelet-based Quantization Watermarking for ECG Signals		

今年經費參加兩個國際研討會，分別說明如下。

1. 參加 2011 年 10 月 14 日至 10 月 16 日於 Dalian, China 由 IEEE Signal Processing Society 共同主辦舉行之 The Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP-2011)，並發表論文。

2. 參加 2012 年 7 月 7 日至 7 月 9 日於 Taiyuan, China 由 IEEE Signal Processing Society 共同主辦舉行之 The International Conference on

Computing, Measurement, Control and Sensor Network (CMCNS 2012)，並發表論文。

一、IIH-MSP-2011 研討會

(一)、參加會議經過

此次會議由 IEEE 訊號處理社群、大連理工學院、高雄應用科技大學等單位共同支助，會議的目的為提供給專家與學者共同討論以下主題：

Track I: Information Hiding and Security

Track II: Multimedia Signal Processing and Networking

Track III: Bio-Inspired Multimedia Technologies and Systems

本人發表的論文名稱為Wavelet-Based Entropy for Digital Audio Watermarking安排在Session A02: Techniques and Algorithms for Multimedia Security, 12/14, 上午10:00-12:00，收錄到EI index的會議論文集當中，論文如附錄所示。

最後要再次感謝國科會的經費補助。為本次會議之議程，請參考會議網址：
http://bit.kuas.edu.tw/~iihmsp11/

(二)、與會心得

此次會議當中邀請到的 Keynote Speaker 有：**Chin-Chen Chang, Jar-Ferr Yang, Volker Roth** 等專家的演講，了解訊息隱藏的最新發展，相當有趣。

(三)、攜回資料名稱及內容

會議手冊一份以及論文集光碟片一份。

二、CMCNS 2012 研討會

(一)、參加會議經過

此次會議由 IEEE 訊號處理社群、太原理工學院、高雄應用科技大學等單位共同辦理，會議的目的為提供給兩岸專家與學者在計算、量測、控制、與網路方面的共同討論的機會，促進國際學術交流與合作機會等，主題涵蓋：

A: Computing and Signal Processing

B: Measurement

C: Control

D: Sensor Network

本人負責一個Session，並發表的數篇，論文全文如附錄所示。

最後要再次感謝國科會的經費補助。為本次會議之議程，請參考會議網址：

http://isic2011.nedu.edu.cn/

（二）、與會心得

此次會議當中邀請到的 Keynote Speaker 有二位：**Chin-Chen Chang、Qingsheng Zeng** 的演講，了解從大型動態系統的智慧觀測與控制，以及多目標最佳化的解決算法。

（三）、攜回資料名稱及內容

會議手冊一份以及論文集光碟片一份。

附錄1

Wavelet-Based Entropy for Digital Audio Watermarking

Shuo-Tsung Chen
Department of Mathematics,
Tunghai University,
Taichung 407, Taiwan (ROC).
shough33@yahoo.com.tw

Huang-Nan Huang
Department of Mathematics,
Tunghai University,
Taichung 407, Taiwan (ROC).
nhuang@thu.edu.tw
(corresponding author)

Chih-Yu Hsu
Department of Information and
Communication Engineering,
Chaoyang University of Technology,
Taichung County 413, Taiwan (ROC).
tccnchsu@gmail.com

Kuo-Kun Tseng
Department of Computer Science
and Technology, Harbin Institute
of Technology, Shenzhen
Graduate School, China.
kuokun.tseng@googlemail.com

Chun-Hua Wu
Chang Bing Show Chwan Memorial
Hospital, Lukang, Changhua County
505, Taiwan (ROC).
anis6699@yahoo.com.tw

Jeng-Shyang Pan
Department of Computer Science
and Technology, Harbin Institute
of Technology, Shenzhen
Graduate School, China.
jengshyangpan@gmail.com

***Abstract*—Unlike traditional entropy in information theory, this work uses the normalized energy instead of probability to obtain a low-frequency amplitude transform (LAT) on coefficients of discrete wavelet transform (DWT). The watermark is embedded based on the properties and characteristics of this transform. Finally, performance of the proposed scheme is assessed by signal-to-watermark (SWR) and bit error rate (BER). Experimental results demonstrate that the embedded data are robust against most signal processing and attacks, such as re-sampling, low-pass filtering, and amplitude-scaling.**

Keywords-entropy; low-frequency amplitude transform; discrete wavelet transform

I. INTRODUCTION

In recent years, many watermarking techniques have been proposed [1-8]. For music copyright protection, audio watermarking has the following requirements: 1) The watermark should be imperceptible in embedded audio. 2) The embedding technique should offer more than 20 dB signal to watermark ratio (SWR). 3) The watermark should prevent common attacks, including filtering, re-sampling, and mp3 compression, etc.

Wu *et al.* [2] used quantization index modulation to embed information into the low-frequency sub-band coefficients of discrete wavelet transform (DWT). This technique has good watermarked audio quality and strong robustness against common signal processing and noise corruption. However, this method is vulnerable to amplitude and time scaling. Xiang *et al.* [3] proposed a DWT-based audio watermarking algorithm robust against the DA/AD conversions. The relative energy relation among different groups of the DWT coefficients in the low-frequency sub-band are utilized in embedding by adaptively controlling the embedding strength. However, the method has low capacity and SNR. Chen *et al.* [4] proposed an optimization-based watermarking scheme robustly against many attacks.

Unlike traditional entropy, this work uses normalized energy instead of probability to form a novel entropy. Based on this concept, this work presents a new technique that embeds information by using low-frequency amplitude transform (LAT). Some properties and characteristic curve of LAT are analyzed and proved to investigate the relationship between LAT and DWT coefficients. Finally, the performance of the proposed scheme is assessed by signal-to-watermark ratio (SWR) and bit error rate (BER). Experimental results demonstrate that the embedded data are robust against most signal processing and attacks.

The remainder of this paper is organized as follows. Section II introduces DWT and LAT. Section III derives the properties and characteristic curve of LAT to analyze the relationship between LAT and DWT coefficients. The proposed embedding and extraction processes are described in Section IV. Experiments are conducted to test the performance of our proposed method in Section V. Finally, conclusions are summarized in Section VI.

II. DWT AND LAT

Discrete wavelet transform (DWT) is first reviewed in this section. Based on the low-frequency DWT coefficients in level seven, which is also referred as the lowest-frequency DWT coefficients, traditional entropy is redefined as a novel low-frequency amplitude transform (LAT).

A. *Discrete-time wavelet transform (DWT)*

Since the conventional fast Fourier transform (FFT) efficiently decomposes a signal into uniform-resolution analysis, it is suitable to analyze the wide-sense-stationary condition but not in non-stationary signal. In this paper, the discrete wavelet transform (DWT) is adapted to decompose the signal into the time-frequency domain. According to the

multi-resolution property of DWT, it leads to low-frequency but high-temporal resolution in high frequency bands and low-temporal but high-frequency resolution in low frequency bands. Therefore, we let the low frequency bands enhance the periodic property by only decomposing low-frequency band in each level. In [9], a method to implement DWT by using filter bank decomposition is proposed.

B. LAT

Before to introduce the proposed watermarking technique, the LAT must be defined and discussed. If there are N non-negative random samples that are shown as $\tilde{X}_N = \{c_i \mid 0 \le i \le N-1\}$, the corresponding probabilities are $P(c_0) = p_0$, $P(c_1) = p_1$, ..., $P(c_{N-1}) = p_{N-1}$. Based on information theory, the entropy of these samples is defined as follows:

$$H_r(\tilde{X}_N) = -\sum_{i=0}^{N-1} p_i \log_r p_i,\ 0 \le i \le N\text{-}1 \qquad (1)$$

where r is a base number of the logarithm function $\log$. This work adopts $r = 10$. Unlike the traditional way, this work uses the normalized energy instead of probability in wavelet domain as follows.

Definition 1. *Suppose that* $X_N = \{|c_i| \mid 0 \le i \le N-1\}$ *is a set of low-frequency coefficients in DWT, low-frequency amplitude transform (LAT) of* X_N *is then defined as*

$$LAT(X_N) = -\sum_{k=0}^{N-1} \left(\frac{|c_k|}{\sum_{j=0}^{N-1} |c_j|} \right) \log \left(\frac{|c_k|}{\sum_{j=0}^{N-1} |c_j|} \right) \qquad (2)$$

where $\dfrac{|c_k|}{\sum_{j=0}^{N-1} |c_j|}$ *is the normalized energy of coefficient* c_k.

From this definition, if the variation of low-frequency coefficients X_N is small, the corresponding $LAT(X_N)$ is big. For an example with $N = 20$ as shown in Figure 1, when X_{20} varies slowly, $LAT(X_{20})$ is 1.2998. However, when the variation of low-frequency coefficients X_N is large, the corresponding $LAT(X_N)$ is small. Figure 2 depicts that when X_{20} ($N = 20$) varies markedly, the corresponding $LAT(X_{20})$ is 1.1589. In this paper, we adopt $N = 2$ to have high embedding payload. Moreover, the standard deviations of this function for $N = 2$ before and after various attacks are approximately invariant. It is expected that the proposed low-frequency amplitude transform is robust against common attacks.

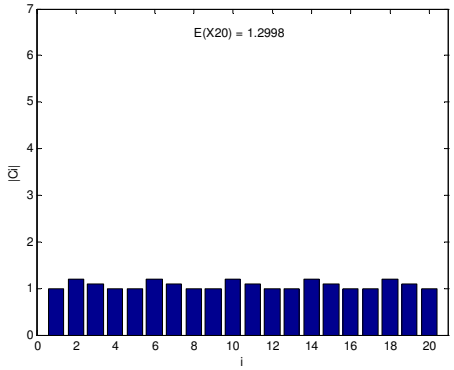


Fig. 1 When X_{20} varies slowly, $LAT(X_{20})$ is 1.2998.

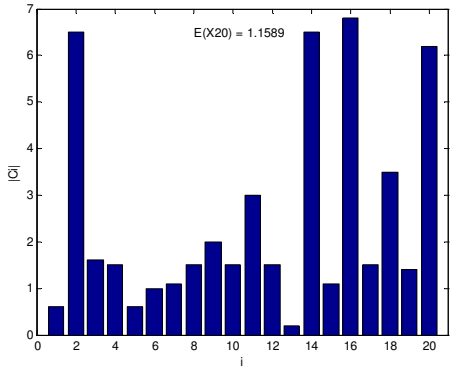


Fig. 2 When X_{20} varies markedly, $LAT(X_{20})$ is 1.1589.

Finally, the performance of our method is measured by signal-to-watermark ratio (SWR) and Bit-error-rate (BER) mathematically. They are defined as follows.

$$\text{SWR} = 10 \log_{10} \left\{ \frac{\sum_n (S(n))^2}{\sum_n (S(n) - \overline{S}(n))^2} \right\}, \qquad (3)$$

where $S(n)$ and $\overline{S}(n)$ denote the original and the modified audio, respectively;

$$\text{BER} = \left(B_{error} / B_{total} \right) \times 100\%, \qquad (4)$$

where B_{error} and B_{total} denote the number of error bits and the number of total bits, respectively.

III. PROPERTIES OF THE CHARACTERISTIC CURVE OF LAT

This section discusses some properties of the characteristic curve of LAT (CCL). Based on these properties and the characteristic curve, this work presents a novel watermarking scheme.

A. Properties of CCL

In the proposed watermarking scheme, the host digital audio signal $S(n)$ is cut into segments. Then, every two low-frequency DWT coefficients in each segment are grouped and sorted according to their absolute value into a vector form $X_2 = [|c_0|, |c_1|]$, where $|c_0| < |c_1|$. Since the value of

$LAT(X_2)$ in (2) is a function of X_2, a weighting matrix W is used to control the variation of $LAT(X_2)$ as follows:

$$\hat{X}_2 = X_2 W \tag{5}$$

where

$$W = \begin{pmatrix} w_0 & 0 \\ 0 & 1 \end{pmatrix} = diag(w_0, 1) \tag{6}$$

In other words, only the smallest value $|c_0|$ will be modified. Hence, the corresponding $LAT(\hat{X}_2)$ is shown as follows.

$$LAT(\hat{X}_2) = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|} \log \frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|} \log \frac{|c_1|}{w_0|c_0|+|c_1|} \right\}$$

with the following property:

Lemma 1. *$LAT(\hat{X}_2)$ has an unique critical point (CP)* $w_0 = |c_1|/|c_0|$.

B. The Characteristic curve of CCL

Based on the previous discussion, the relation between $LAT(\hat{X}_2)$ and w_0 can be described as a CCL. For example, $|c_0| = 100$, $|c_1| = 370$. Their relation is shown in Fig. 3. Based on this CCL, $LAT(\hat{X}_2)$ has a CP at $w_0 = 3.7$ according to Lemma 1. In other words, the maximum of $LAT(\hat{X}_2)$ should occur at $w_0 = |c_1|/|c_0|$ with it value given by

$$LAT(\hat{X}_2)|_{w_0=|c_1|/|c_0|} = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|} \log \frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|} \log \frac{|c_1|}{w_0|c_0|+|c_1|} \right\}\Bigg|_{w_0=|c_1|/|c_0|}$$

$$= -\log \frac{1}{2} \equiv LAT_{\max} \tag{7}$$

Since the minimum of $LAT(\hat{X}_2)$ should be attained when $w_0 \to 0$, i.e.,

$$LAT(\hat{X}_2)|_{w_0 \to 0} = -\left\{ \frac{w_0|c_0|}{w_0|c_0|+|c_1|} \log \frac{w_0|c_0|}{w_0|c_0|+|c_1|} + \frac{|c_1|}{w_0|c_0|+|c_1|} \log \frac{|c_1|}{w_0|c_0|+|c_1|} \right\}\Bigg|_{w_0 \to 0}$$

$$\equiv LAT_{\min} \to 0, \tag{8}$$

And we set $LAT_{\min} = 0.05$ to sufficiently approximate the smallest value of LAT for computational purpose. During the watermarking process, we also set LAT_{mid} to be

$$LAT_{\text{mid}} = \left(LAT_{\max} + LAT_{\min}\right)/2 \tag{9}$$

By Lemma 1, $LAT(\hat{X}_2)$ has two monotone subintervals which are called segment 1 and 2, referred to Fig. 3 as a typical example. In this work, we adopt the range $w_0 \in \left[0.05, |c_1|/|c_0|\right]$ of $LAT(\hat{X}_2)$ in segment 1 to embed data. The detail process will be introduced in the next section.

IV. The Proposed Watermarking Technique

In this section, the novel watermarking technique by using segment 1 in the characteristic curve of CCL is proposed. It contains embedding and extraction processes. These processed are introduced as follows.

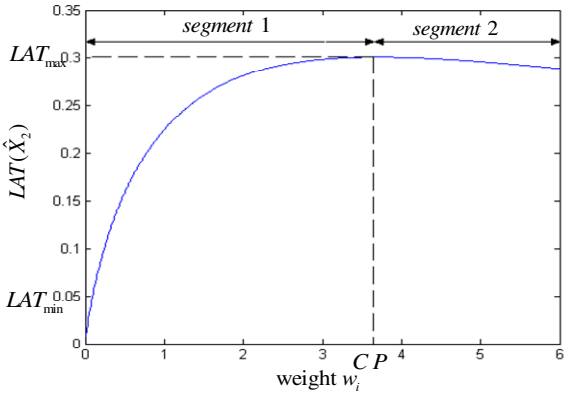


Fig. 3 The characteristic curve of LAT for $|c_0| = 100$, $|c_1| = 370$.

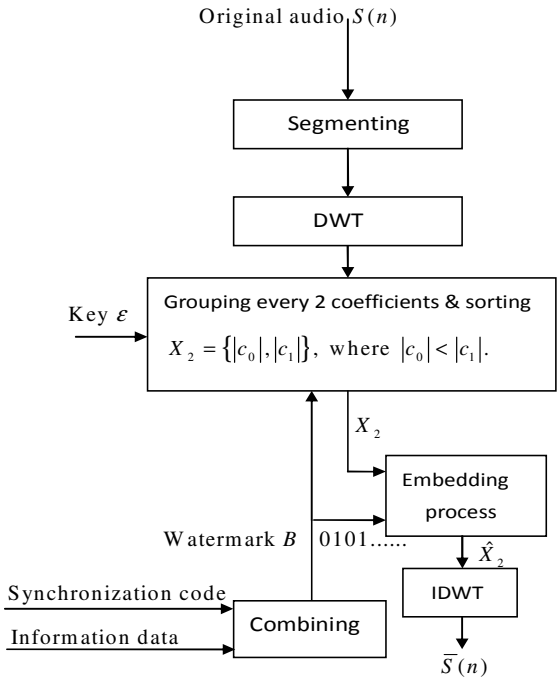


Fig. 4 The flowchart of watermark embedding process.

A. The embedding process

First of all, the synchronization codes and watermark are arranged into a binary pseudo-noise (PN) sequence B, for example, $B = \{0,1,10,1,...\}$. Secondly, as shown in Figure 4, the original audio $S(n)$ is split into proper segments, and DWT is applied to each segment. Then the synchronization codes and watermark are embedded into the lowest-frequency DWT coefficients. In this step, we group every two consecutive coefficients into $X_2 = \{|c_0|, |c_1|\}$ with $|c_0| < |c_1|$. The proposed embedding process is described as follows.

- If the binary bit "$1 \in$ B" is embedded, we choose w_0 such that

$$LAT(X_2W) = LAT(\hat{X}_2) = (LAT_{\max} + LAT_{\text{mid}})/2 + \varepsilon \quad (10)$$

- If the binary bit "$0 \in$ B" is embedded, we choose w_0 such that

$$LAT(X_2W) = LAT(\hat{X}_2) = (LAT_{\min} + LAT_{\text{mid}})/2 - \varepsilon \quad (11)$$

where $\varepsilon \in [0, 0.3]$ is a small positive number which can be used as a secret key.

B. The extraction process

The flowchart of watermark extraction is given in Figure 5. Every two consecutive lowest-frequency DWT coefficients is grouped into $X_2 = \{|c_0|, |c_1|\}$. To extract the watermark $\hat{B} = \{\hat{\beta}\}$, we apply (2) with *N* = *2* as follows.

- If $LAT(X_2) > LAT_{\text{mid}}$, the extracted value $\hat{\beta} = 1$.
- If $LAT(X_2) < LAT_{\text{mid}}$, the extracted value $\hat{\beta} = 0$.

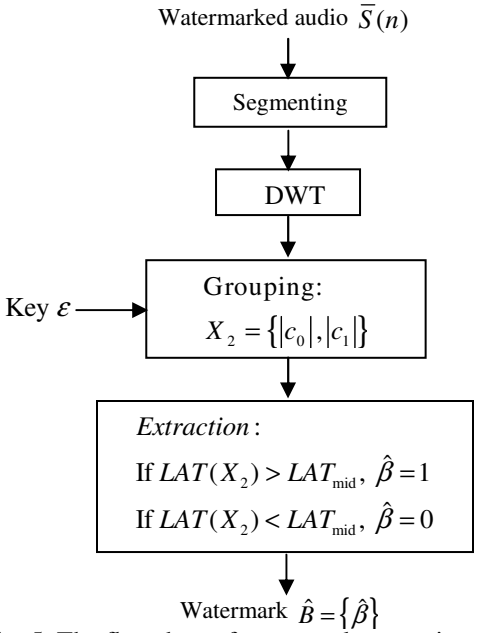


Fig. 5 The flowchart of watermark extraction.

V. EXPERIMENTAL RESULTS

The performance of the proposed audio watermarking technique is tested by using 16-bit mono audio signal sampled at 44.1 kHz. The length of each audio is about 11.6 seconds. We use two kinds of music which are symphony and popular. By setting the parameter ε to be 0.05, the synchronization code and watermark are embedded into the low-frequency DWT coefficients in level seven. Accordingly, the embedding capacity is 2000bits/11.6 secs. The SWR for the two audios are 20.8 dB (symphony) and 21.1 dB (popular). Moreover, we apply three types of attack to test the robustness: (1) re-sampling, (2) amplitude scaling, (3) low-pass filtering. The testing results are listed in TABLES I-III.

TABLE I. BER(%) after Re-Sampling

Rate (Hz)	22050	11025	8000
symphony	3.6	7.5	7.3
popular	9.2	12.9	14.7

TABLE II. BER(%) after Amplitude Scaling

Scaling factor	0.2	0.8	1.1	1.2
symphony	0.5	0.4	0.4	0.4
popular	0.5	0.5	0.4	0.4

TABLE III. BER(%) after Low-Pass Filter

Cut-off frequency(kHz)	3
symphony	24.1
popular	26.8

VI. CONCLUSIONS

A novel audio watermarking technique is proposed to embed the information by using LAT. When embedding the watermark, an analytical formula is provided to determine the weight on DWT coefficients. The experimental results show that the embedded data are robust against some attacks.

ACKNNOWLEDGEMENTS

The work was in part supported by the NSC, TAIWAN, under the grant: NSC 100-2115-M-029-003.

REFERENCES

[1] C.-Yu Yang, C.-H. Lin, and W.-C. Hu, "Embedding Limitations with Digital-audio Watermarking Method Based on Cochlear Delay Characteristics," *Journal of Information Hiding and Multimedia Signal Processing,* Vol. 2, No. 1, January 2011, pp. 1-23.

[2] S. Wu, J. Huang, D. Huang, and Y. Q. Shi, "Efficiently Self-synchronized Audio Watermarking for Assure Audio Data Transmission," *IEEE. Transactions on Broadcasting,* Vol. 51, No.1, March 2005, pp. 69-76.

[3] S. Xiang, and J. Huang, "Robust Audio Watermarking Against the D/A and A/D Conversions," CoRR abs/0707.0397, 2007.

[4] S.-T. Chen, G.-D. Wu, and H.-N. Huang, "Wavelet-Domain Audio Watermarking Scheme using Optimization-Based Quantization," *IET Proceedings on Signal Processing*, vol. 4, no. 6, 2010, pp. 720-727.

[5] H. G. Kaganami, S. K. Ali, and B. Zou, "Reversible Data Hiding By Adaptive IWT-coefficient Adjustment," *Journal of Information Hiding and Multimedia Signal Processing,* Vol. 2, No. 1, January 2011, pp. 24-32.

[6] R. M. Noriega, M. Nakano, B. Kurkoski, and K. Yamaguchi , "High Payload Audio Watermarking: toward Channel Characterization of MP3 Compression," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 2, April 2011, pp. 91-107.

[7] R. Martinez-Noriega, M. Nakano, B. Kurkoski, and K. Yamaguchi, "High Payload Audio Watermarking: Toward Channel Characterization of MP3 Compression," *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 2, No. 2, Apr. 2011, pp. 91-107.

[8] H. C. Huang and Y. H. Chen, "Genetic Fingerprinting for Copyright Protection of Multicast Media," *Soft Computing*, Vol. 13, No. 4, Feb. 2009, pp. 383-391.

[9] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Respresentation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, July 1989, pp.674-693.

附錄2

2012 International Conference on Computing, Measurement, Control and Sensor Network

A Lossless Characteristic Compression Method for Continuous Arterial Pulse Waveforms

Albert C.-Y. Lin[1*], Huang-Nan Huang[2*], Tzu-Min Lin[1],
Pin-Huang Hsu[1], and Ching-Chi Yen[1]

[1]*Department of Automatic Control Engineering, Feng Chia University, Taichung 407, Taiwan (R.O.C.)*
[2]*Department of Mathematics, Tunghai University, Taichung 407, Taiwan (R.O.C.)*

Abstract—**Arterial blood pressure is one of the important biomedical waveforms that possess lots of important clinical and pathological information. This paper aims in designing a technique, which is called characteristic compression method, to perform almost lossless data compression for the continuous arterial blood pressure waveforms. This method uses the spline function to interpolate the original waveform at preselected sample points first. A fourth order spline function is considered with associated formulas in describing the coefficients of the function. A second stage compression is accompanied by using 12-bit digital data representation to store sample points and coefficients for each interval. Numerical simulation confirms that the proposed technique is feasible and does provide a very high compression rate and almost lossless reconstruction for the continuous arterial pulse wave.**

***Keywords*—blood pressure, pulse wave, spline interpolation, sample point, jerk**

I. INTRODUCTION

According to the 2008 Factsheet WHO [1] "Top ten causes of death" divides countries into three groups: high-income, middle-income and low-income. The death due to ischaemic and hypertension heart diseases have already leapt into the top ten causes of death in high-income and middle-income countries and they occupy over 13% of the world's population. Thus the heart disease becomes the most mentioned disease in the whole world and monitoring the continuous arterial pulse waveforms becomes one of effective way for preventing this disease.

The measurement on the continuous blood pressure of the artery can not only provide the SBP and DBP as well as the MBP information, but also obtain the full waveform of the continuous blood pressure for each heart stoke. Based on the arterial pulse waveform, physiologists can determine the corresponding type of heart disease such as hypertension, heart rate volatility (HRV), mitral value regurgitation, etc.. According to the analysis of Windkessel model [2-4] on the continuous arterial pulse wave, one can compute various cardiovascular parameters such as arterial compliance, cardio output, blood volume, and left ventricular ejection time, etc. to aiming for prevention or early treatment on the heart disease.

Though the continuous blood pressure monitoring can reveal exhaustive pathologic information, but the amount of the acquired data is much larger than the static blood pressure obtained by using the traditional sphygmomanometer. With the sampled frequency at 500Hz and storage with 32-bit representation, the blood pressure monitoring system requires at least 168MB hard disk storage every day for round-the-clock monitor on each individual. At the same time, some special pathologic characteristics (such as sudden HRV, autonomic dysfunction, etc.) need a record of long-time measurement in revealing their significances. For some patients with serious disease or just after medical surgery, a long-term continuous blood pressure monitor is required. Therefore, it becomes important to have very high compression rate for storing the recorded continuous blood pressure waveform. In addition, with the rapid development of electric devices such as hand-carrying type continuous blood pressure instruments (for example, Portapres), novel technology in storing and compressing the continuous blood pressure can lighten the new measuring device and accelerate the wirelessly signal transmission speed which make the realization of the physiological monitoring system for home care possible.

The blood pressure reflects the operating status of the cardiovascular system which can assist physiologists in diagnosing the heart disease. Thus the compression technology should guarantee that the reconstruction of the blood pressure waveform from the compressed data is lossless to avoid inadequate judgments on the disease status. Under the lossless requirement, the compression technique seeks for the data compression ratio as higher as possible. The mainstream in compressing the pulse pressure waveform is mainly utilizing the method developed for the electrocardiogram (ECG/EKG) [5-6], such as Turning Point (TP) algorithm [7], Amplitude Zone Time Epoch Coding (AZTEC) algorithm [8-9], Reduction Time Encoding System (CORTES) algorithm [10], and Fan algorithm [11-13]. The compression rate (CR) of these methods lies between 2 and 5; the percent root-mean-square difference (PRD) around 5 to 29%. These methods do not provide high compression rate and their compression is also not lossless. In 2004, Chen etc.[14] classifies the continuous blood pressure signal according to the similarity for this kind of physiological signals, and then proposes to use Huffman coding [15], run-length coding [16] and vector quantization [17-18] for further data compression. Although the compression ratio of their method lies between 14.17 and 34.40, which is apparently higher than traditional algorithms, like TP, AZTEC, CORTES

978-0-7695-4738-1/12 $26.00 © 2012 IEEE
DOI 10.1109/CMCSN.2012.3

etc., however it is still not a lossless compression algorithm for the continuous blood pressure data.

This paper aims in proposing a characteristic compression method (CCM) which is an almost lossless technology to compress the continuous blood pressure data. It consists of two operational stages: the first one use the spline function to interpolate the original waveform at some specified sample points; the second stage uses 12-bit digital data representation to store the coordinates of sample points and the associated spline coefficients for each interval. Our proposed method can reduce the PRD value under 0.5% and maintains good compression efficiency.

The remaining parts of the paper are organized as follows: Section II will introduce mathematical formula in describing coefficients of the spline function. Due to the mismatching between the original and reconstructed waveforms, we add more sample points and derive the corresponding spline function such that to reduce their mismatching. Section III describes the digital data representation to store the selected sample points and computed spline coefficients in each interval. Also the length of binary representation for data storage is determined for second stage compression. Section IV shows our experimental results and their comparison with the previous researches. Finally, some concluding remarks are stated in last section.

II. SPLINE FUNCTION FOR SAMPLE POINTS

The main idea in the characteristic compression method is to locate certain points inside the signal which possess special characteristics such that the original waveform of the signal can be lossless reconstructed by utilizing the designed mathematical function passing to interpolate the wave through these data points. Thus, we only need to store coordinates of these characteristic points (which is referred as sample points in our paper) together with the coefficients or parameters of the designed mathematical function in each interval for the signal reconstruction. That is, the original waveform of the signal can be represented by a small amount of data which means the original signal is compressed at all. To implement this idea, two questions are raised in advance. The first one is how to select the so-called sample points inside the waveform of the signal? The other is how to design the mathematical function for lossless reconstruction of the original waveform based on the selected sample points?

To answer these two questions, we review the data interpolation theory from calculus and numerical analysis. Since the continuous waveform of the blood pressure is continuous and almost smooth everywhere, the peak/valley points, inflection points and jerk points are considered to possess most important features of the waveform. Thus we select the sample point to be the peak/valley point, inflection point or jerk point on the waveform of the signal. In our case, we consider the jerk point as the vanished third derivative occurs. In order to capsulate the smoothness property of the waveform, a nature choice is to pick the spline function, as described typically in equation (1), as our candidate for mathematical function for data reconstruction.

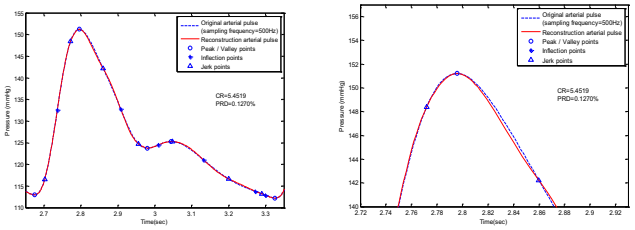
$$S_i(x)=b_i(x-x_i)^3+c_i(x-x_i)^2+d_i(x-x_i)+e_i,\ i=1,\ldots,n-1. \quad (1)$$

For the jerk point we obtain $b_i=0$. Although there is the possibility for two jerk points adjacent to each other, the computation error is too high and we neglect the point transition in this type. As we move on the index for sample points from $P(x_i,y_i)$ to $Q(x_{i+1},y_{i+1})$, six possible relationships between coefficients are listed below:

1. From peak/valley point *P* to inflection point *Q*:
   We obtain the following relationship for coefficients:
$$\begin{cases} y_{i+1}-y_i=b_i(x_{i+1}-x_i)^3+c_i(x_{i+1}-x_i)^2+0 \\ d_{i+1}-0=3b_i(x_{i+1}-x_i)^2+2c_i(x_{i+1}-x_i) \\ 0-c_i=3b_i(x_{i+1}-x_i) \end{cases} \quad (2)$$
2. From peak/valley point *P* to jerk point *Q*:
$$\begin{cases} y_{i+1}-y_i=a_i(x_{i+1}-x_i)^4+b_i(x_{i+1}-x_i)^3+c_i(x_{i+1}-x_i)^2+0 \\ d_{i+1}-0=4a_i(x_{i+1}-x_i)^3+3b_i(x_{i+1}-x_i)^2+2c_i(x_{i+1}-x_i) \\ c_{i+1}-c_i=6a_i(x_{i+1}-x_i)^2+3b_i(x_{i+1}-x_i) \\ 0-b_i=4a_i(x_{i+1}-x_i) \end{cases} \quad (3)$$
3. From inflection point *P* to peak/valley point *Q*:
$$\begin{cases} y_{i+1}-y_i=b_i\left(x_{i+1}-x_i\right)^3+0+d_i\left(x_{i+1}-x_i\right) \\ 0-d_i=3b_i\left(x_{i+1}-x_i\right)^2+0 \\ c_{i+1}-0=3b_i\left(x_{i+1}-x_i\right) \end{cases} \quad (4)$$
4. From inflection point *P* to jerk point *Q*:
$$\begin{cases} y_{i+1}-y_i=a_i(x_{i+1}-x_i)^4+b_i(x_{i+1}-x_i)^3+0+d_i(x_{i+1}-x_i) \\ d_{i+1}-d_i=4a_i(x_{i+1}-x_i)^3+3b_i(x_{i+1}-x_i)^2+0 \\ c_{i+1}-0=6a_i(x_{i+1}-x_i)^2+3b_i(x_{i+1}-x_i) \\ 0-b_i=4a_i(x_{i+1}-x_i) \end{cases} \quad (5)$$
5. From jerk point *P* to peak/valley point *Q*:
$$\begin{cases} y_{i+1}-y_i=a_i(x_{i+1}-x_i)^4+0+c_i(x_{i+1}-x_i)^2+d_i(x_{i+1}-x_i) \\ 0-d_i=4a_i(x_{i+1}-x_i)^3+0+2c_i(x_{i+1}-x_i) \\ c_{i+1}-c_i=6a_i(x_{i+1}-x_i)^2+0 \\ b_{i+1}-0=4a_i(x_{i+1}-x_i) \end{cases} \quad (6)$$
6. From jerk point *P* to inflection point *Q*:
$$\begin{cases} y_{i+1}-y_i=a_i(x_{i+1}-x_i)^4+0+c_i(x_{i+1}-x_i)^2+d_i(x_{i+1}-x_i) \\ d_{i+1}-d_i=4a_i(x_{i+1}-x_i)^3+0+2c_i(x_{i+1}-x_i) \\ 0-c_i=6a_i(x_{i+1}-x_i)^2+0 \\ b_{i+1}-0=4a_i(x_{i+1}-x_i) \end{cases} \quad (7)$$

Based on previous discuss, not only the parameters a_i, b_i and c_i are obtained from (2) ~ (7) for the spline in the given interval, but also the parameters c_{i+1} and/or d_i for next interval. Fig. 1 gives us the comparison between the original waveform and the reconstruction wave by using the fourth order spline function. The PRD is 0.127%, and compression ratio is 5.4515. It indicates that the fourth order spline function can almost reconstruct the original waveform, i.e., the data compression of continuous blood pressure by utilizing the

fourth order spline function together with sample points is almost lossless. But in comparison to the previous method, the compression ratio is not particularly remarkable. In order to raise the compression rate, a so-called digital data representation (DDR) is applied to compress the stored data again, i.e., DDR is adopted to represent sample points and interval coefficients of the fourth order spline.



(a) (b)

Fig. 1. Comparison between the original waveform and its reconstruction by a fourth order spline function: (a) a typical cycle, (b) enlarged portion around the peak at time 2.8 sec.

III. Digital Data Representation

Traditionally, the digital data with decimal point can be stored in the binary form with two different types of realizations: fixed-point representation system and floating-point representation system. As shown in Fig. 2, the first one uses 16 bits to represent the data with 8 bits for integer part before decimal point and other 8 bits for the fraction part after decimal point; while the second one uses 32 bits to represent a floating point number according to ANSI standard. Since there are only 8 bits use in Fig. 2(a) in representing the integer part of the data value, hence only 0~255 of the data range is allowed which may not be adequate for hypertension disease. Although the second one given in Fig. 2(b) has the resolution about 1/224 (≈5.96×10-7) but it uses 32 bits for one numerical data representation such that the compression rate can't be reduced further.

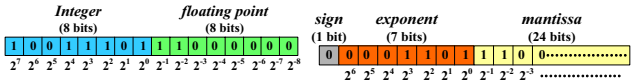


Fig. 2. Binary representation of a decimal number:(a) the fixed-point representation system, (b) the floating-point representation system according to the ANSI standard.

We conduct a survey on using how many digits to represent the data will be sufficient to provide the required performance for data compression. For the 32-bit float point number, we take its log value first and then store the rounding value into an n-bit long representation with its corresponding binary value. The reconstruction process is the same but with reverse order of operations to reconstruct the original waveform.

Our experiment selects 30 young people aged 25±5 years old. Their continuous arterial blood pressures are recorded for 450 sec long at the sampling frequency 500Hz by the Millar Instruments PCU-2000 Pressure Control Unit and SPT-301 pressure sensor. A 10Hz low-pass filter is utilized for data preprocessing. As shown in Figs. 3 and 4, both the PDR and the CR are decreasing as the number of bits increases. The PDR value seems to be stable when is greater or equal to 12. At the same time, the CR is still decreasing linearly as increases beyond 12. In conclusion, we select 12 bits for our numerical data representation, and the corresponding PRD is 0.1270% with standard deviation 0.0629% and the CR value is 18.4577 with standard deviation 2.5680.

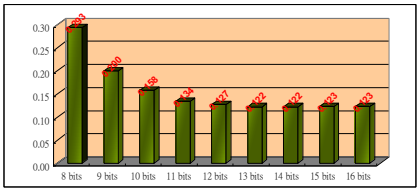


Fig. 3. The variation of PDR vs. the length n used in binary data representation. As increasing, the PDR is reduced as well and reaches a stable value when $n = 12$.

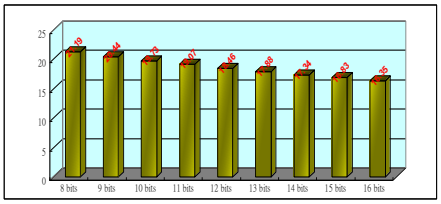
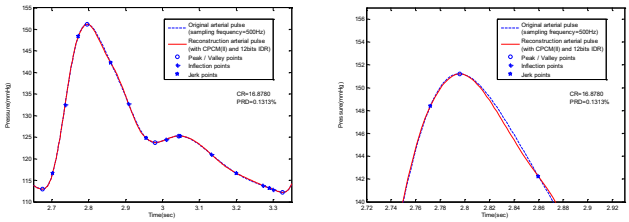


Fig. 4. The CR decreases proportionally to the length n used in binary data representation

The effect of using DDR with 12-bit representation for data storage is shown in Fig. 5 for the fourth order spline function. It clearly indicates that the compression rate increases dramatically from 5.4515 to 16.8780 and at the same time the error between reconstruction and original waves is almost remain the same, i.e., it changes from 0.1212% to 0.1313%. These results confirm the usage of DDR is a good strategy for the second stage compression of the arterial continuous pulse waveform.



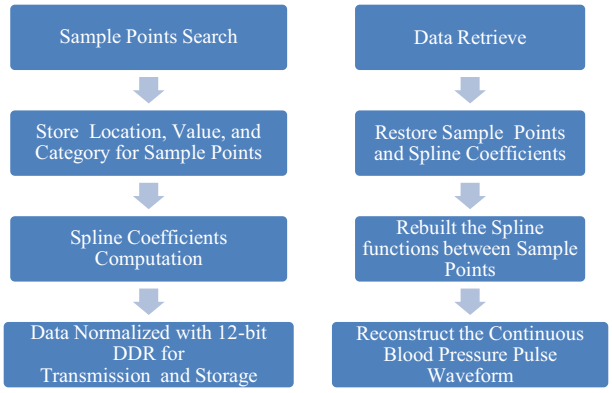
(a) (b)

Fig. 5. Comparison between the continuous pulse waveform and its reconstruction by a fourth order spline function and DDR representation: (a) the typical cycle, (b) enlarged portion around the peak at time 2.8 sec.

IV. Results and Comparison

Fig.6 describes the computational flow chart of our CCM method for both the data compression with spline function and DDR data representation and data reconstruction processes. Table 1 shows the comparison of the experimental result of CCM method and previous studies. Although two different spline functions by CCM without DDR has a very small PRD (<0.5%) than other methods, their CR's are not so good as VLC, VLC in DCT domain, and VQ coding. But after we use

the 12-bit DDR for further compression, the CR of 4th order CCM increases from 5.96 to 18.46. We want to mention that when the measured data without preprocessing by a 10Hz low-pass filter, the CR is increased to 24.4799±2.3643 and the associated PRD is 0.0793±0.0238%. Therefore the proposed technique does not have the lowest CR but have the lowest PRD which is very small when compared with other approaches. Based on these discussion we confirm that the proposed CCM is a good technique to provide not only the adequate CR but also a very small PRD.



(a) Data compression (b) Data reconstruction

Fig. 6. The flow chart of the proposed algorithm

Table 1. Comparison between the proposed method and other compression approaches

Method	CR	PRD (%)
TP	2	7.7
AZTEC	5	29.0
CORTES	4.6	8.4
Fan	2.5	5.9
VLC	14.17	3.90
VLC in DCT domain	26.42	7.02
VQ coding	34.40	5.99
4 order CCM	5.96±0.83	0.12±0.06
4 order CCM with12-bit DDR	18.46 ±2.58	0.13 ±0.06

V. Concluding Remarks

An almost lossless compression technique for continuous blood pressure pulse waveforms is proposed in the present study. The key idea is utilizing the spline function to interpolate the original waveform at the preselected sample points. At the same time, we can iteratively compute all the coefficients for each interval instead of solving a system of equations. This result reduces the computational complexity in evaluating spline coefficients. Another advantage of the proposed CCM method is that the stored data size depends only on the shape variation of the waveform and is independent of the sampling frequency for data acquisition. Thus, if the data are sampled at a higher frequency, the CR of applying CCM to this data set will also increase as well.

Our experimental result confirms that the CCM method can provide almost lossless compression effect (PRD<1%), and the CR will reach 18 when CCM with 12-bit DDR for data storage. Therefore the proposed CCM with 12-bit DDR is indeed a high compression technique for the lossless data reconstruction of the continuous arterial blood pressure.

References

[1] World Health Organization, 2008 Factsheet–"The Top Ten Causes of Death", May 2011. Available at:
http://www.who.int/entity/mediacentre/factsheets/fs310_2008.pdf

[2] T.S. Manning, B.E. Shykoff, and J.L. Izzo, Jr, "Validity and Reliability of Diastolic Pulse Contour Analysis (Windkessel Model) in Humans," *Hypertension*, Vol. 39, pp. 963-968, 2002.

[3] V.P. Crabtree and P.R. Smith, "Physiological Models of the Human Vasculature and Photoplethysmography," Electronic Systems and Control Division Research, pp.60-63, 2003.

[4] J.N. Cohn, D.J. Morgan, and C.W. Bratteli, "Apparatus and method for blood pressure pulse waveform contour analysis," International Application Published under the Patent Cooperation Treaty(PCT), International Patent Classification: G06F 17/00, A61B 5/021, International Publication Number: WO 99/48023, International Publication Date: 23 September 1999.

[5] J.M.S. Jalaleddine, "ECG Data Compression Techniques--A Unified Approach," IEEE Transactions on Biomedical Engineering, Vol. 37, No.4, pp. 329-343, 1990.

[6] C.-C. Sun, "ECG Compression algorithms utilizing the inter-beat correlation," Dissertation for Doctor of Philosophy, Department of Electrical Engineering, National Cheng Kung University, Taiwan, ROC, July, 2005.

[7] W.C. Mueller, "Arrhythmia Detection Software for an Ambulatory. ECG Monitor," Biomed. Sci. inst., Vol. 14, pp. 81-85, 1978.

[8] J.R. Cox, et al., "AZTEC: a preprocessing program for real-time ECG rhythm analysis," IEEE Transactions on Biomedical Engineering, Vol. 15, pp.128-129, April 1968.

[9] V. Kumar, S.C. Saxena, V.K. Giri, and D. Singh, "Improved modified AZTEC technique for ECG data compression: Effect of length of parabolic filter on reconstructed signal," Computers and Electrical Engineering, Vol. 31, pp. 334-344, 2005.

[10] J.P. Abenstein and W.J. Tompkins, "A New Data-Reduction Algorithm for Real-Time ECG Analysis," IEEE Transactions on Biomedical Engineering, Vol. BME-29, No. 1, pp. 43-48, 1982.

[11] L.W. Gardenhire, "Redundancy reduction the key to adaptive telemetry," Proceedings of the 1964 National Telemetry Conference, pp.1-16, 1964.

[12] L.N. Bohs and R.C. Barr, "Prototype for real-time adaptive sampling using the fan algorithm," Computer Methods and Programs in Biomedicine ., Vol. 26, pp. 574-583, 1988.

[13] M. Giallorenzo, J. Cohen, F. Mora, G. Passariello and L.O. Lara, "Ambulatory monitoring device using the fan method as data-compression algorithm," Medical and Biological Engineering and Computing, Vol. 26, No. 4, pp. 439-443, 1988.

[14] W.-S. Chen, L. Hsieh, and S.-Y. Yuan, "High performance data compression method with pattern matching for biomedical ECG and arterial pulse waveforms," Computer Methods and Programs in Biomedicine, Vol. 74, pp. 11-27, 2004.

[15] D. Huffman, "A Method for Construction of Minimum-Redundancy Codes," Proceedings of the IRE, Vol. 40, No.10, pp. 1098-1101, September 1952.

[16] C. Tu, J. Liang, and T.D. Tran, "Adaptive Runlength Coding," IEEE Signal Processing Letters, Vol. 10, No. 3, pp. 61-64, March 2003.

[17] C.P. Mammen and B. Ramamurthi, "Vector quantization for compression of multichannel ECG," IEEE Transactions on Biomedical Engineering, Vol. 37, pp. 821-825, September 1990.

[18] B. Wang and G. Yuan, "Compression of ECG data by vector quantization," IEEE Transactions on Biomedical Engineering, Vol. 37, pp. 23-26, 1997.

[19] J.P. Abenstein and W.J. Tompkins, "A New Data-Reduction Algorithm for Real-Time ECG Analysis," IEEE Transactions on Biomedical Engineering, Vol. 29, No. 1, pp. 43-48, 1982.

附錄3

2012 International Conference on Computing, Measurement, Control and Sensor Network

Wavelet-based Quantization Watermarking for ECG Signals

Xialong He[1], Kuo-Kun Tseng[1], Huang-Nan Huang[2*], Shuo-Tsung Chen[2], Shu-Yi Tu[3], Fufu Zeng[1] and Jeng-Shyang Pan[1]

[1]Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen Graduate School, China
[2]Department of Mathematics, Tunghai University, Taichung 40704, Taiwan(ROC)
[3]Mathematics Department, University of Michigan, Flint MI 48502, USA
[*]Corresponding author, nhuang@thu.edu.tw

Abstract—**In this article, we use a self-synchronized watermark technology [7], to achieve the purpose of protection of electrocardiogram (ECG) signal. A Harr wavelet transform with 7 levels decomposition is adopted to transform the ECG signal and the synchronization code, combined with watermark, are quantized embedded in the low-frequency sub-band of level 7. The signal to noise ratio (SNR) between the embedded ECG and original one is greater than 30 such that the difference between these two ECG signals is very small and negligible in general. To test the robustness under the network transfer of ECG data, a white noise attack with various strengths is simulated that the bit error rate is quite small unless the SNR of the noise is very large. This study confirms the use of wavelet-based quantization watermarking scheme on ECG signal for patient protection is adequate.**

Index Terms- ***Electrocardiogram; watermarking; wavelet; data transmission; self-synchronization.***

I. INTRODUCTION

Nowadays in the society, we pay more attention to a variety of copyright protection and the emphasis on personal information. Due to the change in landscape of medical environment, the delivery and transfer of medical data between hospitals or clinics do occur very frequently.

In the past, the patient data was randomly stored inside the hospital without any protection. However, with the development of science and technology, it is found that the patient data contains important private information, and villains even can use some information. Therefore, protection measures should be taken in the process of storing, transmitting, or browsing the information. Thus the protection of medical data through data hiding technique is undoubted an important issue.

Although some data hiding algorithms can embed data into medical data, the original information may be distorted permanently. But in medical diagnosis, these changes are not allowable. Thus we are only concerned on data hiding method for which the lossless original media can be restored from marked media.

Watermarking technology is the most widely used data hiding technology in the field of multimedia. Digital watermarking technology refers to directly embedding some identification information (watermark) into the carrier (including multimedia, documents, software, etc.). It does not affect the usage of the original carrier and is hard to be perceived by ordinary perception system such as visual or auditory system. The hidden information in the carrier can help us to confirm the content creators, buyers, carrier's transmission secret information to determine whether the carrier is altered or not during its transmitting process. Digital watermarking is an important research direction of information hiding technology.

Electrocardiogram (ECG) reflects the process of the electrical activity of our heart, which can be taken as a reference for the study of cardiac pathology and cardiovascular system diagnostics. With ECG signals, we can analyze and identify various heart diseases, such as arrhythmias, myocardial damage etc. ECG has high requirements for accuracy. Thus ECG is one of the very important bio-information to be protected.

In 1998, application of watermarking technique in medical image was proposed by Anand and Niranjan [1] to embed the patient information. In 2005, Engin et al. [2] proposed a very elementary watermarked technique for ECG signal to resist the white noise attack. At present, the research on the protection of ECG information with watermarking technique is still in its infancy stage, there are few related researches. All these works utilize wavelet-based digital watermarking encryption technology

Nambakhsh et al. [3] proposes a novel blind watermarking method combined with the EZW-based wavelet coder to embed ECG signals as secret key into medical CT and MRI images. Zheng and Qian [4] developed a wavelet-based algorithm to watermark ECG signals in non-QRS complex region to guarantee the restore of almost un-distorted ECG signals. Kaur et al. [5] constructed a blind digital watermarking to ensure the safe transmission of ECG signals in wireless network that the embedded watermark can be fully removed by the receiver. Ibaida [6] developed an watermarked algorithm such that the ECG signals are watermarked with patient biomedical information to confirm patient/ECG linkage integrity and is suitable for a wearable sensor-net health monitoring system.

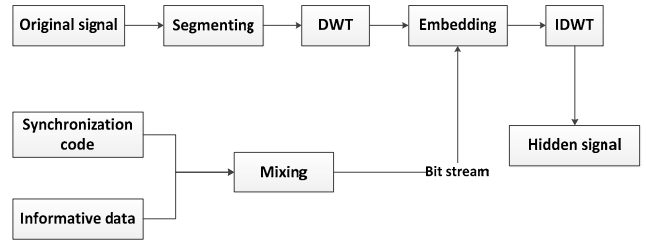
In this paper we preliminarily study the effect of applying wavelet-based quantization watermarking scheme on ECG signal with self-synchronization mechanism [7]. Although this type of watermarking technique is not a reversible technique, but if the change in ECG signal is small, then it is acceptable. The organization of this paper is as follows. Section 2 describes the proposed algorithm of this paper. Section 3 describes the experimental result of the proposed algorithm by utilizing the MIT-BIH database. Some conclusions are drawn in the last section.

II. PROPOSED ALGORITHM

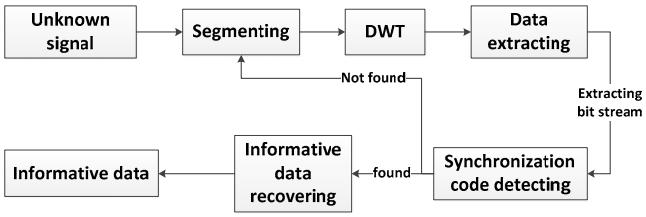
Since ECG signals are one-dimensional, various watermarked techniques for audio signals can be considered as the candidate. Consideration on the safety transmission of

978-0-7695-4738-1/12 $26.00 © 2012 IEEE
DOI 10.1109/CMCSN.2012.119

ECG data through the network, the self-synchronized audio watermarking technique [7] is adopted here. The embedding and extraction process is described in Fig. 1.



(a) watermark embedding



(b) watermark extraction

Fig.1. Watermark embedding and extraction process

In this process, we use the synchronization code. It can be used to locate hidden information, to prevent the unpredictable attacks. Supposing $A = \{a_i\}$ is a source-synchronous code and $\{b_i\}$ is an unknown code with the same length of A. If the difference between $\{a_i\}$ and $\{b_i\}$ are less than the determined threshold, then $\{b_i\}$ will be identified as the synchronization code.

Fig. 2. Wavelet decomposition

Let $S = \{s_1, s_2, \dots, s_N\}$ denote the ECG signal with total length N sample points. Fig. 2 illustrates the DWT decomposition process and we use 7 levels DWT on ECG signals. The watermark together with the synchronization code (which is called PN sequence $\{m_i\}$) will be embedded into 7th level low- frequency sub-band (denoted by $\{c_i\}$), i.e., the coefficients A7.

The rule of embedding is as follows.

$$\tilde{c}_i = \begin{cases} \lfloor c_i/Q \rfloor \cdot Q + 3Q/4, & \text{if } m_i = 1 \\ \lfloor c_i/Q \rfloor \cdot Q + Q/4, & \text{if } m_i = 0 \end{cases} \quad (1)$$

where $\{c_i\}$ and $\{\tilde{c}_i\}$ are the level 7 low-frequency DWT coefficients before and after embedding, respectively, and Q is the embedding strength. By applying the IDWT, the corresponding watermarked ECG signal is obtained and denoted by $\tilde{S} = \{\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_N\}$. The insertion of watermark will affect the original signal and we use the signal to noise ratio (SNR) to measure the effect:

$$SNR = -10 \log \frac{\sum(\tilde{s}_i - s_i)^2}{\sum s_i^2} \quad (2)$$

When extracting the data, we divide the ECG signal $S^* = \{s_1^*, s_2^*, \dots, s_N^*\}$, which have been embedded with watermark, into several sections, of which at least includes one synchronization code segment. Then we performed DWT transform on each section. Suppose $\{c_i^*\}$ is the coefficient of level 7 low-frequency sub-band, we use the following rules extract sequence $\{m_i^*\}$ from $\{c_i^*\}$:

$$m_i^* = \begin{cases} 1, & \text{if } c_i^* - \lfloor c_i^*/Q \rfloor \cdot Q \geq Q/2 \\ 0, & \text{if } c_i^* - \lfloor c_i^*/Q \rfloor \cdot Q < Q/2 \end{cases} \quad (3)$$

III. Experimental Results

We selected four sets of data from the MIT-BIH Arrhythmia Database [8], i.e., data 100-103. The sampling rate of the ECG signal is 360Hz. In each data set, we select a fragment of length 4096 to be tested. The PN sequence consists of 8 bits synchronization code and 32 bits watermark. The Haar wavelet transform is applied to the signal down with 7 levels decomposition. Then we utilize the quantization in (1) to embed the PN sequence into the ECG signal with the embedding strength $Q = 4096$. It is noted that the ECG signal is adjusted to have zero mean first, and then is scaled to the resolution with 16-bit representation.

Fig. 3 shows the original and watermarked signals for data 100 look almost distinguishable. And we enlarge the portion in Fig. 3 around the first second and plot both on the same graph as drawn in Fig. 4. The difference is quite small, i.e., the PQRST complex from the watermarked ECG signal is almost the same as the original one. Figs. 5-7 compare the original and watermarked signals for data 101-103. Table I gives us the SNR of the watermarked signals for data 100-103, which are all larger than 30. And the relative error due to watermark is around 3% under 2-norm measure. Although this quantization technique is not a reversible scheme, but the change due to watermark is negligible. On the other hand, we can adjust the embedding strength to increase the SNR and whence the relative error is reduced. But the robustness of the watermark will be reduced.

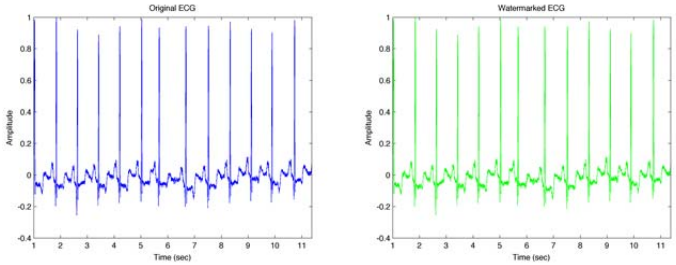


Fig. 3. Original and watermarked signals for data 100

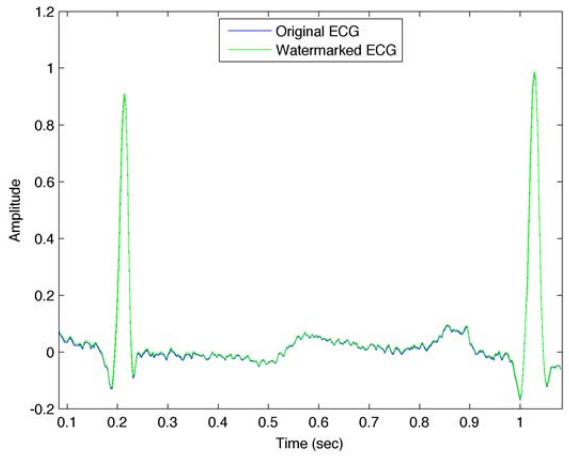


Fig. 4. Original and watermarked signals for data 100 in [0.09,1.09]

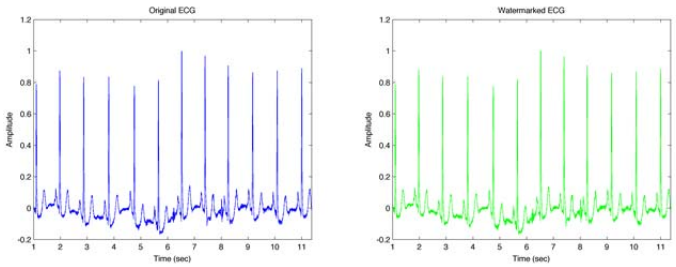


Fig. 5. Original and watermarked signals for data 101

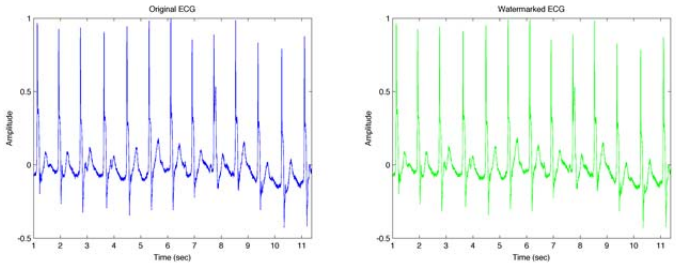


Fig. 6. Original and watermarked signals for data 102

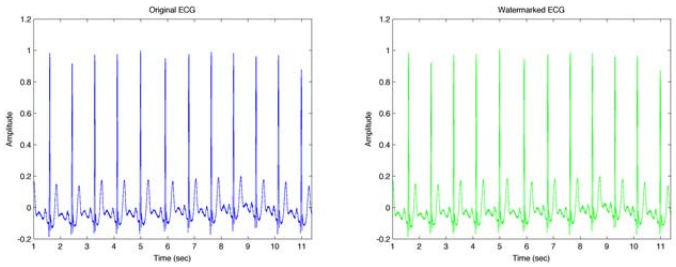


Fig. 7. Original and watermarked signals for data 103

TABLE I
IMPACT OF WATERMARK ON ECG SIGNAL

Data ID	SNR	$\sqrt{\sum(\tilde{s}_i - s_i)^2 / \sum s_i^2}$
100	30.3960	0.0299
101	30.6291	0.0354
102	31.7761	0.0327
103	32.0753	0.0269

Since the ECG data may be transferred using network, we consider the white noise attack to test the robustness of the watermark technique, i.e.,

$$s_i^* = \tilde{s}_i + \alpha \cdot y_i$$

where $\{\tilde{s}_i\}$ is the watermarked signal and $\{s_i^*\}$ is the attacked signal which is influenced by the white noise $\{y_i\}$ with zero mean and standard deviation one. Here α is considered as the strength of the white noise, i.e., $\{\alpha y_i\}$ is zero-mean white noise with standard deviation α.

The bit error rate (BER) is used to measure the correctness of watermark after some attack is presented which is an indication of the robustness of the watermark.

$$BER = \frac{\text{Number of error bits}}{\text{Number of total bits}} \times 100\%$$

Table II describes the robustness of the watermarked ECG signal under the attack of white noise with different standard deviation. For all data set, the white noise with α less or equal to 500 does not change the watermark and only when α is greater than 500 it will produce error bits in watermark extraction. The corresponding SNR for $\alpha = 500$ is around 15 which means the noise is very large which will not happen for the normal network transfer.

TABLE II.
ROBUSTNESS TEST VIA WHITE NOISE ATTACK

Data ID	White Noise deviation α (SNR)	Error Bits	BER(%)
100	1 (72.6337)	0	0
100	500 (18.6543)	0	0
100	750 (15.1325)	4	12.5000
100	1000 (12.6337)	7	21.8750
101	1 (71.8875)	0	0
101	50 (17.9081)	0	0
101	500 (14.3873)	4	12.5000
101	1000 (11.8875)	7	21.8750
102	1 (73.9786)	0	0
102	500 (19.9992)	0	0
102	500 (16.4773)	4	12.5000
102	1000 (13.9786)	7	21.8750
103	1 (73.7127)	0	0
103	500 (19.7333)	0	0
103	750 (16.2115)	4	12.5000
103	1000 (13.7127)	7	21.8750

IV. CONCLUSION

In this paper, we apply the self-synchronized quantization watermarked scheme [7] to embed watermark into the ECG signal. After tested with four data set from MIT-BIH arrhythmia database, the difference between the watermarked ECG and original one is very small and negligible. We also use the white noise with various standard deviations to test the watermarked ECG which shows a very strong robustness. This confirms that the application of wavelet-based quantization scheme to ECG signal is successful. In the future, not only more data set should be used to verify our conclusion and the detail influence of the watermarking technique on the features like QPRST complex should be examined as well.

REFERENCES

[1] D. Anand and U. C. Niranjan, "Watermarking medical images with patient information," in *Proceeding of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Hong Kong, China, 29 Oct.-1 Nov. 1998, Vol. 2, pp.703-706.

[2] M. Engin, O. Çidam and E. Z. Engin, "Wavelet transformation based watermarking technique for human electrocardiogram (ECG)", *Journal of Medical Systems*, Vol. 29(6), pp. 589-594, 2005.

[3] M. S. Nambakhsh, A. Ahmadian, M. Ghavami, R. S. Dilmaghani, and S. Karimi-Fard, "A novel blind watermarking of ECG signals on medical images using EZW algorithm," in *Proceedings of the 28th Annual International Conference of the IEEEEngineering in Medicine and Biology Society* (EMBS'06), New York, USA, 30 Aug-3 Sept. 2006, pp. 3274-3277.

[4] K.-M. Zheng and X. Qian, "Reversible data hiding for electrocardiogram signal based on wavelet transforms," in *Proceedings: 2008 International Conference on Computational Intelligence and Security* (CIS2008), Suzhou, China,, 13-17 Dec. 2008 pp. 295-299.

[5] S. Kaur, O. Farooq, R. Singhal, and B. S. Ahuja, "Digital watermarking of ECG data for secure wireless communication," in *Proceedings: 2010 International Conference on Recent Trends in Information, Telecommunication and Computing* (ITC2010), Kochi, Kerala, India, 12-13 March 2010, pp. 140-144.

[6] A. Ibaida, I. Khalil, and R. van Schyndel, "A low complexity high capacity ECG signal watermark for wearable sensor-net health monitoring system, " in *Proceeding: Computing in Cardiology* (CinC), Hangzhou, China, 18-21 Sept. 2011, pp. 393-396.

[7] S. Wu, J. Huang, D. Huang, and Y.Q. Shi, "Efficiently self-synchronized audio watermarking for sssured audio data transmission," *IEEE Transactions on Broadcasting*, Vol. 51(1), pp. 69-76, 2005.

[8] G. B. Moody, R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, 2001.

附錄4

2012 International Conference on Computing, Measurement, Control and Sensor Network

ECG Human Identification with Statistical Support Vector Machines

He Chen[1], Fufu Zeng[1], Kuo-Kun Tseng[1], Huang-Nan Huang[2*], Shu-Yi Tu[3] and Jeng-Shyang Pan[1]
[1]Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen Graduate School, China
[2]Department of Mathematics, Tunghai University, Taichung 40704, Taiwan (ROC)
[3]Mathematics Department, University of Michigan, Flint MI 48502, USA
[*]Corresponding author, nhuang@thu.edu.tw

***Abstract*—Electrocardiogram (ECG) as a biological information, it has some special feature. Different people will have different ECG information, even one person has different ECG when he is under different body state. In this paper we use the Electrocardiogram (ECG) to identify disease or to detect different person. Firstly, we collect the ECG information form different body state of the different people. Secondly we will preprocess the ECG data by using a method of statistical. Thirdly we can use the support vector machine to train the data, and then classify different people's data into different class. And finally when there are one new ECG data, we can also use SVM to identify the new data. Because even one people have several ECG signal, with our statistical method, the classifier may gets better robust.**

Keywords - ECG; human identification; SVM

I. INTRODUCTION

A lot of biological information has been widely used for clinical application. Electrocardiogram (ECG) is one of the very important biological signals. ECG signal is one-dimensional data to represent the time electrical change of the voltage variation, which is detected on the skin. Besides of utilizing ECG to identify the human physiological status, it has been recently adopted in human identification. As ECG is so special than the other biological information that the ECG signal varies from time to time, and even for the same person different waveforms of ECG signal will be presented at different body state. Thus the biometric designed based on ECG signal will be dynamical which is different from the other identification method by utilizing the static features of fingerprint, face and iris etc.. This dynamical property indicates that the ECG biometric is hard to be copied than the other biometrics. Thus various research works [1-10] reveals that ECG biometric is the realizable and reliable.

The SVM algorithm is proposed by Boser, Guyon and Vapnik in 1992 [11]. SVM has a good performance on classifying. It's base on the statistic theory of machine learning algorithm. SVM could automatically look for those support vector who has a good ability on distinguishing different classes. The classifier base on SVM can maximize the distance between different classes.

Fig. 1 is our taxonomy of related researches of using support vector machine (SVM) to analyze ECG signals. There are three aspects: applications, type of SVM, and feature extraction.

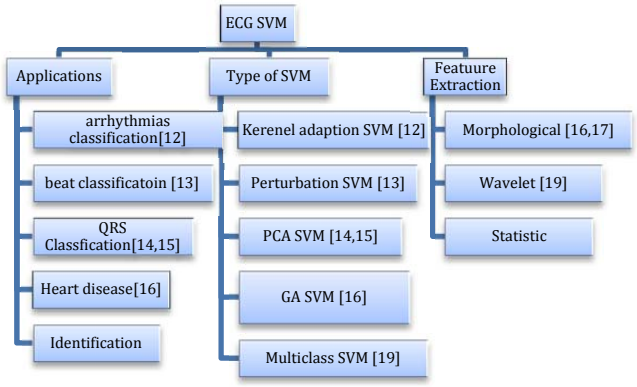


Fig. 1. Related studies of ECG analysis via SVM

ECG signals contain many important bio-information and scientists have using ECG in many applications with disease detection as a typical one. Moavenian and Khorrami [12] use ECG signal for arrhythmias classification for patients. Beat classification was introduced by Acir [13] to classify various types of beat. Mehta and Lingayat [14, 15] proposed two SVM methods on QRS classification of ECG signals; the SVM was applied as a classifier to delineate QRS and non-QRS regions. Polat and Gunes [16] developed an algorithm to find the heart disease.

Different types of SVM have been applied in ECG analysis. Moavenian and Khorrami [12] use the kernel-adaption algorithm to aid SVM for ECG arrhythmias classification. The SVM is much faster than MLP (multi-layered-perception) in training stage, and several times higher in performance. But MLP's mean square error is three times less than SVM. In the paper of Acır[13], they have used perturbation method to extract the feature the ECG data, and then apply SVM with PCA to classify four types of ECG beats. Polat and Gunes [16] develop an algorithm base on PCA (principle component analysis) and least square SVM. Zhang and Zhang [17] extract the principle characteristic of the ECG signal by using PCA technique and then SVM is used to classify the ECG data into four categories of heart disease. A method combining SVM and genetic algorithm is proposed by Nasiri and Naghibzadeh etc [18] where twenty-two features were extracted from the ECG signal, and then using SVM with genetic algorithm to searching for the best value of the

978-0-7695-4738-1/12 $26.00 © 2012 IEEE
DOI 10.1109/CMCSN.2012.120

parameters, and looking for the best subset of the feature that optimizes the classification function. Ubeyli [19] combined the wavelet coefficients and multiclass SVM method as a classifier for four types of the ECG beats are obtained.

In this paper, the human identification using SVM classifier is developed based on the rank order statistical feature of the ECG signals. The organization of this paper is as follows. Section 2 describes the proposed algorithm of this paper including the frequency and rank order statistics and support vector machine. Section 3 describes the experimental result of applying the proposed algorithm to classify individual by utilizing the MIT-BIH database. Some conclusions are drawn in the last section.

II. Proposed Algorithm

Fig. 2 shows the structure of utilizing SVM in ECG human identification, which consists of two steps. Firstly we transfer the input ECG signals into reduced binary pattern, and counting and ranking the appearance of patterns, which is regarded as features of ECG signals. Secondly, the feature data for various persons are used to training the SVM classifier and then for matching test to identify an unknown input ECG signal. These steps are described in detail as following.

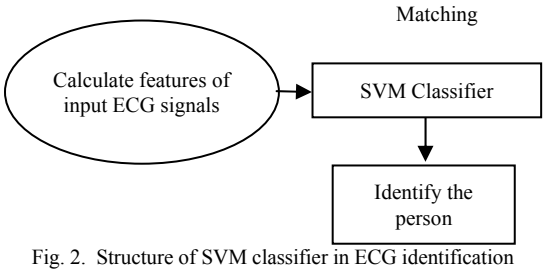


Fig. 2. Structure of SVM classifier in ECG identification

1) Frequency and Rank order Statistics

Consider an ECG signal as $S = \{x_1, x_2, \dots, x_i, \dots, x_n\}$ where real-valued x_i corresponds to the i^{th} input data. Compare each pair of consecutive input signal and categorize the data into one of the two cases: decrease or increase in x_i. A preliminary reduced function then maps these two cases to 0 or 1, respectively, according to the rule:

$$y_i = \begin{cases} 0\,, x_{i+1} \leq x_i \\ 1\,, x_{i+1} > x_i \end{cases} \quad (1)$$

That is, this procedure converts the ECG signal of length N to a binary sequence $Y = \{y_1, y_2, \dots, y_{N-1}\}$ of length $N-1$. Group every m bits in Y into a reduced binary sequence of length m, referred as an m-bit word; collect all such words to form a reduced binary pattern $B = \{b_1, b_2, \dots, b_k, \dots, b_{N-m}\}$ where $b_k = \{y_k, y_{k+1}, \dots, y_{k+m-1}\}$. We then convert each m-bit word b_k to its decimal expansion w_k.

Next, count the occurrences of all w_k and then sort them in the order of descending frequency. For $k = 1, 2, \cdots, N-m$, define $j = w_k$. It is obvious that, values of j range 0 over $2^m - 1$. Let $p(j)$ be the corresponding relative frequency of j, $p(j) = \frac{n_j}{N-m}$ and $\sum_{j=0}^{(2^m-1)} n_j = N - m$. Next, rank j according to its frequency n_j, from the largest to the smallest. For instance, if $R(j) = 1$ corresponds to a specific j, it mean that the m-bit words b_k who convert to the same decimal expansion j are those appear the most in the reduced binary pattern.

The relative frequency array p and the rank array R are considered as the features of ECG signals, and will be the input to the SVM classifier for training purpose and person identification. The process is described in Fig. 3.

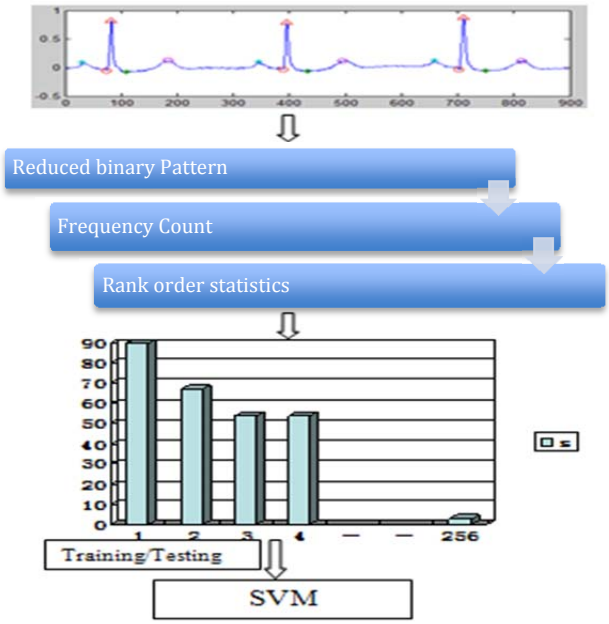


Fig. 3. Process of ECG human identification with statistical support vector machine

2) SVM classification

The purpose of SVM is to find a hyperplane, which can separate the training data set and get a maximal distance against the direct of the edge orthogonal to the hyperplane. SVM has good performance on small number of samples.

Given a training data set $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ where $x_i \in \mathbb{R}$ and y_i is either 1 or -1 indicating the class to which the point x_i belongs. Let $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$. The construction of the hyperplane for a linearly separable problem is $\mathbf{w}^T\mathbf{x} + b = 0$ where $\mathbf{w}$ is the normal vector to the hyperplane and the parameter $\frac{b}{\|\mathbf{w}\|}$ determines the offset of the hyperplane from the origin along the normal vector $\mathbf{w}$. Thus the margin between the hyperplane and the nearest point is maximized and can be posed as following problem:

$$\begin{aligned} &\min_{\mathbf{w},b,\xi} \tfrac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i \\ &\text{subject to } y_i(w^T x_i + b) \geq 1 - \xi_i,\ i = 1,2,\dots,n \quad (1) \\ &\xi_i \geq 0 \end{aligned}$$

where C is a user define constant as the penalty parameter of the error term. The SVM requires the optimal solution. We use the LIBSVM [21] to solve this optimization problem with its user's guide given in [22].

III. Experimental Result and Discussion

In real application, the ECG data should be collected from persons under various body states. For simplicity, we select the MIT-BIH Arrhythmia Database [23]. This database includes 48 groups, within two-lead ECG recordings for half an hour, a total of up to 24 hours of information. The data contains 47 individuals' ECG information (dataset ID 201 and 202 are duplicated); subjects consist of 25 men aged between 32 to 89 and 22 women aged from 23 to 89. These ECG data has a sampling rate of 360Hz and a 12-bit binary representation.

For each individual, 8 segments of 10 sample periods long are obtained from the record of its ECG signal in the database. Thus 3600 sample points in each segment are selected for frequency and rank order statistics. We set $m = 8$, i.e., the reduced binary pattern consists of 8-bit words and there are in total 256 different 8-bit words for frequency and rank order calculation. Afterward, for each individual there are 8 data sets and 256 features for each data set to training via SVM. The input file for the LIBSVM program consists of a matrix with 8×48 rows and 256 columns; each row stores 256 statistic features of the corresponding person and each individual has eight rows with the same label to represent the frequency and rank order of each segment in his/her ECG signal. The process in training the SVM for MIT-BIH database is shown in Fig. 4.

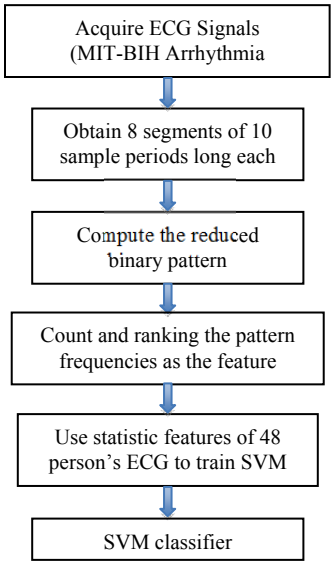


Fig. 4. Training SVM classifier for MIT-BIH database

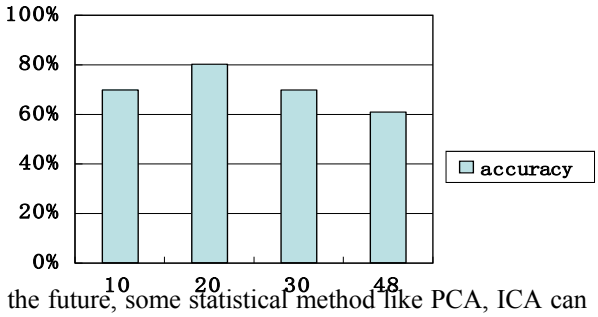
After training the SVM with the statistical features of ECG signals, we get a SVM classifier to identify individuals. The test data for identification are also acquired from the same MIT-BIH database. For each individual, we recapture 10 segments of 3600 sample points long. Note that these 10 segments are obtained at different location of the ECG waveform, i.e., none of them are overlapped with previous selected segments in training process. Each segment is pass through SVM classifier for matching test. Thus there are 10 matching tests for each individual.

Table I shows some result of matching test. The second row in Table I denotes that 10 segments for individual No. 1 (denoted by ID1 for simplicity) are all identified correctly, thus it is 100% accurate. But for the third row, it indicates that the 2nd, 6th, 8th, and 9th segments of ID2 are identified as ID6, while the last segment is identified as ID3. Thus only 4 segments of ID2 are identified correctly, i.e., the accuracy is 40%. In summary, the accuracy for some individuals are very high but for some of others are very low. We set the population size M to denote the first M individuals from 48 individuals and take their average as the group accuracy. Fig 5 summarizes the group accuracy of various group population sizes. It shows that the group accuracy for matching test lies between 60-80%, which is moderately acceptable. The possible reason is 256 features are too large for SVM. Even an alternative ECG signal not from MIT-BIH database is pass through the SVM, the classifier will identify it as one of the 48 individuals. One of the advantage is the proposed algorithm is executed very fast and has less computational complexity.

TABLE I
Result of Matching Test for Individual Identification

indiv.\data	1	2	3	4	5	6	7	8	9	10	Accuracy
1	1	1	1	1	1	1	1	1	1	1	100%
2	2	6	2	2	2	6	3	6	6	3	40%
3	3	1	3	3	3	3	3	1	3	3	80%
4	4	4	4	4	4	4	4	4	4	4	100%
5	25	32	40	21	2	32	1	5	5	25	20%
6	6	6	6	6	6	6	6	6	6	6	100%
7	35	35	35	6	5	7	6	11	32	7	20%
8	8	8	8	8	8	8	37	8	8	8	90%
9	9	9	1	3	1	3	1	31	1	31	20%
10	10	10	10	10	10	10	10	10	10	10	100%
11	11	7	11	11	7	7	11	11	11	11	70%
12	12	18	12	12	12	39	12	12	12	12	80%
13	11	13	13	13	13	13	13	13	13	13	90%
14	14	14	14	14	41	14	14	14	14	14	90%
15	15	15	15	15	15	15	23	15	15	15	90%
16	16	16	16	16	16	16	16	16	16	16	100%
17	17	17	17	17	17	22	17	17	17	17	90%
18	46	46	46	46	48	48	18	46	46	48	10%
19	19	45	19	19	17	17	19	19	19	19	70%
20	20	15	20	20	20	20	20	20	15	20	80%
21	21	21	21	21	21	21	21	21	21	9	90%
22	22	17	22	17	22	17	17	22	22	22	60%
23	23	17	22	39	23	23	23	17	23	39	50%
24	21	43	24	21	24	21	21	24	24	24	50%
25	1	1	32	1	1	1	3	32	25	1	10%
26	26	26	26	40	26	40	40	26	26	26	70%
27	27	36	36	36	27	35	30	10	27	10	30%
28	28	28	28	28	28	28	28	28	28	28	100%
29	4	9	36	36	21	25	4	25	4	36	0%
30	13	30	33	35	33	10	27	35	13	30	20%
31	31	46	43	43	31	31	31	31	31	31	70%
32	32	32	32	32	32	32	35	32	32	32	90%
33	36	31	31	36	43	43	36	24	24	24	0%
34	34	34	34	34	34	34	34	34	34	34	100%
35	13	35	35	35	35	35	35	13	11	13	60%
36	36	36	36	43	36	36	36	36	21	36	80%
37	37	37	37	37	37	37	37	37	10	37	90%
38	38	38	38	38	38	38	38	38	38	19	90%
39	22	22	39	48	48	18	39	39	39	39	50%
40	40	40	40	40	40	40	40	40	40	40	100%
41	41	14	14	14	14	41	41	14	41	41	50%

42	42	12	42	42	42	42	18	12	18	12	50%
43	46	21	21	21	21	43	21	25	46	43	20%
44	12	17	12	12	17	12	18	12	39	45	0%
45	17	17	19	45	15	23	15	20	20	15	10%
46	22	46	39	48	48	22	46	48	46	48	30%
47	34	34	37	34	34	34	47	37	47	37	20%
48	48	48	48	48	48	48	48	48	48	48	100%



In the future, some statistical method like PCA, ICA can be applied to reduce the feature size and then use SVM for identification.

Fig. 5. Group accuracy for various group population sizes

IV. CONCLUSIONS

In this paper, a statistical based support vector machine algorithm is applied to ECG signal for human identification. We first convert the ECG signal into reduced binary pattern and count the frequency and rank the order. A 48 individuals ECG data are used to training the SVM classifier, which is utilized for matching test of the unknown input ECG signal. The performance of the proposed method is around 60-80%, which is moderately acceptable. Thus the proposed method in the present stage can be used in some situations that not require high accuracy. The main advantage of the proposed algorithm is executed very fast and has less computational complexity.

REFERENCES

[1] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG Analysis: A new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, Vol. 50(3), pp. 808-812, 2001.

[2] S. A. Israel, J. M. Irvine, C. Andrew, D. W. Mark and K. W. Brenda, "ECG to Identify Individuals," *Pattern Recognition*, Vol. 38(1), pp. 133-142, 2005.

[3] S. Saechia, J. Koseeyaporn, P. Wardkein, "Human identification system based ECG signal," *IEEE Region 10 TENCON 2005*, Melbourne, Australia, 21-24 Nov. 2005, pp. 1-4.

[4] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Person identification using electrocardiograms," in *Proceedings: 2006 Canadian Conference on Electrical and Computer Engineering* (CCECE'06), Ottawa, Canada, 7-10 May 2006, pp. 1-4.

[5] H. Silva, H. Gamboa, and A. Fred, "One lead ECG based personal identification with feature subspace ensembles," in *Proceeding: 5th International Conference on Machine Learning and Data Mining in Pattern Recognition* (MLDM 2007), Leipzig, German, 18-20 July 2007, LNAI 4571, pp. 770-783.

[6] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet distance measure for person identification using electrocardiograms," *IEEE Transactions on Instrumentation and Measurement*, Vol. 57(2), pp. 248- 253, 2008.

[7] Y. N. Singh and P. Gupta, "ECG to Individual Identification," in *Proceeding of the 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems* (BTAS 2008), Arlington, Virginia, USA, 29 Sept.-1 Oct. 2008, pp. 1-8.

[8] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram ECG for biometric recognition", *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, Article No. 20, 2008.

[9] P. Sasikala and R. S. D. Wahidauanu, "Identification of individuals using electrocardiogram," *International Journal of Computer Science and Network Security*, Vol. 10(12), pp. 147-153, 2010.

[10] C. Ye, M. T. Coimbra, B. V. K. V. Kumar, "Investigation of human identification using two-lead electrocardigram (ECG)," in *Proceedings on the Fourth IEEE International Conference on Biometrics: Theory Applications and Systems* (BTAS 2010), Washington, DC, USA, 27-29 Sept. 2010, pp. 1-8.

[11] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in D. Haussler (Ed.): Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory (COLT 1992), Pittsburgh, PA, 27-29 July 1992, ACM Press, pp. 144-152.

[12] M. Moavenian and H. Khorrami, "A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification," *Expert Systems with Application*, Vol. 37(4), pp. 3088-3093, 2009.

[13] N. Acir, "A support vector machine classifier algorithm based on a perturbation method and its application to ECG beat recognition systems," *Expert Systems with Application*, Vol. 31(1), pp. 150-158, 2006

[14] S. S. Mehta and N. S. Lingayat, "Combined entropy based method for detection of QRS complexes in 12-lead electrocardiogram using SVM", *Computer Computer in Biology and Medicine*, Vol. 38(1), pp 138-145, January 2008.

[15] S. S. Mehta and N. S. Lingayat, "SVM-based algorithm for recognition of QRS complexes in electrocardiogram," *IRBM*, Vol 29(5), pp 310-317, 2008.

[16] K. Polat and S. Gunes, "Detection of ECG arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine," *Applied Mathematics and Computation*, Vol. 186(1), pp. 898-906, 2007.

[17] H. Zhang and L. Q. Zhang, "ECG analysis based on PCA and support vector machines," in *Proceedings: 2005 International Conference on Neural Networks and Brain* (ICNN&B05), Beijing, China, 13-15 Oct. 2005, pp. 743-747.

[18] J. A. Nasiri and M. Naghibzadeh, "ECG arrhythmia classification with support vector machines and genetic algorithm," in Proceedings on 2009 Third UKSim European Symposium on Computer Modeling and Simulation, Athens, Greece. 25-27 Nov. 2009, pp. 187-192.

[19] E. D. Übeyli, "ECG beats classification using multiclass support vector machines with error correcting output codes," *Digital Signal Processing*, Vol. 17(3), pp.675-784, 2007.

[20] R. Besrour and Z. Lachiri, "ECG Beat classifier using support Vector machine," in *Proceeding: 2008 3rd Information and Commutation Technologies: From Theory to Applications* (ICTTA2008), 7-11 April 2008, Damascus, Syria, pp. 1-5.

[21] C.-C. Chang and C.-J. Lin, LIBSVM - A Library for Support Vector Machines, Version 3.12, http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[22] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, *A Practical Guide to Support Vector Classification*, 2010. http://www.csie.ntu.edu.tw/~cjlin

[23] G. B. Moody, R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, 2001.

國科會補助計畫衍生研發成果推廣資料表

日期:2013/01/01

國科會補助計畫	計畫名稱: 對稱多盤上之頻譜Caratheodory-Fejer插值函數
	計畫主持人: 黃皇男
	計畫編號: 100-2115-M-029-003- 學門領域: 複變函數

無研發成果推廣資料

100 年度專題研究計畫研究成果彙整表

計畫主持人：黃皇男	計畫編號：100-2115-M-029-003-
計畫名稱：對稱多盤上之頻譜 Caratheodory-Fejer 插值函數	

成果項目			量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
			實際已達成數（被接受或已發表）	預期總達成數(含實際已達成數)	本計畫實際貢獻百分比		
國內	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	
		博士生	0	1	0%		由於該生暫時休學，進行論文研究，因此無法執行以博士生聘任，而轉聘其臨時工來協助研究。
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	1	1	100%	篇	投稿到 IET Signal Processing.
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		已經發表在 The Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP-2011)
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	

		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)		本年計畫除執行原有計畫外，並和英國學者 N. J. Young 探討我們理論發展至今已經解決有關 mu 合成的哪些數學理論基礎。					

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估 ■達成目標 □未達成目標（請說明，以 100 字為限） □實驗失敗 □因故實驗中斷 □其他原因 說明：
2. 研究成果在學術期刊發表或申請專利等情形： 論文：□已發表 □未發表之文稿 ■撰寫中 □無 專利：□已獲得 □申請中 ■無 技轉：□已技轉 □洽談中 ■無 其他：(以 100 字為限)
3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）(以 500 字為限)