



HAL
open science

Inequalities related to norm and numerical radius of operators

Tahere Azimi Roushan

► **To cite this version:**

Tahere Azimi Roushan. Inequalities related to norm and numerical radius of operators. Complex Variables [math.CV]. Université de Pau et des Pays de l'Adour; University of Mazandaran, 2020. English. NNT: 2020PAUU3002 . tel-02891958

HAL Id: tel-02891958

<https://theses.hal.science/tel-02891958>

Submitted on 7 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**UNIVERSITÉ DE PAU ET DES PAYS DE LADOUR
ÉCOLE DOCTORALE DES SCIENCES EXACTES ET DE LEURS APPLICATIONS
(UPPA, France)**

**UNIVERSITÉ DE MAZANDARAN, FACULTÉ DES SCIENCES MATHÉMATIQUES
(UMZ, Iran)**

THÈSE

présentée pour obtenir le grade de
DOCTEUR DE UPPA
Spécialité : Mathématiques Appliquées

par

Tahere AZIMI ROUSHAN

Inequalities related to norm and numerical radius of operators

soutenue publiquement le 08 Janvier 2020

Devant la commission d'examen composée de:

M. GUESSAB Allal	Professeur à l'UPPA	Directeur de thèse
M. TAGHAVI Ali	Professeur à l'UMZ	Directeur de thèse
M. BECKER Roland	Professeur à l'UPPA	Présidente
M. BABAKHANI Azizollah	Professeur à l'Université de technologie Babol Noshirvani	Rapporteur
M. TAGHIZADEH Nasir	Professeur à l'Université de Guilan	Rapporteur
M. ALIZADEH AFROUZI Ghasem	Professeur à l'UMZ	Examineur
M. ONDAMI Bienvenu	Maître de Conférences à l'UPPA	Examineur

**Travail effectué au sein du Faculté des Sciences Mathématiques, Université de Mazandaran et
Laboratoire de Mathématiques et de leurs applications (LMAP)**

Dedication

This dissertation work is dedicated to:

- **My dear mother and father**
- **Prof. Vahid Darvish**

Acknowledgements

First and foremost, I would like to thank God, for all the blessings he has given me.

I consider myself so blessed for having many nice people around me who helped me to reach at this stage. I would like to thank all those people. There are so many whose their encouragements and supports have made this work possible and to whom I owe gratitude. They are either in Iran or in France, the places I worked on my PhD project.

Firstly, I would like to address my sincere gratitude to my supervisors, Prof. Allal Guessab and Prof. Ali Taghavi for their patience, support, encouragements and their valuable guidance.

Many thanks must go also to the rest of my thesis committee, Prof. Ghasem Alizadeh Afrouzi, Prof. Azizollah Babakhani, Prof. Roland Becker, Prof. Bienvenu Ondami, Prof. Nasir Taghizadeh for their precious time given for reading my dissertation, insight comments and hard questions.

My sincere thanks also go to Prof. Vahid Darvish for his encouragement, collaboration and valuable advice.

I would also like to express my gratitude and appreciation to the director of the Laboratory of Mathematics and their Applications of Pau, Prof. Gilles Carbou, for accepting me into his laboratory and for his very moral behaviour and support.

I address my gratitude and appreciation to Prof. Mamoud Azizi, Prof. Afshin Babaei, Prof. Asghar Talebi and Mrs. Chehreh Sadeghi for all their efforts in this co-supervision program. As well, my sincere thanks go to Prof. Isabelle Baraille and Prof. Sophie Mercier for their permission in this scientific cooperation and my special thanks go to Mrs. Jacqueline Petitbon who did all her best in this Doctoral program.

My thanks and appreciation for Faculty of Mathematical science, University of Mazandaran and of exact sciences of Pau University, which have significantly contributed in my scientific training through numerous training programs during my thesis. My thanks also to all the teaching staff of the laboratory, in particular Mrs. Sylvie Berton and Mr. Bruno Demoisy.

A very special thanks to my family, particularly, to my mother and father, for their kindness, continued support and encouragement. I would also like to thank my friends and lab mates in both universities, with whom I had a very good time, I wish all happiness and full success in their personal and professional life.



Abstract

In this thesis, after expressing concepts and prerequisites, we investigate Hermite-Hadamard inequality for geometrically convex functions. Then, by introducing operator geometrically convex functions, we extend the results and prove Hermite-Hadamard type inequality for these kind of functions. In the following, by proving the log-convexity of some functions which are based on the unitarily invariant norm and considering the relation between geometrically convex functions and log-convex functions, we present several refinements for some well-known operator norm inequalities. We discuss about the numerical radius of an operator which is equivalent with the operator norm and we state some related results, and by obtaining some upper bounds for the Berezin number of an operator which is contained in the numerical range of that operator, we finish this discussion. In the sequel, we prove operator version of some numerical results, which were obtained for approximating a class of convex functions, as an application, we refine Hermite-Hadamard inequality for a class of operator convex functions.

Approximation of integrals of multivariate functions is a notoriously difficult tasks and satisfactory error analysis is far less well studied than in the univariate case. We propose a method to approximate the integral of all strongly convex function which are continuously differentiable functions with Lipschitz continuous gradients by a family of multi-dimensional numerical integration formulas (cubature formulas). We attempt to quantify their sharp approximation errors. Moreover, we show that how the Delaunay triangulation is used as a convenient partition of the integration domain for constructing the best piecewise cubature formulas.



Résumé

Dans cette thèse, après la présentation des notions et des introductions nécessaires, nous étudierons l'inégalité Hermite-Hadamard pour quelques opérateurs linéaires. Ensuite, nous développerons des résultats en introduisant la fonction convexe géométrique et nous confirmerons des inégalités de type Hermite-Hadamard pour ce type de fonctions. Puis nous proposerons des améliorations de certaines inégalités pour les opérateurs bien étudiés dans la littérature, en montrant le rôle convexe logarithmique de quelques fonctions. Nous considérerons aussi le lien entre les fonctions convexes géométriques et les fonctions logarithmiques et discuterons de quelques propriétés d'image et rayon numérique. Les bornes supérieures du nombre Berezin d'un opérateur seront également données. Dans ce qui suit, nous confirmerons les résultats numériques obtenus pour approcher une classe de fonctions convexes pour les fonctions d'opérateur.

Les méthodes de quadrature multidimensionnelle jouent un rôle important, voire fondamental, en analyse numérique. Nous proposerons une méthode d'approximation de l'intégrale des fonctions fortement convexes avec une famille de formules d'intégration numérique multidimensionnelles. Nous quantifierons leurs erreurs d'approximation. De plus, nous montrons que la triangulation de Delaunay, utilisée comme partition du domaine d'intégration, génère les meilleures formules d'intégration.

Contents

Acknowledgements	i
Abstract (English)	iii
Résumé (Français)	v
Table of contents	vii
Notations	ix
Introduction (English)	1
Introduction (Français)	5
1 Introductory concepts	9
1.1 Convex functions and some related inequalities	9
1.2 Concepts of operator theory and operator inequality	18
1.3 Operator norms	22
Bibliography	27
2 Hermite-Hadamard inequality for operator geometrically convex functions	29
2.1 Inequalities for geometrically convex functions	29
2.2 Operator geometrically convex functions	31
2.2.1 Inequalities for operator geometrically convex functions	35
2.2.2 Examples of operator geometrically convex functions	42
Bibliography	45
3 Refinements for some inequalities related to unitarily invariant norm	47
3.1 Some inequalities for arithmetic-geometric convex functions	47
3.2 Inequalities related to unitarily invariant norms of operators	49
Bibliography	55
	vii

Contents

4	Extension of some inequalities to the numerical radius version	57
4.1	Numerical radius inequalities via convexity	58
4.2	Some upper bounds for the Berezin number of an operator	62
4.2.1	Main results	64
4.2.2	Additional results	67
	Bibliography	71
5	Approximation of differentiable convex functions	73
5.1	Convex polytopes	73
5.2	Barycentric coordinates	75
5.3	Generalized barycentric coordinates on polytopes	76
5.4	Approximation of a class of convex functions	77
5.5	On a generalization of an approximation operator defined by A. Guessab	82
	Bibliography	93
6	Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration	95
6.1	Introduction, motivation and terminology	95
6.2	General setting	100
6.3	Characterization of snd-cubature formulas in terms of the existence of a set of barycentric coordinates	105
6.4	Integral Approximation using barycentric coordinates	106
6.5	Practical Construction of snd-Cubature Formulas	110
6.6	Numerical experiments in $3D$	112
	Bibliography	115
	Conclusion and final remarks	117

Notations

General notations

\mathbb{R}	the set of real numbers
$\mathbb{R}^+ = \{x \in \mathbb{R} x \geq 0\}$	the set of all nonnegative real numbers
\mathbb{R}^d	d -dimensional Euclidean space
$\mathbf{x} = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$	element of \mathbb{R}^d
$\ \mathbf{x}\ = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$	the Euclidean norm of \mathbf{x} in \mathbb{R}^d
$[\mathbf{x}, \mathbf{y}]$	the line segment between \mathbf{x} and \mathbf{y}
$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \frac{\partial f}{\partial x_2}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_d}(\mathbf{x}) \right)$	the gradient of f at \mathbf{x}
$\nabla^2 f(\mathbf{x})$	the Hessian matrix of f at \mathbf{x} with the matrix elements given by $\nabla^2 f(\mathbf{x})_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}), i = 1, \dots, d, j = 1, \dots, d$
conv	the convex hull of a finite point set
ext	the set of all extreme points of
vert	the set of all vertices of
$a(\cdot)$	any affine function

Notations

Del, Dt	Delaunay triangulation
Ω°	the interior of Ω
$ \Omega $	measure of Ω
$C(\Omega)$	all real-valued continuous functions on Ω
$C^k(\Omega)$	all functions which are k times continuously differentiable, where $k \in \mathbb{N}$
$L^1(\Omega)$	all Lebesgue integrable functions
$C^{1,1}$	the subclass of all functions which are continuously differentiable with Lipschitz continuous gradients
$T(f)$	the trapezoid rule
$ A $	the absolute value of operator A
$ A $	the unitarily invariant norm of operator A
$\ A\ $	the operator norm of A
s_j	the singular value
$\text{Sp}(A)$	the spectrum of operator A
$M_n(\mathbb{C})$	the algebra of all $n \times n$ complex matrices
\mathcal{H}	the Hilbert space
$\mathbb{B}(\mathcal{H})$	the set of all bounded linear operators from \mathcal{H} into \mathcal{H}
x^*	the adjoint of x
$\mathbb{B}(\mathcal{H})^+$	the set of all positive operators in $\mathbb{B}(\mathcal{H})$

$\mathbb{B}(\mathcal{H})^{++}$	the set of all strictly positive operators in $\mathbb{B}(\mathcal{H})$
$\mathbb{B}_{sa}(\mathcal{H})$	the set of all self-adjoint elements in $\mathbb{B}(\mathcal{H})$
$r(A)$	the spectral radius of A
\sharp_t	the t -weighted geometric mean
\sharp	the geometric mean
$OC^{1,1}$	the subclass of all operator functions which are continuously differentiable with Lipschitz continuous gradients
$\text{diag}(\lambda_i)$	the diagonal matrix with entries λ_i
$S_1(\mathcal{H})$	the set of all trace class operators
$S_1(\mathcal{H})^*$	the dual space of $S_1(\mathcal{H})$
$w(\cdot)$	the numerical radius

Introduction

Operator theory and matrix analysis are important research areas which are of interest to many mathematicians and have many applications in various sciences including scientific computing, operations research, mathematical physics, control and systems theory, statistics, economics and engineering disciplines.

Many theorems in matrix analysis appear in the form of inequalities, and one of the fundamental fields of research in this area is matrix and operator inequality. In fact, it can be said that matrix inequalities reflect matrix analysis from a quantitative perspective.

Inequalities can be traced in almost all areas of pure and applied mathematics. In some related fields such as differential equations theory, calculus of variations and geometry, efforts have been made to extend and find more suitable forms than classical examples.

The concept of convex functions has indeed found an important place in Modern Mathematics as can be seen in a large number of research articles and books devoted to the field these days. Also, the Hermite-Hadamard inequality, which, we can say, is the first fundamental result for convex functions with a natural geometrical interpretation and many applications, has attracted and continues to attract much interest in elementary mathematics. So, many mathematicians have devoted their efforts to generalize, refine, counterpart and extend it for different classes of functions.

What we are going to do in this thesis is to find an operator version of some numerical inequalities. It also attempts to introduce new inequalities that improve or generalize previous results.

In applying inequalities for operators we should be careful extremely because the operator version of an inequality is not necessarily established. For example, the following inequality holds for any pair of positive integers a and b :

$$|a - b| \leq a + b.$$

The operator version is expected to be as follows

$$|A - B| \leq A + B, \tag{1}$$

where A and B are positive operators and $|\cdot|$ represents the second positive root of the operator.

But considering matrices such as $A = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$. It can be shown that the inequality (1) is not always true. Maybe if we use the singular value version (instead of the matrix itself), we get the result:

$$s_j(A - B) \leq s_j(A + B), \quad 1 \leq j \leq n. \quad (2)$$

But the above examples still violate the above inequality.

The weaker version of the inequality (2) can be stated as follows:

$$\| \|A - B\| \| \leq \| \|A + B\| \|,$$

where $\| \cdot \|$ is the unitarily invariant norm. The above inequality is established in (R. Bhatia, F. Kittaneh, The matrix arithmetic-geometric mean inequality revisited, *Linear Algebra Appl.* 428 (2008) 2177-2191).

In the first chapter, the concepts, basic definitions and principles needed for the following chapters are presented.

In the second chapter, we discuss Hermite-Hadamard inequality for geometrically convex functions, then by providing a definition of the operator geometrically convex functions, we prove the Hermite-Hadamard type inequality for such functions and by deriving some examples of operator geometrically convex functions and presenting an application of the result to the operator norm we will finish this chapter.

In Chapter 3, we first discuss the relationship between geometrically convex functions and logarithmic convex functions, and present the Hermite-Hadamard type inequality for logarithmic convex functions. Then, by proving the log-convexity of some functions which are based on the unitarily invariant norm we present several refinements for some well-known operator norm inequalities.

In Chapter 4, we consider the equivalent norm of the operator norm, i.e., the numerical radius norm, and study the numerical radius inequalities which are an extension of some unitarily invariant norms inequalities to the context of numerical radius, and by giving contradictory examples we show that due to weak unitary invariance of the numerical radius, these extended inequalities will be weaker than the corresponding unitarily invariant norm versions. What's more, we introduce the Berezin number of an operator, which is a subset of the numerical range of that operator, and prove inequalities based on the Berezin number.

In Chapter 5, we prove operator version of some numerical results, which were obtained for approximating a class of convex functions, as an application, we refine Hermite-Hadamard inequality for a class of operator convex functions.

In Chapter 6, we provide a necessary and sufficient condition for the existence of the cubature formulas to approximate the integral of all strongly convex function which are continuously differentiable functions with Lipschitz continuous gradients. We use the generalized barycentric coordinates to construct a multivariate version of the classical trapezoidal rule in

arbitrary higher-dimensional polytopes. As a result, we get explicit lower and upper bounds for the approximation error. Indeed, analogously to the one-dimensional estimates, we offer sharp error estimates which only depend on the parameter of the strong convexity, the Lipschitz constants of the gradients and the error associated with using the quadratic function. In the following, by using the Delaunay triangulation as a partition of a polytope, we present an explicit construction of our sharp cubature schemes. Finally, we provide a numerical example to illustrate the efficiency of this approach.

Contributions

The following articles are extracted from this thesis:

1- A. Taghavi, V. Darvish, T. Azimi Roushan, *Hermite-Hadamard Type Inequalities For Operator Geometrically Convex Functions II*, Kragujevac Journal of Mathematics, 45(2) (2021) 191-202.

2- A. Taghavi, T. Azimi Roushan, V. Darvish, *Some Refinements For The Arithmetic-Geometric Mean and Cauchy-Schwarz Matrix Norm Interpolating Inequalities*, Bulletin of the Iranian Mathematical Society, 44(4) (2018) 927-936.

3- A. Taghavi, T. Azimi Roushan, V. Darvish, *Some upper bounds for the Berezin number of Hilbert space operators*, Filomat, 33(14) (2019) 4353-4360.

4- A. Taghavi, T. Azimi Roushan, *On a generalization of an approximation operator defined by A. Guessab*, Linear and Multilinear Algebra, 67(11) (2019) 1-13.

5- A. Guessab, T. Azimi Roushan, *Generalized barycentric coordinates and sharp strongly negative definite multi-dimensional numerical integration*, Accepted for publication in a Springer volume, entitled: Approximation theory and analytic inequalities, (2019).

Introduction générale

La théorie des opérateurs et l'analyse des matrices sont des domaines importants de la recherche. Attirantes pour pas mal de mathématiciens, elles ont de nombreux emplois dans les différentes sciences entre autres les calculs scientifiques, la recherche des opérations, la physique mathématique, la théorie du système et le contrôle, la statique, l'économie et les génies.

Dans l'analyse matricielle, beaucoup de théories apparaissent dans le cadre des inégalités de sorte que les inégalités fonctionnelles et matricielles constituent un des centres privilégiés de recherche de ce domaine. À dire vrai, les inégalités matricielles reflètent l'analyse matricielle du point de vue quantitatif.

Les inégalités sont présentes dans la quasi-totalité des domaines des mathématiques pures et appliquées. Cela est plus visible dans la théorie des équations différentielles, des calculs des variations, de la géométrie pour trouver des exemples plus convenables que les exemples classiques.

La théorie des fonctions convexes a pris une place de plus en plus importante dans les mathématiques modernes. On peut voir beaucoup de livres et d'articles consacrés à ce sujet. De plus, l'inégalité Hermite-Hadamard avec son interprétation géométrique naturelle et sa vaste utilisation qui sont le résultat de l'utilisation des fonctions convexes a attiré l'attention des chercheurs, ce qui est constant dans les mathématiques élémentaires. Ainsi, la plupart des mathématiciens travaillent afin de développer, améliorer et de généraliser cette notion pour les autres catégories des fonctions.

Dans cette thèse, nous essayons de trouver l'opération de certaines inégalités mathématiques et de présenter de nouvelles inégalités qui aboutissent à l'amélioration ou à la généralisation des résultats précédents.

Il faut faire le maximum d'attention dans l'utilisation des inégalités pour les opérateurs car l'opération de chaque inégalité n'existe pas forcément. En guise d'exemple, l'inégalité suivante est correcte pour les deux nombres positifs a et b

$$|a - b| \leq a + b.$$

Introduction (Français)

Alors on peut penser que son opération sera ainsi :

$$|A - B| \leq A + B, \quad (3)$$

où A et B des opérateurs positifs et $|\cdot|$ représentent la racine carrée positive de l'opérateur.

Mais en considérant les matrices comme $A = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$. On peut montrer que l'inégalité (3) n'est pas toujours correcte. Les valeurs singulières pourraient nous conduire au résultat, c'est-à-dire

$$s_j(A - B) \leq s_j(A + B), \quad 1 \leq j \leq n. \quad (4)$$

Mais les exemples ci-dessus confirment l'exactitude de l'inégalité mentionnée.

Le cas plus faible que l'inégalité (4) est ainsi

$$\| \|A - B\| \| \leq \| \|A + B\| \|,$$

qui est $\| \|\cdot\| \|$ de la norme stable. L'exactitude de l'inégalité ci-dessus est confirmée dans (R. Bhatia, F. Kittaneh, la matrice arithmétique-géométrique veut dire l'inégalité révisée, *l'algèbre linéaire d'Apple*. 428 (2008) 2177-2191).

Dans le premier chapitre, il y a des définitions et des principes nécessaires pour les chapitres suivants.

Le deuxième chapitre concerne l'étude de l'inégalité Hermite-Hadamard pour les fonctions convexes géométriques puis en donnant une définition de la fonction convexe géométrique opératrice, nous confirmons l'inégalité Hermite-Hadamard pour ces types de fonction, enfin nous terminons le chapitre en donnant des exemples des fonctions convexes géométriques et en présentant un de leurs emplois selon la norme opératrice.

Le chapitre 3 est commencé par la relation entre des fonctions convexes géométriques et des fonctions convexes logarithmiques ; et la présentation de l'inégalité Hermite-Hadamard pour les fonctions convexes logarithmiques. Ensuite, nous présentons l'amélioration du cas normatif de plusieurs inégalités opérationnelles célèbres en confirmant le rôle convexe et logarithmique de certaines fonctions classées selon la norme stable.

Dans le 4 chapitre, nous considérons la norme équivalente de la norme opérateur, c'est-à-dire la norme du rayon numérique et aussi le rayon numérique de plusieurs inégalités confirmées par la norme stable. En plus, en donnant des contre-exemples, nous montrons que l'équation du rayon de ces inégalités n'est pas toujours correcte. Ensuite, nous confirmons la version faible à l'aide des théorèmes. La dernière partie de ce chapitre est consacré à la présentation du nombre Berezin d'un opérateur et nous confirmons des inégalités selon ce nombre.

Dans le chapitre 5, nous essayons d'étudier les résultats numériques pour approcher une catégorie des fonctions convexes à la notion de la fonction réelle pour confirmer ensuite l'opération des résultats, enfin nous présentons un emploi des résultats afin d'améliorer l'inégalité Hermite-Hadamard.

Dans le chapitre 6, nous fournissons une condition nécessaire et suffisante pour l'existence des formules cubiques pour se rapprocher de l'intégrale de toutes les fonctions fortement convexes qui sont continuellement différentiables. fonctionne avec les gradients continus de Lipschitz. Nous utilisons la méthode généralisée pour construire une version multivariable de la règle trapézoïdale classique en des polytopes arbitraires de dimensions supérieures. Par conséquent, nous obtenons des limites inférieures et supérieures explicites pour l'erreur d'approximation. En effet, comme pour les estimations unidimensionnelles, nous proposons des estimations d'erreurs nettes qui ne dépendent que du paramètre de la forte convexité, le Constantes de Lipschitz des gradients et de l'erreur associée à l'utilisation de la fonction quadratique. Dans ce qui suit, en utilisant la triangulation Delaunay comme partition d'un polytope, nous présentons une construction explicite de nos schémas cubiques aiguisés. Enfin, nous donnons un exemple numérique pour illustrer l'efficacité de cette approche.

Contributions de la thèse

Les articles suivants sont extraits de cette thèse :

1- A. Taghavi, V. Darvish, T. Azimi Roushan, *Hermite-Hadamard Type Inequalities For Operator Geometrically Convex Functions II*. Kragujevac Journal of Mathematics, 45(2) (2021) 191-202.

2- A. Taghavi, T. Azimi Roushan, V. Darvish, *Some Refinements For The Arithmetic-Geometric Mean and Cauchy-Schwarz Matrix Norm Interpolating Inequalities*. Bulletin of the Iranian Mathematical Society, 44(4) (2018) 927-936.

3- A. Taghavi, T. Azimi Roushan, V. Darvish, *Some upper bounds for the Berezin number of Hilbert space operators*, Filomat, 33(14) (2019) 4353-4360.

4- A. Taghavi, T. Azimi Roushan, *On a generalization of an approximation operator defined by A. Guessab*, Linear and Multilinear Algebra 67(11) (2019) 1-13.

5- A. Guessab, T. Azimi Roushan, *Generalized barycentric coordinates and sharp strongly negative definite multi-dimensional numerical integration*, Accepté pour dans un volume de Springer, intitulé: Approximation theory and analytic inequalities, (2019).

Introductory concepts

In this chapter, we describe the concepts and definitions which are required for our results in the forthcoming chapters. Also, some applicable theorems used to obtain the results of the following chapters are presented just by referring to references without proof.

1.1 Convex functions and some related inequalities

Using the standard notation \mathbb{R} for the set of real numbers, a real-valued function f on an interval $\mathcal{I} \subset \mathbb{R}$ is said to be convex if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad (1.1)$$

for $x, y \in \mathcal{I}$ and $\lambda \in [0, 1]$. Also f is said to be concave if the inequality (1.1) be valid conversely, or in other words, $-f$ be convex function.

The functions $f(x) = x^2$ and exponential function $g(x) = e^x$ are two familiar examples of convex functions.

In the d -dimensional Euclidean space, \mathbb{R}^d , a set $C \subset \mathbb{R}^d$ is said to be convex if the inequality

$$(1 - \lambda)x + \lambda y \in C,$$

holds for every $x, y \in C$ and $\lambda \in [0, 1]$.

One of the simple characterizations of a convex set in \mathbb{R}^d is that it is closed under any convex combinations of its elements, and this is proved in [14] by induction.

Definition 1.1.1 *Let $C \subset \mathbb{R}^d$ be a convex set. A point $x \in C$ is called an extreme point of C if $x = ty + (1 - t)z$ for $y, z \in C$ and $t \in (0, 1)$ implies $x = y = z$.*

Compact convex sets can be described via their extreme points as following.

Theorem 1.1.2 *(Krein-Milman theorem)[2] Let $C \subset \mathbb{R}^d$ be a compact convex set. Then the set*

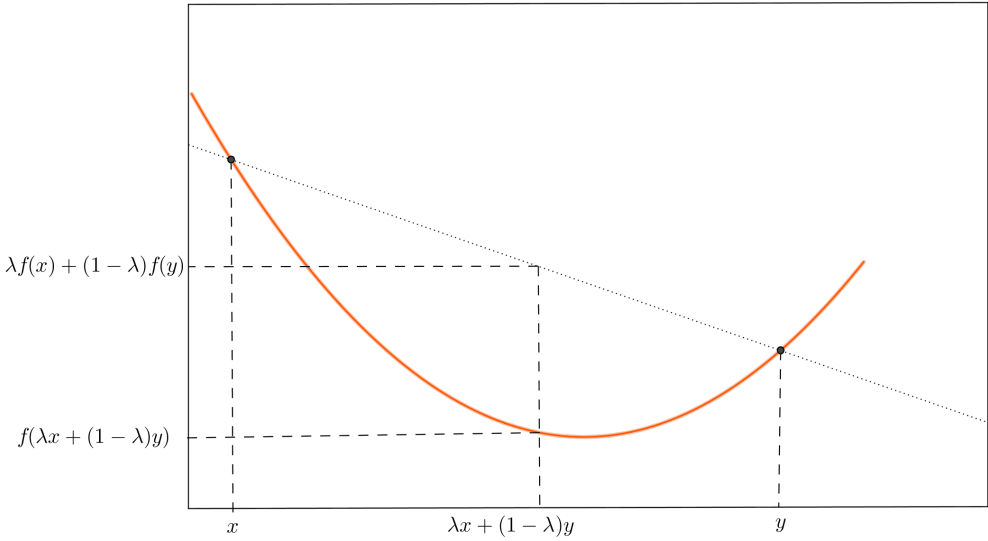


Figure 1.1 – Illustration of the inequality $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$.

of extreme points of C is not empty. Furthermore, every $x \in C$ may be expressed as a convex combination of finitely many extreme points of C .

A real-valued function $f : C \rightarrow \mathbb{R}$ on a convex set $C \subset \mathbb{R}^d$ is said to be convex if it satisfies on inequality (1.1) for every $x, y \in C$ and $\lambda \in [0, 1]$.

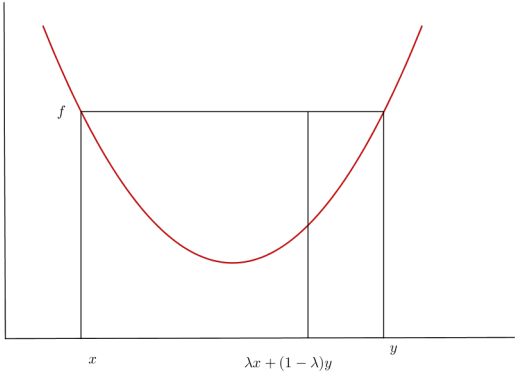


Figure 1.2 – Convex function.

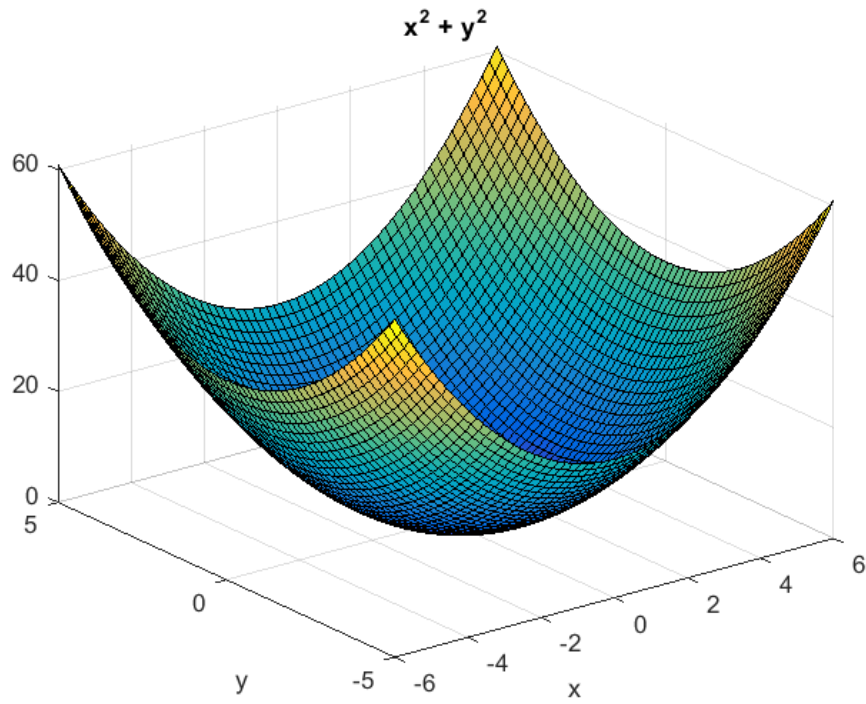


Figure 1.3 – Convex function.

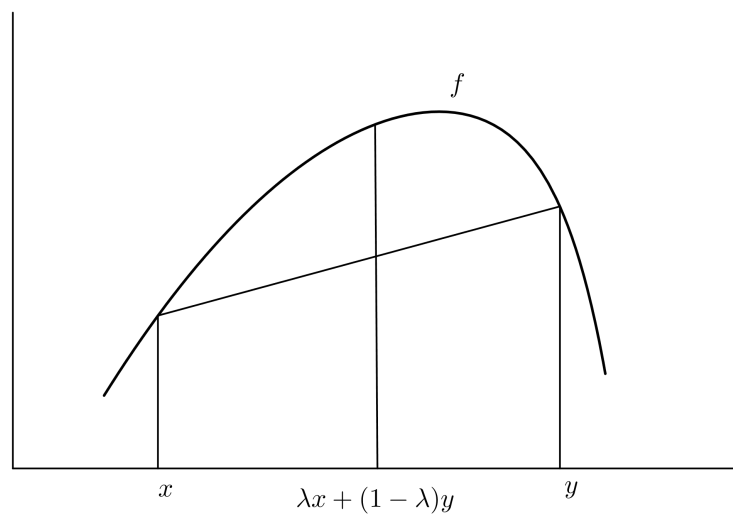


Figure 1.4 – Concave function.

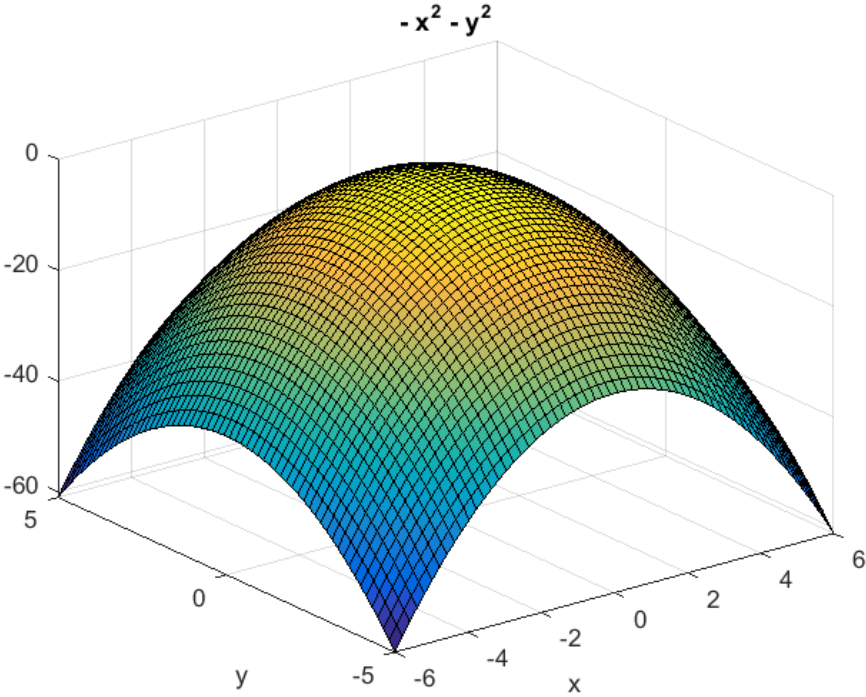


Figure 1.5 – Concave function.

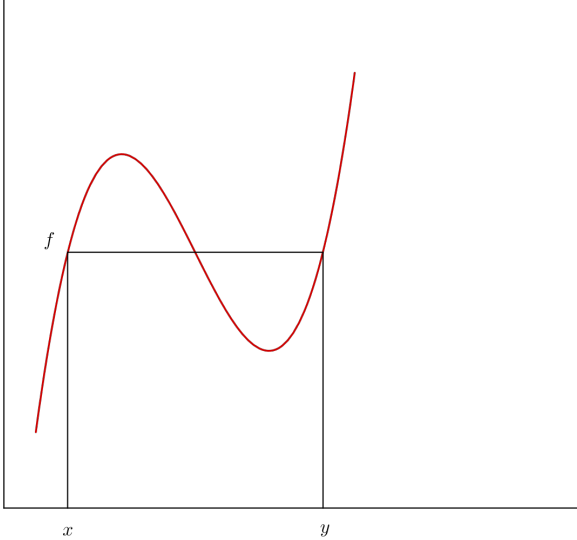


Figure 1.6 – Neither convex nor concave.

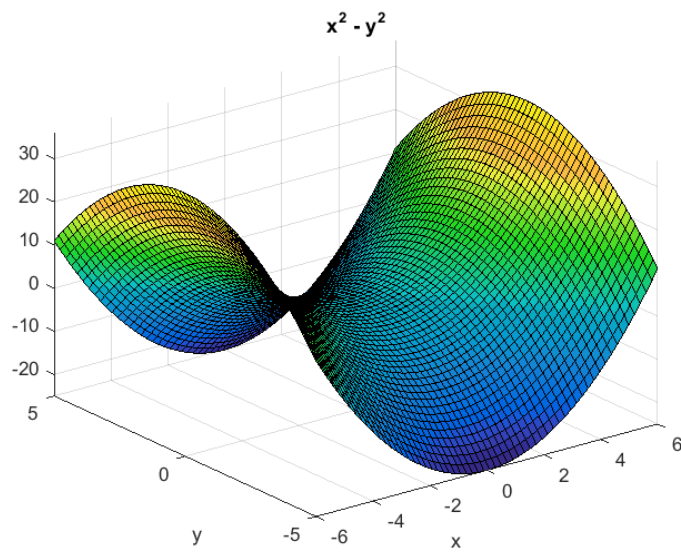


Figure 1.7 – Neither convex nor concave.

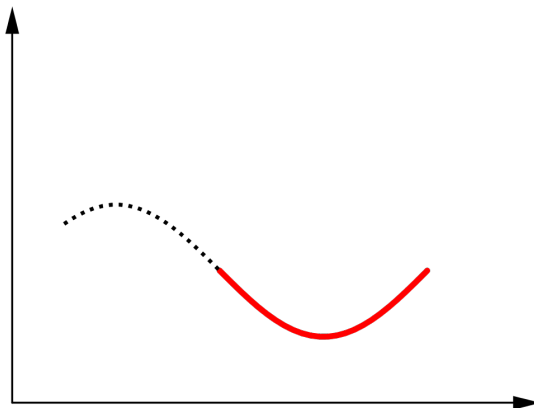


Figure 1.8 – Convex and concave.

The function in Fig.1.2 is convex, Fig.1.4 is concave and Fig.1.6 is neither. The function in Fig.1.8 is convex on the part where it is solid and concave on the part where it is dotted.

Continuity is an important property of a convex function:

Introductory concepts

Theorem 1.1.3 [16] Let $C \subset \mathbb{R}^d$ be an open convex set. Let $f : C \rightarrow \mathbb{R}$ be convex. Then f is continuous on C .

For a continuously differentiable function f defined on an open convex set $C \subset \mathbb{R}^d$, we denote its gradient at $x \in C$, by

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

We say that f is continuously differentiable on C if it is continuously differentiable on an open set containing C .

Checking that a function is convex or not is not very easy, but fortunately several useful criteria are available. Probably the simplest one is the following:

Theorem 1.1.4 [14] Let $f : \mathcal{I} \rightarrow \mathbb{R}$ be a continuous function. Then f is convex if and only if f is midpoint convex, that is,

$$x, y \in \mathcal{I} \text{ implies } f\left(\frac{x+y}{2}\right) \leq \frac{f(x)+f(y)}{2}.$$

The following theorem gives the first order characterization of convex functions.

Theorem 1.1.5 (The first-order condition for convexity: Gradient inequality)[5] Let $C \subset \mathbb{R}^d$ be a nonempty open convex set and let $f : C \rightarrow \mathbb{R}^d$ be a continuously differentiable function. Then f is convex if and only if for any $x, y \in C$ we have

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

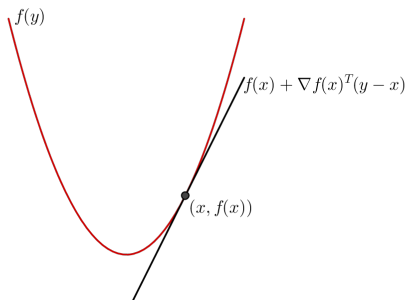


Figure 1.9 – The first order characteristic of a convex function f .

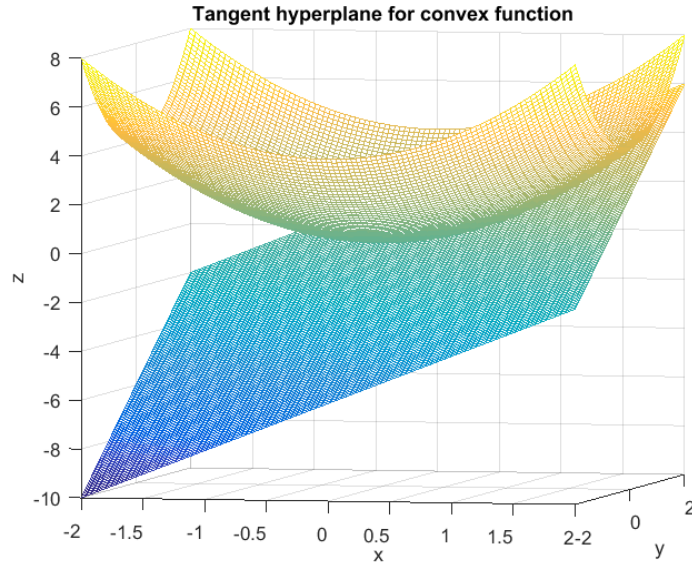


Figure 1.10 – The first order characteristic of a convex function f .

A mapping $F : D \subset \mathbb{R}^d \rightarrow \mathbb{R}^d$ is increasing on D if

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in D. \quad (1.2)$$

By letting $F = \nabla f$ in the above definition, then we conclude the following theorem which states that ∇f of a convex function f is an increasing mapping.

Theorem 1.1.6 [5] Suppose $C \subset \mathbb{R}^d$ be a nonempty open convex set and let $f : C \rightarrow \mathbb{R}^d$ be a continuously differentiable function. Then f is convex on C if and only if its gradient ∇f is increasing on C , i.e.,

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0, \quad \forall x, y \in C.$$

We will now characterize convex functions in terms of their Hessian matrices.

Theorem 1.1.7 (The second-order condition for convexity) [5] Let $C \subset \mathbb{R}^d$ be a nonempty open convex set, and let $f : C \rightarrow \mathbb{R}^d$ be twice continuously differentiable in C . Then, f is convex on C if and only if $\nabla^2 f(x)$ is positive semi-definite for all $x \in C$.

Convex functions play an important role in the inequality theory. In the following, we present two famous inequalities related to convex functions.

Theorem 1.1.8 (Jensen inequality)[2] Let $C \subset \mathbb{R}^d$ is a convex set and $f : C \rightarrow \mathbb{R}$ be a convex function. If $\lambda_1, \lambda_2, \dots, \lambda_k \in [0, 1]$, $\sum_{i=1}^k \lambda_i = 1$, and $x_1, x_2, \dots, x_k \in C$, then

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i). \quad (1.3)$$

Introductory concepts

In the above theorem, suppose $f(x) = -\ln x$ on the interval $\mathcal{I} \subset \mathbb{R}^+$ which \ln is a logarithmic function. Then we have

$$a^\nu b^{1-\nu} \leq \nu a + (1-\nu)b, \quad (1.4)$$

for $a, b \in \mathcal{I}$ and $\nu \in [0, 1]$, which is famous for Young inequality.

In the especial case $\nu = \frac{1}{2}$, we derive the following inequality which is known as arithmetic-geometric mean inequality:

$$\sqrt{ab} \leq \frac{a+b}{2}.$$

Theorem 1.1.9 (Hermite-Hadamard inequality)[16] *If $f : \mathcal{I} \rightarrow \mathbb{R}$ be a convex function and $a, b \in \mathcal{I}$ with $a < b$, then*

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a} \int_a^b f(x) dx \leq \frac{f(a)+f(b)}{2}. \quad (1.5)$$

We remark that the above inequalities have been studied deeply in the literature. And there has been a great deal of efforts in refining and extending them, among which we can refer the reader to [1, 4, 7, 5, 6, 8, 18] as a sample of recent work treating such inequalities.

Now we present the notion of strong convexity, which generalizes the classical definition of convexity.

Definition 1.1.10 *Let $C \subset \mathbb{R}^d$ be nonempty convex set. A function $f : C \rightarrow \mathbb{R}$ is μ -strongly convex with convexity parameter $\mu > 0$ if the strong Jensen inequality holds: for any $\mathbf{x}, \mathbf{y} \in C$ and $t \in [0, 1]$,*

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) - \frac{\mu}{2}t(1-t)\|\mathbf{x} - \mathbf{y}\|^2. \quad (1.6)$$

Obviously, every strongly convex function is convex, or in the other words, the set of strongly convex functions is contained in the set of convex functions. Observe also that, for instance, affine functions are not strongly convex and if $\mu = 0$, we can get the classical definition of convexity.

Remark 1.1.11 *For any positive real number μ the following functions are μ -strongly convex functions:*

1. $\frac{\rho}{2} \|\cdot\|^2, (\mu \leq \rho)$.
2. *Addition of a convex function to a strongly convex function gives a strongly convex function with the same modulus of strong convexity. Therefore, adding a convex function to $\frac{\mu}{2} \|\cdot\|^2$ does not affect μ -strong convexity.*

1.1. Convex functions and some related inequalities

Throughout this thesis, for every point (vector) in d -dimensional Euclidean space, $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$, $\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_d^2}$ is the Euclidean norm. Also, for $x, y \in \mathbb{R}^d$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_d y_d$ is the usual scalar product on \mathbb{R}^d . Here we state a result concerning the Euclidean norm.

Lemma 1.1.12 [12] *Let u, v be in \mathbb{R}^d and $t \in [0, 1]$. Then the following identity holds:*

$$(1-t)\|x\|^2 + t\|y\|^2 - \|(1-t)x + ty\|^2 = t(1-t)\|x-y\|^2. \quad (1.7)$$

According to Definition 1.1.10 and identity (1.7), we can conclude that μ -strong convexity of f is equivalent to the convexity of $g := f - \frac{\mu}{2}\|\cdot\|^2$. Indeed, if f is μ -strongly convex, then by identity (1.7) we get

$$\begin{aligned} g(tx + (1-t)y) &= f(tx + (1-t)y) - \frac{\mu}{2}\|tx + (1-t)y\|^2 \\ &\leq tf(x) + (1-t)f(y) - \frac{\mu}{2}t(1-t)\|x-y\|^2 - \frac{\mu}{2}\|tx + (1-t)y\|^2 \\ &= tf(x) + (1-t)f(y) - \frac{\mu}{2}t\|x\|^2 - \frac{\mu}{2}(1-t)\|y\|^2 \\ &= tg(x) + (1-t)g(y), \end{aligned}$$

which yields the convexity of g . Conversely, if g is a convex function, then again by Lemma 1.1.12 we have

$$\begin{aligned} f(tx + (1-t)y) &= g(tx + (1-t)y) + \frac{\mu}{2}\|tx + (1-t)y\|^2 \\ &\leq tg(x) + (1-t)g(y) + \frac{\mu}{2}(t\|x\|^2 + (1-t)\|y\|^2 - t(1-t)\|x-y\|^2) \\ &= t(g(x) + \frac{\mu}{2}\|x\|^2) + (1-t)(g(y) + \frac{\mu}{2}\|y\|^2) - \frac{\mu}{2}t(1-t)\|x-y\|^2 \\ &= tf(x) + (1-t)f(y) - \frac{\mu}{2}t(1-t)\|x-y\|^2, \end{aligned}$$

which shows that f is strongly convex with convexity parameter μ .

When the function is differentiable, an alternative characterization of strong convexity is in terms of the gradient inequality as following.

Theorem 1.1.13 [12] *Let f be a continuously differentiable function defined on an open convex set $C \subset \mathbb{R}^d$. Then f is strongly convex with parameter $\mu > 0$ if and only if for any $x, y \in C$ we have*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|x - y\|^2.$$

The following result is a characterization of strongly convex functions via the strong mono-

tonicity of the gradient.

Theorem 1.1.14 [5] *Let f be a continuously differentiable function defined on an open convex set $C \subset \mathbb{R}^d$. Then f is strongly convex with parameter $\mu > 0$ if and only if its gradient ∇f is uniformly monotone i.e.,*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|^2, \quad x, y \in C.$$

1.2 Concepts of operator theory and operator inequality

Let $\mathbb{B}(\mathcal{H})$ stand for algebra of all bounded linear operators on a complex Hilbert space \mathcal{H} and assume that operator multiplication corresponds to operator composition. Then by operator norm

$$\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\},$$

$\mathbb{B}(\mathcal{H})$ is a Banach space.

Definition 1.2.1 [20] *A Banach algebra is a algebra \mathcal{A} over the field of complex numbers \mathbb{C} such that*

1. \mathcal{A} be a Banach space with the norm $\|\cdot\|$,
2. for every x and y in \mathcal{A} , $\|xy\| \leq \|x\|\|y\|$.

Whatsmore, \mathcal{A} is called unital if it has a unit e . In this case, it should be $\|e\| = 1$.

Example 1.2.2 [20]

1. If \mathcal{X} is a compact Hausdorff space, then $C(\mathcal{X})$ (the set of all continuous functions from \mathcal{X} to \mathbb{C}) is a Banach algebra with the sup-norm and pointwise operations. The constant function with value 1 is the multiplicative unit.
2. If \mathcal{X} be a Banach space, then $\mathbb{B}(\mathcal{X})$ (the set of all bounded linear operator from \mathcal{X} into \mathcal{X}) with operator norm and operator multiplication (i.e., composition) is a Banach algebra and the identity operator I is the multiplicative unit. As a result, $\mathbb{B}(\mathcal{H})$ is also a Banach algebra.

Let \mathcal{A} is a Banach algebra and $x \in \mathcal{A}$. Spectrum of x denoted by $\sigma(x)$ (or $\text{Sp}(x)$), and defined as follows:

$$\sigma(x) = \{\lambda \in \mathbb{C} : \lambda e - x \text{ is not invertible in } \mathcal{A}\},$$

whatsmore, for $x \in \mathcal{A}$ we let $r(x)$ denote the spectral radius of it and defined as following:

$$r(x) = \sup\{|\lambda| : \lambda \in \sigma(x)\}.$$

1.2. Concepts of operator theory and operator inequality

Definition 1.2.3 [20] A C^* -algebra is a Banach algebra \mathcal{A} together with a mapping $x \rightarrow x^*$ on \mathcal{A} satisfying the following conditions:

- (a) $(x^*)^* = x$ for all $x \in \mathcal{A}$,
- (b) $(ax + by)^* = \bar{a}x^* + \bar{b}y^*$ for all $x, y \in \mathcal{A}$ and $a, b \in \mathbb{C}$,
- (c) $(xy)^* = y^*x^*$ for all $x, y \in \mathcal{A}$,
- (d) $\|x^*x\| = \|x\|^2$ for all $x \in \mathcal{A}$.

Any mapping $x \rightarrow x^*$ on an algebra satisfying (a), (b), and (c) is called an involution on the algebra. The element x^* is usually called the adjoint of x .

Example 1.2.4 $\mathcal{A} = \mathbb{C}$ with $z^* = \bar{z}$ is the simplest C^* -algebra.

Example 1.2.5 $\mathcal{A} = \mathbb{B}(\mathcal{H})$ with the usual adjoint operation as involution is a C^* -algebra. If \mathcal{H} is n -dimensional, then $\mathbb{B}(\mathcal{H}) = M_n(\mathbb{C})$ is the algebra of all $n \times n$ complex matrices and the adjoint of a matrix in $M_n(\mathbb{C})$ is its conjugate transpose.

Definition 1.2.6 [20] Suppose \mathcal{A} is a C^* -algebra and $x \in \mathcal{A}$.

- (a) We say that x is self-adjoint if $x = x^*$.
- (b) We say that x is normal if $x^*x = xx^*$.
- (c) We say that x is unitary if $x^*x = xx^* = 1_{\mathcal{A}}$, or equivalently $x^* = x^{-1}$.
- (d) We say that x is contraction if $x^*x \leq 1_{\mathcal{A}}$.
- (e) We say that x is positive if $x = y^*y$ for some y in \mathcal{A} .

It is clear that both self-adjoint and unitary elements are normal. It is also clear that positive elements are self-adjoint.

According to part (e) of above definition, an operator $A \in \mathbb{B}(\mathcal{H})$ is said to be positive (semi-definite positive) and denoted by $A \geq 0$ if for all $x \in \mathcal{H}$, $\langle Ax, x \rangle \geq 0$. Also, we say it is strictly positive (definite positive) if for every non-zero $x \in \mathcal{H}$ we have $\langle Ax, x \rangle > 0$.

The positive operator $A \in \mathbb{B}(\mathcal{H})$ is strictly positive ($A \in \mathbb{B}(\mathcal{H})^{++}$) if and only if it is invertible.

Theorem 1.2.7 [20] If $A \in \mathbb{B}(\mathcal{H})$ is normal, then $r(A) = \|A\|$.

Introductory concepts

Theorem 1.2.8 [15] *If $A \in \mathbb{B}(\mathcal{H})$ be a self-adjoint operator, then $\sigma(A) \subseteq \mathbb{R}$. Also, spectrum of a positive operator included in the set of non-negative real numbers.*

*If $A \in \mathbb{B}(\mathcal{H})$ is a positive operator, $A^{\frac{1}{2}}$ is the unique positive element B which $B^2 = A$. We also write $|A| = \sqrt{A^*A}$ be absolute value of operator $A \in \mathbb{B}(\mathcal{H})$. It is clear that $|A| \geq 0$ for all A .*

The mapping $\phi : \mathbb{B}(\mathcal{H}) \rightarrow \mathbb{B}(\mathcal{H})$ is called positive, if for any positive operator $A \in \mathbb{B}(\mathcal{H})$, $\phi(A)$ be positive operator in $\mathbb{B}(\mathcal{H})$.

Definition 1.2.9 [15] *Let \mathcal{X} is a Banach space, we say $T \in \mathbb{B}(\mathcal{X})$ is finite rank, if it has a finite dimensional range.*

Eigenvalues of the self-adjoint operator A in $M_n(\mathbb{C})$ are arranged as $\lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A)$ and the singular values of an arbitrary operator $A \in M_n(\mathbb{C})$ as $s_1(A) \geq s_2(A) \geq \dots \geq s_n(A)$ are the eigenvalues of the positive operator $|A|$.

Definition 1.2.10 [20] *Suppose \mathcal{A} and \mathcal{B} are C^* -algebra. A mapping $\Phi : \mathcal{A} \rightarrow \mathcal{B}$ is said to be a C^* -homomorphism if*

- $\Phi(ax + by) = a\Phi(x) + b\Phi(y)$ for all $a, b \in \mathbb{C}$ and $x, y \in \mathcal{A}$,
- $\Phi(xy) = \Phi(x)\Phi(y)$ for all $x, y \in \mathcal{A}$,
- $\Phi(x^*) = \Phi(x)^*$ for all $x \in \mathcal{A}$,
- Φ maps the unit in \mathcal{A} to the unit in \mathcal{B} .

If Φ is further one-to-one, we say that Φ is a C^ -isomorphism. Two C^* -algebras are C^* -isomorphic if there exists a C^* -isomorphism from one onto the other.*

Let A be a self-adjoint operator on a Hilbert space \mathcal{H} . The Gelfand map [20] establishes a C^ -isometrically isomorphism ϕ between the set $C(\text{Sp}(A))$ of all continuous functions defined on the spectrum of A and the C^* -algebra $C^*(A)$, generated by A and the identity operator I , such that for all $f, g \in C(\text{Sp}(A))$ and $\alpha, \beta \in \mathbb{C}$ we have*

- $\phi(\alpha f + \beta g) = \alpha\phi(f) + \beta\phi(g)$,
- $\phi(fg) = \phi(f)\phi(g)$ and $\phi(\bar{f}) = \phi(f^*)$,
- $\|\phi(f)\| = \|f\| := \sup_{t \in \text{Sp}(A)} |f(t)|$,
- $\phi(f_0) = 1_{\mathcal{H}}$ and $\phi(f_1) = A$ where $f_0(t) = 1$ and $f_1(t) = t$, for $t \in \text{Sp}(A)$.

With this notation we define

$$f(A) = \phi(f), \quad \forall f \in C(\text{Sp}(A)),$$

1.2. Concepts of operator theory and operator inequality

and we call it the continuous functional calculus for a self-adjoint operator A , [9].

Self-adjoint elements in $\mathbb{B}(\mathcal{H})$ form a real subspace which denoted by $\mathbb{B}_{sa}(\mathcal{H})$ and can be considered by a partial order which for $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ we say that $A \leq B$ if and only if $B - A \geq 0$. Also, $A < B$ if and only if $B - A > 0$. Especially, if m, M be two real numbers, then $m \leq A \leq M$ if and only if $m \leq \langle Ax, x \rangle \leq M$ for every unit vector $x \in \mathcal{H}$. This partial order is called operator order and the inequality by this order is called as operator inequality. In fact, in operator inequality both sides of the inequality deal with self-adjoint operators, hence continuous functional calculus is a very useful tool for achieving these kind of inequalities. However, it should be noted that many numerical inequalities are not extendable to operators or their operator version are different from numerical form. For instance, triangular inequality for two matrices A and B as $|A + B| \leq |A| + |B|$ is not always true.

If A is a self-adjoint operator and f is a real-valued continuous function on $Sp(A)$ then $f(t) \geq 0$ for any $t \in Sp(A)$ implies that $f(A) \geq 0$, i.e., $f(A)$ is a positive operator. Moreover, if $g(t)$ be a real-valued continuous function on $Sp(A)$ such that $f(t) \leq g(t)$ for any $t \in Sp(A)$ implies that $f(A) \leq g(A)$.

Definition 1.2.11 [19] A real-valued continuous function on an interval \mathcal{I} is said to be operator convex if

$$f(\lambda A + (1 - \lambda)B) \leq \lambda f(A) + (1 - \lambda)f(B),$$

for $\lambda \in [0, 1]$ and self-adjoint operators A and B on the Hilbert space \mathcal{H} with spectrum in \mathcal{I} . Notice that a function f is operator concave if $-f$ is operator convex.

Definition 1.2.12 [19] A real-valued continuous function on an interval \mathcal{I} is said to be operator monotone with respect to the operator order, i.e., $A \leq B$ with $Sp(A), Sp(B) \subset \mathcal{I}$ imply $f(A) \leq f(B)$.

We notice that $f : [0, \infty) \rightarrow [0, \infty)$ is operator monotone if and only if it is operator concave, see [11, Theorem 2.5.2].

Example 1.2.13 [9] The function $f(t) = t^r$ is operator monotone on $(0, \infty)$ if and only if $0 \leq r \leq 1$. In addition, the function $f(t) = t^r$ is operator convex on $(0, \infty)$ if and only if $r \in [-1, 0] \cup [1, 2]$. The logarithmic function $f(t) = \ln t$ is operator monotone on $(0, \infty)$. The exponential function $f(t) = e^t$ is neither operator convex nor operator monotone.

Remark 1.2.14 It should be noted that every operator convex function is a convex function in the sense of the real-valued function, but it does not hold vice versa. Moreover, every operator monotone function is an increasing function in the sense of real-valued function, but it is not valid conversely.

Example 1.2.15 The function $f(t) = t^3$ on $[0, \infty)$ is convex, but it is not operator convex. Be-

cause, by considering

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

then

$$\frac{A^3 + B^3}{2} - \left(\frac{A+B}{2}\right)^3 = \frac{1}{4} \begin{pmatrix} 11 & 9 \\ 9 & 7 \end{pmatrix} \not\geq 0.$$

Whatsmore, by utilizing these two matrices A and B we can conclude that the increasing function $f(t) = t^2$ on $[0, \infty)$ is not operator monotone. Because, $A \geq B \geq 0$ but $A^2 \not\geq B^2$. Namely

$$A^2 - B^2 = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix} \not\geq 0.$$

For the basic results on operator convex functions (operator concave) and operator monotone, we can refer the reader to [9].

As we know, in the operator inequalities both sides of the inequality should be self-adjoint. Therefore, we are not allowed to multiply both sides in a self-adjoint operator, because the self-adjoint property may not be valid after multiplication. The following lemma shows how the multiplication of an operator in the both sides can be done.

Lemma 1.2.16 [15] Let $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ and $A \leq B$. Then for $C \in \mathbb{B}(\mathcal{H})$ we have

$$C^* A C \leq C^* B C.$$

1.3 Operator norms

One of the remarkable topics in the field of inequalities, are norm inequalities that deal with operator norms. In this section, we introduce some operator norms.

As we know, in spaces with a finite dimension, all norms are equivalent. In the other words, for every pair of norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$, there are some constants like c_1 and c_2 such that for every x

$$c_1 \|x\|_\alpha \leq \|x\|_\beta \leq c_2 \|x\|_\alpha.$$

Therefore in a finite dimension vector space all norms produce the same topology, hence upon their properties they will be selected

The norm $\|\cdot\|$ on $M_n(\mathbb{C})$ is said to be unitarily invariant if $\|UAV\| = \|A\|$ for every A and unitaries U, V . The singular value decomposition theorem states that for every $A \in M_n(\mathbb{C})$ there are unitary matrices U and V such that $A = U(s_1(A), s_2(A), \dots, s_n(A))V$. As a result, unitarily invariant norms are functions of singular values. In [13, 3], Von Neumann proved that these are symmetric gauge functions, i.e., these norms correspond to the function Φ on \mathbb{R}^n , which have the following properties.

- $\Phi(Px) = \Phi(x)$ for all permutable matrices P and $x \in \mathbb{R}^n$.

- If $\epsilon_j = \pm 1$, then $\Phi(\epsilon_1 x_1, \epsilon_2 x_2, \dots, \epsilon_n x_n) = \Phi(x_1, x_2, \dots, x_n)$.

There is a one-to-one correspondence between unitarily invariant norm $||| \cdot |||_\Phi$ and symmetric gauge function Φ which are related together as following:

$$|||A|||_\Phi = \Phi(s_1(A), s_2(A), \dots, s_n(A)),$$

for every $A \in M_n(\mathbb{C})$.

In fact, this norm is closed in the topology generated by $||| \cdot |||_\Phi$ and is called the norm ideal dependent on $||| \cdot |||_\Phi$.

For $1 \leq p \leq \infty$, Schatten p -norm of the matrix A is defined as following:

$$\|A\|_p = (\text{Tr}|A|^p)^{\frac{1}{p}} = \left\{ \sum_{j=1}^n (s_j(A))^p \right\}^{\frac{1}{p}},$$

where Tr is the trace functional. The Schatten p -norms are important examples of unitarily invariant norms, where $p = 1$, $p = 2$, and $p = \infty$ correspond to the trace norm, the Hilbert-Schmidt norm, and the spectral (usual operator) norm, respectively. Namely, based on Schatten p -norms, operator norm defines as follows:

$$\|A\| = \|A\|_\infty \equiv s_1(A) = \lim_{p \rightarrow \infty} \|A\|_p.$$

Another kind of unitarily invariant norms, are Fan k -norms which are defined as

$$\|A\|_{(k)} = \sum_{j=1}^k s_j(A), \quad 1 \leq k \leq n,$$

and by this notation, operator norm is given by:

$$\|A\| = \|A\|_\infty = \|A\|_{(1)}.$$

Lemma 1.3.1 (Fan dominance principle)[3] Let A and B be two matrices in $M_n(\mathbb{C})$. If

$$\|A\|_{(k)} \leq \|B\|_{(k)}$$

for $k = 1, 2, \dots, n$, then

$$|||A||| \leq |||B|||,$$

holds for all unitarily invariant norms.

For every three matrices A , B , and C we have

$$|||ABC||| \leq |||A||| |||B||| |||C|||.$$

Introductory concepts

Now, we discuss about another norm on $\mathbb{B}(\mathcal{H})$ which is equivalent to the operator norm. The numerical range of an operator $A \in \mathbb{B}(\mathcal{H})$ denoted by $W(A)$ is the subset of the complex numbers \mathbb{C} given by:

$$W(A) = \{\langle Ax, x \rangle : x \in \mathcal{H}, \|x\| = 1\}.$$

Based on the definition of $W(A)$ the following properties are immediate.

- $W(\alpha I + \beta A) = \alpha + \beta W(A)$, for every $\alpha, \beta \in \mathbb{C}$.
- $W(UAU^*) = W(A)$, for every unitary operator U .
- $W(A^*) = (W(A))^*$.

The following interesting properties of numerical range are proved in [10].

Proposition 1.3.2 Suppose $A \in \mathbb{B}(\mathcal{H})$. Then

- $W(A)$ is a convex subset of \mathbb{C} .
- The spectrum of A is contained in the closure of its numerical range, i.e., $\sigma(A) \subseteq \overline{W(A)}$.

The numerical radius $w(A)$ of an operator A in $\mathbb{B}(\mathcal{H})$ is defined by

$$w(A) = \sup\{|\lambda|, \lambda \in W(A)\} = \sup\{|\langle Ax, x \rangle| : x \in \mathcal{H}, \|x\| = 1\}. \quad (1.8)$$

It is well-known that $w(\cdot)$ is a norm on the $\mathbb{B}(\mathcal{H})$:

- For every $A \in \mathbb{B}(\mathcal{H})$ we have $w(A) \geq 0$, and $w(A) = 0$ if and only if $A = 0$,
- $w(\lambda A) = |\lambda|w(A)$ for every $\lambda \in \mathbb{C}$ and $A \in \mathbb{B}(\mathcal{H})$,
- $w(A + B) \leq w(A) + w(B)$ for every A and B in $\mathbb{B}(\mathcal{H})$.

This norm is equivalent to the operator norm. In fact, we have the following exact result [10].

Theorem 1.3.3 For every $A \in \mathbb{B}(\mathcal{H})$ we have

$$w(A) \leq \|A\| \leq 2w(A). \quad (1.9)$$

Both inequalities in (1.9) are sharp. The first inequality becomes an equality if A is a self-adjoint operator. The second inequality becomes an equality if $A^2 = 0$. Also, it is a basic fact that $w(\cdot)$ satisfies the power inequality

$$w(A^n) \leq w(A)^n,$$

for every A and every natural number n . Another basic property of the numerical radius, is that it is weakly unitarily invariant, i.e.,

$$w(UAU^*) = w(A),$$

for every unitary operator U . One of the interesting property of $w(\cdot)$ is that it is not submultiplicative, i.e.,

$$w(AB) \leq w(A)w(B)$$

does not hold in general, even if A and B commute to each other.

Example 1.3.4 Suppose that $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Then $w(A) < 1$ and $w(A^2) = w(A^3) = \frac{1}{2}$, therefore $w(A^3) \not\leq w(A)w(A^2)$.

In the following proposition submultiplicative property of the numerical radius is expressed [10].

Proposition 1.3.5 Let $A, B \in \mathbb{B}(\mathcal{H})$. Then

- $w(AB) \leq 4w(A)w(B)$.
- If $AB = BA$, then $w(AB) \leq 2w(A)w(B)$.
- If $AB = BA$ and A be a normal operator, then $w(AB) \leq w(A)w(B)$.

Bibliography

- [1] M. Alomari, M. Darus, *Hadamard-type inequalities for s-convex functions*. *Int. Math. Forum*, 3 (2008) 1965-1975.
- [2] A. Beck, *Introduction to Nonlinear Optimization-Theory, Algorithms, and Applications with MATLAB*, MOS-SIAM Series on Optimization 19, SIAM (2014).
- [3] R. Bhatia, *Matrix Analysis*, GTM 169, Springer-Verlag, New York, (1997).
- [4] M. Bombardelli, S. Varoanec, *Properties of h-convex functions related to the Hermite-Hadamard-Fejér inequalities*. *Comput. Math. Appl.* 58(9) (2009) 1869-1877.
- [5] S.S. Dragomir, *An inequality improving the first Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products*, *J. Inequal. Pure and Appl. Math.* 3(2) (2002) Article 31, 8 pp.
- [6] S.S. Dragomir, *An inequality improving the second Hermite-Hadamard inequality for convex functions defined on linear spaces and applications for semi-inner products*, *J. Inequal. Pure and Appl. Math.* 3(3) (2002) Art. 35.
- [7] S.S. Dragomir, *Refinements of the Hermite-Hadamard integral inequality for log-convex functions*, *Austral. Math. Soc. Gaz.* 28(3) (2001) 129-134.
- [8] A. El Farissi, *Simple proof and refinement of Hermite-Hadamard inequality*, *J. Math. Ineq.* 4(3) (2010) 365-369.
- [9] T. Furuta, J. Mii Hot, J. Peari, Y. Seo, *Mond-Peari Method in Operator Inequalities. Inequalities for Bounded Selfadjoint Operators on a Hilbert Space*, Element, Zagreb, (2005).
- [10] K.E. Gustafson, D.K.M. Rao, *Numerical Range*, Springer-Verlag, New York, (1997).
- [11] F. Hiai, *Matrix analysis: matrix monotone functions, matrix means, and majorization*, *Interdisciplinary Information Sciences*, 16(2) (2010) 139248.

Bibliography

- [12] J.B. Hiriart-Urruty, C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer, Berlin (2001).
- [13] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, (1991).
- [14] J. L. W. V. Jensen, *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, *Acta Math.* 30 (1906) 175-193.
- [15] G.J. Murphy, *C* -Algebras and operator theory*, Academic Press, (1990).
- [16] J. E. Pečarić, F. Proschan, Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, Academic Press Inc., San Diego, (1992).
- [17] R.T. Rockafeller, *Convex Analysis*, Princeton University Press, (1970).
- [18] W.T. Sulaiman, *Refinements to Hadamard's inequality for log-convex functions*, *Applied Mathematics*, 2 (2011) 899-903.
- [19] X. Zhan, *Matrix Inequalities*, Springer-Verlag, Berlin, (2002).
- [20] K. Zhu, *An introduction to operator algebras*, CRC Press, (1993).

Hermite-Hadamard inequality for operator geometrically convex functions

In this chapter, first we discuss Hermite-Hadamard inequality for geometrically convex functions and summarize the obtained results. Then we introduce operator geometrically convex functions and present Hermite-Hadamard type inequality for this kind of functions. At the end, as an application of the result which we proved, we refine a norm inequality.

2.1 Inequalities for geometrically convex functions

Definition 2.1.1 [11] A continuous function $f : \mathcal{I} \subseteq \mathbb{R}^+ = (0, \infty) \rightarrow \mathbb{R}^+$ is said to be geometrically convex if

$$f(a^\lambda b^{1-\lambda}) \leq f(a)^\lambda f(b)^{1-\lambda},$$

for $a, b \in \mathcal{I}$ such that $a < b$ and $\lambda \in [0, 1]$.

Every polynomial $P(x)$ with non-negative coefficients is a geometrically convex function on $(0, \infty)$. More generally, every real analytic function $f(x) = \sum_{n=0}^{\infty} c_n x^n$ with non-negative coefficients is geometrically convex function on $(0, R)$ where R denotes the radius of convergence [11].

The author of [5] proved the following Hermite-Hadamard type inequality for operator geometrically convex functions.

Theorem 2.1.2 Let $f : \mathcal{I} \subseteq \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a geometrically convex function and $a, b \in \mathcal{I}$ which

$a < b$. If $f \in L^1[a, b]$, then

$$\begin{aligned} f(\sqrt{ab}) &\leq \frac{1}{\ln b - \ln a} \int_a^b \frac{1}{t} \sqrt{f(t)f\left(\frac{ab}{t}\right)} dt \\ &\leq \frac{1}{\ln b - \ln a} \int_a^b \frac{f(t)}{t} dt \\ &\leq \frac{f(b) - f(a)}{\ln f(b) - \ln f(a)} \\ &\leq \frac{f(a) + f(b)}{2}. \end{aligned}$$

By changing the variable $t = a^\lambda b^{1-\lambda}$ in the above theorem, we have

$$\frac{1}{\ln b - \ln a} \int_a^b \frac{f(t)}{t} dt = \int_0^1 f(a^\lambda b^{1-\lambda}) d\lambda.$$

Remark 2.1.3 For positive scalars a and b , we have

$$\min\{a, b\} \leq G(a, b) = \sqrt{ab} \leq L(a, b) = \frac{b-a}{\ln b - \ln a} \leq A(a, b) = \frac{a+b}{2} \leq \max\{a, b\}.$$

The following lemma shows the relation between convex and geometrically convex functions.

Lemma 2.1.4 [11] Let \mathcal{I} be an interval in \mathbb{R}^+ and $f : \mathcal{I} \rightarrow (0, \infty)$ be a geometrically convex function. Then

$$F = \log \circ f \circ \exp : \log(\mathcal{I}) \rightarrow \mathbb{R}$$

is convex function. Conversely, if \mathcal{J} be a subinterval in \mathbb{R}^+ and $F : \mathcal{J} \rightarrow \mathbb{R}$ be a convex function. Then

$$f = \exp \circ F \circ \log : \exp(\mathcal{J}) \rightarrow \mathbb{R}^+$$

is a geometrically convex function.

On the other hand, in [12] the authors by considering Hermite-Hadamard inequality for convex function $f : [a, b] \rightarrow \mathbb{R}$ separately on each intervals $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$, and then by summation resulted inequalities, they presented a refinement of Hermite-Hadamard inequality for

convex function $f : [a, b] \rightarrow \mathbb{R}$ as following

$$\begin{aligned}
 f\left(\frac{a+b}{2}\right) &\leq \frac{1}{2}\left(f\left(\frac{3a+b}{4}\right) + f\left(\frac{a+3b}{4}\right)\right) \\
 &\leq \frac{1}{b-a} \int_a^b f(x) dx \\
 &\leq \frac{1}{2}\left(f\left(\frac{a+b}{2}\right) + \frac{f(a)+f(b)}{2}\right) \\
 &\leq \frac{f(a)+f(b)}{2}.
 \end{aligned} \tag{2.1}$$

Also, in [10] based on Lemma 2.1.4 and inequality (2.1) and by considering $\mathcal{I} = [a, b]$ with $\log \mathcal{I} = [\log a, \log b]$, the following refinement of Hermite-Hadamard inequality for geometrically convex function is proved.

Theorem 2.1.5 Let f be a geometrically convex function on $[a, b]$ such that $0 < a < b$. Then

$$\begin{aligned}
 f(\sqrt{ab}) &\leq \sqrt{\left(f(a^{\frac{3}{4}}b^{\frac{1}{4}})f(a^{\frac{1}{4}}b^{\frac{3}{4}})\right)} \\
 &\leq \exp\left(\frac{1}{\log b - \log a} \int_a^b \frac{\log f(t)}{t} dt\right) \\
 &\leq \sqrt{f(\sqrt{ab})} \cdot \sqrt[4]{f(a)} \cdot \sqrt[4]{f(b)} \\
 &\leq \sqrt{f(a)f(b)}.
 \end{aligned} \tag{2.2}$$

2.2 Operator geometrically convex functions

In [4] Dragomir investigated operator version of Hermite- Hadamard inequality for operator convex functions and proved the following result.

Theorem 2.2.1 Let $f : \mathcal{I} \rightarrow \mathbb{R}$ be an operator convex function defined on \mathcal{I} . Then for any self-adjoint operators A and B with spectra in \mathcal{I} , the following inequalities hold:

$$\begin{aligned}
 f\left(\frac{A+B}{2}\right) &\leq \frac{1}{2}\left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right)\right] \\
 &\leq \int_0^1 f((1-t)A + tB) dt \\
 &\leq \frac{1}{2}\left[f\left(\frac{A+B}{2}\right) + \frac{f(A)+f(B)}{2}\right] \\
 &\leq \frac{f(A)+f(B)}{2}.
 \end{aligned} \tag{2.3}$$

Hermite-Hadamard inequality for operator geometrically convex functions

Proof Since the function f is continuous, the operator valued integral $\int_0^1 f((1-t)A + tB)dt$ exists for any selfadjoint operators A and B with spectra in \mathcal{I} .

According to the definition of an operator convex function, we have

$$\begin{aligned} f\left(\frac{C+D}{2}\right) &\leq \frac{1}{2} [f((1-t)C + tD) + f((1-t)D + tC)] \\ &\leq \frac{1}{2} [f(C) + f(D)], \end{aligned} \quad (2.4)$$

for $t \in [0, 1]$ and selfadjoint operators A and B with spectra in \mathcal{I} .

Integrating the inequality (2.4) over $[0, 1]$ and taking into account that

$$\int_0^1 f((1-t)C + tD)dt = \int_0^1 f((1-t)D + tC)dt,$$

then we deduce the Hermite-Hadamard inequality for operator convex functions

$$f\left(\frac{C+D}{2}\right) \leq \int_0^1 f((1-t)C + tD)dt \leq \frac{f(C) + f(D)}{2}, \quad (2.5)$$

that holds for any $C, D \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} .

Now, on making use of the change variable $u = 2t$ we have

$$\int_0^{\frac{1}{2}} f((1-t)A + tB)dt = \frac{1}{2} \int_0^1 f\left((1-u)A + u\frac{A+B}{2}\right)du,$$

and by the change of variable $u = 2t - 1$ we get

$$\int_{\frac{1}{2}}^1 f((1-t)A + tB)dt = \frac{1}{2} \int_0^1 f\left((1-u)\frac{A+B}{2} + uB\right)du.$$

Hence, by utilising inequality (2.5) we obtain

$$f\left(\frac{3A+B}{4}\right) \leq \int_0^1 f\left((1-u)A + u\frac{A+B}{2}\right)du \leq \frac{1}{2} \left[f(A) + f\left(\frac{A+B}{2}\right) \right],$$

and

$$f\left(\frac{A+3B}{4}\right) \leq \int_0^1 f\left((1-u)\frac{A+B}{2} + uB\right)du \leq \frac{1}{2} \left[f(B) + f\left(\frac{A+B}{2}\right) \right],$$

which by summation and division by two we get the desired result. ■

Dragomir in [4], presented another proof of above theorem. He proved that

$$\varphi(t) = \langle f(tA + (1-t)B)x, x \rangle$$

for all $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} and $x \in \mathcal{H}$ which $\|x\| = 1$, is a convex function on $[0, 1]$. Then by utilising Hermite-Hadamard inequality for real-valued convex function φ on

2.2. Operator geometrically convex functions

each subintervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, produced the desired result.

As we mentioned in chapter one, the function $f(t) = t^r$ is operator convex on $(0, \infty)$ if either $-1 \leq r \leq 0$ or $1 \leq r \leq 2$. Dragomir in [4] by replacing this function as a sample of operator convex function in inequality (2.3), obtained the following inequalities:

$$\begin{aligned}
 \left(\frac{A+B}{2}\right)^r &\leq \frac{1}{2} \left[\left(\frac{3A+B}{4}\right)^r + \left(\frac{A+3B}{4}\right)^r \right] \\
 &\leq \int_0^1 ((1-t)A + tB)^r dt \\
 &\leq \frac{1}{2} \left[\left(\frac{A+B}{2}\right)^r + \frac{A^r + B^r}{2} \right] \\
 &\leq \frac{A^r + B^r}{2}, \quad r \in [-1, 0] \cup [1, 2].
 \end{aligned} \tag{2.6}$$

Now we are going to introduce operator geometrically convex functions.

For any two strictly positive operators A and B in $\mathbb{B}(\mathcal{H})$ and $t \in \mathbb{R}$, $A\sharp_t B$ is defined as following:

$$A\sharp_t B = A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t A^{\frac{1}{2}}, \tag{2.7}$$

which is a strictly positive operator in $\mathbb{B}(\mathcal{H})$. In the case $0 \leq t \leq 1$, the operator $A\sharp_t B$ is called the t -weighted geometric mean of A and B . In particular, for $t = \frac{1}{2}$, the operator $A\sharp B := A\sharp_{\frac{1}{2}} B$ is called the geometric mean of A and B .

In [8] the following properties of $A\sharp_t B$ is represented.

Lemma 2.2.2 [8] If $A, B \in \mathbb{B}(\mathcal{H})^{++}$ and $t \in \mathbb{R}$. Then

- (i) $A\sharp_t B = A^{1-t} B^t$ for $AB = BA$.
- (ii) $(aA)\sharp_t (bB) = a^{1-t} b^t (A\sharp_t B)$ for $a, b > 0$.
- (iii) $A\sharp_t B = B\sharp_{1-t} A$ and $(A\sharp_t B)^{-1} = A^{-1}\sharp_t B^{-1}$.

Proposition 2.2.3 [13] Let $A, B \in \mathbb{B}(\mathcal{H})^{++}$ such that $Sp(A), Sp(B) \subseteq \mathcal{I}$ and $t \in [0, 1]$. Then $Sp(A\sharp_t B) \subseteq \mathcal{I}$.

Proof Let $\mathcal{I} = [m, M]$ for some positive real numbers m and M with $m < M$. Since $Sp(A), Sp(B) \subseteq \mathcal{I}$, we have

$$m1_{\mathcal{H}} \leq A \leq M1_{\mathcal{H}}, \tag{2.8}$$

$$m1_{\mathcal{H}} \leq B \leq M1_{\mathcal{H}}. \tag{2.9}$$

Hermite-Hadamard inequality for operator geometrically convex functions

So according to Theorem 1.2.16 and by multiplying both sides of inequality (2.9) by $A^{-\frac{1}{2}}$, we obtain

$$\begin{aligned} A^{-\frac{1}{2}}mA^{-\frac{1}{2}} &\leq A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq A^{-\frac{1}{2}}MA^{-\frac{1}{2}} \\ \Rightarrow mA^{-1} &\leq A^{-\frac{1}{2}}BA^{-\frac{1}{2}} \leq MA^{-1}. \end{aligned}$$

Now by considering operator monotonicity property of function $f(x) = x^t$ on $(0, \infty)$ for $t \in [0, 1]$ we have

$$m^t A^{-t} \leq (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^t \leq M^t A^{-t}, \quad (2.10)$$

by multiplying inequality (2.10) in $A^{\frac{1}{2}}$ from both sides, inequalities

$$m^t A^{1-t} \leq A\sharp_t B \leq M^t A^{1-t},$$

hold, so according to inequality (2.8) we have

$$m^t m^{1-t} 1_{\mathcal{H}} \leq m^t A^{1-t} \leq A\sharp_t B \leq M^t A^{1-t} \leq M^t M^{1-t} 1_{\mathcal{H}},$$

or

$$m 1_{\mathcal{H}} \leq A\sharp_t B \leq M 1_{\mathcal{H}},$$

which is equivalent to

$$\sigma(A\sharp_t B) \subseteq \mathcal{I} = [m, M].$$

■

Now by considering Proposition (2.2.3), we represent the following definition.

Definition 2.2.4 [13] A continuous function $f : \mathcal{I} \subseteq \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is said to be operator geometrically convex function if

$$f(A\sharp_t B) \leq f(A)\sharp_t f(B)$$

for $A, B \in \mathbb{B}(\mathcal{H})^{++}$ such that $Sp(A), Sp(B) \subseteq \mathcal{I}$ and $t \in [0, 1]$.

Now the question is that, whether we can present Hermite-Hadamard inequality for operator geometrically convex functions or not?

In the following we present some results which we derived in answering the above question.

2.2.1 Inequalities for operator geometrically convex functions

We need the following lemmas and theorem for proving our results.

Lemma 2.2.5 [7, 8] Let $A, B \in \mathbb{B}(\mathcal{H})^{++}$ and $t, s, u \in \mathbb{R}$. Then

$$(A\sharp_t B)\sharp_s(A\sharp_u B) = A\sharp_{(1-s)t+su} B.$$

Lemma 2.2.6 [7] Let A, B, C, D be operators in $\mathbb{B}(\mathcal{H})^{++}$ and $t \in \mathbb{R}$. Then

$$A\sharp_t B \leq C\sharp_t D,$$

for $A \leq C$ and $B \leq D$.

According to Lemma 2.2.6, we can prove Proposition 2.2.3 in a shorter way as following:

Proof Second proof of Proposition 2.2.3: By utilizing inequalities (2.8) and (2.9), since $A \leq M1_{\mathcal{H}}$ and $B \leq M1_{\mathcal{H}}$, according to lemma 2.2.6 we obtain

$$A\sharp_t B \leq (M1_{\mathcal{H}})\sharp_t(M1_{\mathcal{H}}) = M. \quad (2.11)$$

We also have $m1_{\mathcal{H}} \leq A$ and $m1_{\mathcal{H}} \leq B$. By using lemma 2.2.6 again, we have

$$m = (m1_{\mathcal{H}})\sharp_t(m1_{\mathcal{H}}) \leq A\sharp_t B. \quad (2.12)$$

According to the inequalities (2.11) and (2.12), the proof will be completed. ■

The following characterization of operator convex functions holds.

Theorem 2.2.7 [4] (Operator Jensen inequality) Let \mathcal{H} and \mathcal{K} be Hilbert spaces. Let f be a real valued continuous function on an interval \mathcal{I} . Let A and A_i be selfadjoint operators on \mathcal{H} with spectra contained in \mathcal{I} , for each $i = 1, 2, \dots, k$. Then the following conditions are mutually equivalent:

- (i) f is operator convex on \mathcal{I} .
- (ii) $f(C^*AC) \leq C^*f(A)C$ for every selfadjoint operator $A: \mathcal{H} \rightarrow \mathcal{H}$ and isometry $C: \mathcal{K} \rightarrow \mathcal{H}$, i.e., $C^*C = 1_{\mathcal{K}}$.

Lemma 2.2.8 [13] Let $A, B \in \mathbb{B}(\mathcal{H})^{++}$. If $f: \mathcal{I} \subseteq \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a continuous function, then

$$\int_0^1 f(A\sharp_t B)\sharp f(A\sharp_{1-t} B) dt \leq \left(\int_0^1 f(A\sharp_t B) dt \right)\sharp \left(\int_0^1 f(A\sharp_{1-t} B) dt \right) \quad (2.13)$$

which $Sp(A), Sp(B) \subseteq \mathcal{I}$.

Proof Since the function $t^{\frac{1}{2}}$ is operator concave, we can write

$$\begin{aligned}
 & \left(\left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \left(\int_0^1 f(A_{\#u}B) du \right) \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \quad \text{By change of variable } v = 1 - u \\
 & = \left(\left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} \left(\int_0^1 f(A_{\#u}B) du \right) \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & = \left(\int_0^1 \left(\int_0^1 f(A_{\#v}B) dv \right)^{\frac{1}{2}} f(A_{\#u}B) \left(\int_0^1 f(A_{\#v}B) dv \right)^{\frac{1}{2}} du \right)^{\frac{1}{2}} \\
 & = \left(\int_0^1 \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} (f(A_{\#1-u}B))^{\frac{1}{2}} \left((f(A_{\#1-u}B))^{-\frac{1}{2}} f(A_{\#u}B) (f(A_{\#1-u}B))^{-\frac{1}{2}} \right) \right. \\
 & \quad \left. \times (f(A_{\#1-u}B))^{\frac{1}{2}} \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} du \right)^{\frac{1}{2}} \\
 & \quad \text{By the operator Jensen inequality} \\
 & \geq \int_0^1 \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} (f(A_{\#1-u}B))^{\frac{1}{2}} \left((f(A_{\#1-u}B))^{-\frac{1}{2}} f(A_{\#u}B) (f(A_{\#1-u}B))^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \quad \times (f(A_{\#1-u}B))^{\frac{1}{2}} \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} du \\
 & = \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} \int_0^1 (f(A_{\#1-u}B))^{\frac{1}{2}} \left((f(A_{\#1-u}B))^{-\frac{1}{2}} f(A_{\#u}B) (f(A_{\#1-u}B))^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \quad \times (f(A_{\#1-u}B))^{\frac{1}{2}} du \left(\int_0^1 f(A_{\#v}B) dv \right)^{-\frac{1}{2}} \\
 & \quad \text{By change of variable } u = 1 - v \\
 & = \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \int_0^1 (f(A_{\#1-u}B))^{\frac{1}{2}} \left((f(A_{\#1-u}B))^{-\frac{1}{2}} f(A_{\#u}B) (f(A_{\#1-u}B))^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \quad \times (f(A_{\#1-u}B))^{\frac{1}{2}} du \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} .
 \end{aligned}$$

So, we got

$$\begin{aligned}
 & \left(\left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \left(\int_0^1 f(A_{\#u}B) du \right) \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \geq \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} \int_0^1 (f(A_{\#1-u}B))^{\frac{1}{2}} \left((f(A_{\#1-u}B))^{-\frac{1}{2}} f(A_{\#u}B) (f(A_{\#1-u}B))^{-\frac{1}{2}} \right)^{\frac{1}{2}} \\
 & \quad \times (f(A_{\#1-u}B))^{\frac{1}{2}} du \left(\int_0^1 f(A_{\#1-u}B) du \right)^{-\frac{1}{2}} .
 \end{aligned}$$

Multiplying both sides of above inequality by $\left(\int_0^1 f(A\sharp_{1-u}B)du\right)^{\frac{1}{2}}$, we concluded

$$\left(\int_0^1 f(A\sharp_uB)du\right)\sharp\left(\int_0^1 f(A\sharp_{1-u}B)du\right)\geq\int_0^1 f(A\sharp_uB)\sharp f(A\sharp_{1-u}B)du.$$

■

Before giving our theorems in this section, we mention the following remark.

Remark 2.2.9 Let $p(x) = x^t$ and $q(x) = x^s$ on $[1, \infty)$, where $0 \leq t \leq s$. If $f(A) \leq f(B)$, then $Sp(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}) \subseteq [1, \infty)$. By functional calculus, we have

$$p\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)\leq q\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right).$$

So,

$$\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t\leq\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^s.$$

Now, we are ready to prove Hermite- Hadamard type inequality for operator geometrically convex functions.

Theorem 2.2.10 [13] Let f be an operator geometrically convex function. Then, we have

$$f(A\sharp B)\leq\int_0^1 f(A\sharp_tB)dt\leq\int_0^1 f(A)\sharp_t f(B)dt. \tag{2.14}$$

Moreover, if $f(A) \leq f(B)$, then we have

$$\int_0^1 f(A\sharp_tB)dt\leq\int_0^1 f(A)\sharp_t f(B)dt\leq\frac{1}{2}((f(A)\sharp f(B)) + f(B)) \tag{2.15}$$

for $A, B \in \mathbb{B}(\mathcal{H})^{++}$.

Proof Let f be an operator geometrically convex function, then we have

$$\begin{aligned} f(A\sharp B) &= f((A\sharp_tB)\sharp(A\sharp_{1-t}B)) \quad \text{By lemma 2.2.5} \\ &\leq f(A\sharp_tB)\sharp f(A\sharp_{1-t}B) \quad f \text{ is operator geometrically convex} \end{aligned}$$

Taking integral of the both sides of the above inequalities on $[0, 1]$ we obtain

$$\begin{aligned} f(A\sharp B) &\leq \int_0^1 f(A\sharp_t B)\sharp f(A\sharp_{1-t} B) dt \\ &\leq \left(\int_0^1 f(A\sharp_t B) dt\right)\sharp\left(\int_0^1 f(A\sharp_{1-t} B) dt\right) \text{ By Lemma 2.13} \\ &= \int_0^1 f(A\sharp_t B) dt \\ &\leq \int_0^1 f(A)\sharp_t f(B) dt. \end{aligned}$$

For the case $f(A) \leq f(B)$, by applying Remark 2.2.9 for $s = \frac{1}{2}$ we have

$$\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t \leq \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^{\frac{1}{2}}.$$

By integrating the above inequality over $t \in [0, \frac{1}{2}]$, we obtain

$$\int_0^{\frac{1}{2}} \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t dt \leq \frac{1}{2} \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^{\frac{1}{2}}.$$

Multiplying both sides of the above inequality by $f(A)^{\frac{1}{2}}$, we have

$$\begin{aligned} &\int_0^{\frac{1}{2}} f(A)^{\frac{1}{2}} \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t f(A)^{\frac{1}{2}} dt \\ &\leq \frac{1}{2} \left(f(A)^{\frac{1}{2}} \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^{\frac{1}{2}} f(A)^{\frac{1}{2}}\right). \end{aligned}$$

It follows that

$$\int_0^{\frac{1}{2}} f(A)\sharp_t f(B) \leq \frac{f(A)\sharp f(B)}{2}. \tag{2.16}$$

On the other hand, by considering Remark 2.2.9 for $s = 1$ we have

$$\left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t \leq f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}.$$

Integrating the above inequality over $t \in [\frac{1}{2}, 1]$, we get

$$\int_{\frac{1}{2}}^1 \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right)^t dt \leq \frac{1}{2} \left(f(A)^{-\frac{1}{2}}f(B)f(A)^{-\frac{1}{2}}\right).$$

By multiplying both side of above inequality by $f(A)^{\frac{1}{2}}$, we have

$$\int_{\frac{1}{2}}^1 f(A)^{\frac{1}{2}} \left(f(A)^{-\frac{1}{2}} f(B) f(A)^{-\frac{1}{2}} \right)^t f(A)^{\frac{1}{2}} dt \leq \frac{f(B)}{2}.$$

It follows that

$$\int_{\frac{1}{2}}^1 f(A) \sharp_t f(B) \leq \frac{f(B)}{2}. \quad (2.17)$$

From inequalities (2.16) and (2.17) we obtain

$$\begin{aligned} \int_0^{\frac{1}{2}} f(A \sharp_t B) dt + \int_{\frac{1}{2}}^1 f(A \sharp_t B) dt &\leq \int_0^{\frac{1}{2}} f(A) \sharp_t f(B) dt + \int_{\frac{1}{2}}^1 f(A) \sharp_t f(B) dt \\ &\leq \frac{f(A) \sharp f(B)}{2} + \frac{f(B)}{2}. \end{aligned}$$

It follows that

$$\int_0^1 f(A \sharp_t B) dt \leq \int_0^1 f(A) \sharp_t f(B) dt \leq \frac{1}{2} ((f(A) \sharp f(B)) + f(B)).$$

■

By making use of inequalities (2.14) and (2.15), we have the following result.

Corollary 2.2.11 [13] Let f be an operator geometrically convex functions. Then, if $f(A) \leq f(B)$ we have

$$f(A \sharp B) \leq \int_0^1 f(A \sharp_t B) dt \leq \frac{1}{2} ((f(A) \sharp f(B)) + f(B)) \quad (2.18)$$

for $A, B \in \mathbb{B}(\mathcal{H})^{++}$.

Remark 2.2.12 It is interesting to note that we can not replace the last part of either inequalities (2.15) or (2.18) with a smaller bound $f(A) \sharp f(B)$. Because, by considering $f(t) = t$, $A = I$ and $B = eI$ as an example, which e is the Euler constant, we have

$$\int_0^1 (f(A) \sharp_t f(B)) dt = \int_0^1 I^{1-t} \times (eI)^t dt = \int_0^1 e^t dt = e - 1.$$

On the other hand

$$f(A) \sharp f(B) = I \sharp (eI) = I^{\frac{1}{2}} \times (eI)^{\frac{1}{2}} = e^{\frac{1}{2}}.$$

So, it results

$$e - 1 = 1.71 \not\leq e^{\frac{1}{2}} = 1.64.$$

Theorem 2.2.13 [13] *Let f be an operator geometrically convex function. Then, for $A, B \in \mathbb{B}(\mathcal{H})^{++}$ we have*

$$f(A\#B) \leq \int_0^1 f(A\#_t B) \# f(A\#_{1-t} B) dt \leq f(A) \# f(B). \quad (2.19)$$

Proof *In fact*

$$\begin{aligned} f(A\#B) &= f((A\#_t B) \# (A\#_{1-t} B)) \quad \text{By Lemma 2.2.5} \\ &\leq f(A\#_t B) \# f(A\#_{1-t} B) \quad f \text{ is operator geometrically convex function} \\ &\leq (f(A) \#_t f(B)) \# (f(A) \#_{1-t} f(B)) \quad \text{By Lemma 2.2.6} \\ &= f(A) \# f(B). \end{aligned}$$

It follows that

$$f(A\#B) \leq f(A\#_t B) \# f(A\#_{1-t} B) \leq f(A) \# f(B).$$

Integrating the above inequality over $t \in [0, 1]$ we obtain the desired result. \blacksquare

We divide the interval $[0, 1]$ to the interval $[\nu, 1 - \nu]$ when $\nu \in [0, \frac{1}{2})$ and to the interval $[1 - \nu, \nu]$ when $\nu \in (\frac{1}{2}, 1]$. Then we have the following inequalities.

Theorem 2.2.14 [13] *Let $A, B \in \mathbb{B}(\mathcal{H})^{++}$ such that $f(A) \leq f(B)$. Then we have*

1. *For $\nu \in [0, \frac{1}{2})$*

$$f(A) \#_\nu f(B) \leq \frac{1}{1-2\nu} \int_\nu^{1-\nu} f(A) \#_t f(B) dt \leq f(A) \#_{1-\nu} f(B). \quad (2.20)$$

2. *For $\nu \in (\frac{1}{2}, 1]$*

$$f(A) \#_{1-\nu} f(B) \leq \frac{1}{2\nu-1} \int_{1-\nu}^\nu f(A) \#_t f(B) dt \leq f(A) \#_\nu f(B). \quad (2.21)$$

Proof *Let $\nu \in [0, \frac{1}{2})$, then by Remark 2.2.9 we have*

$$\begin{aligned} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^\nu &\leq \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^t \\ &\leq \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^{1-\nu} \end{aligned}$$

2.2. Operator geometrically convex functions

for $\nu \leq t \leq 1 - \nu$, and $A, B \in \mathbb{B}(\mathcal{H})^{++}$ such that $Sp(A), Sp(B) \subseteq \mathcal{I}$. By integrating the above inequality over $t \in [\nu, 1 - \nu]$ we obtain

$$\begin{aligned} \int_{\nu}^{1-\nu} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^{\nu} dt &\leq \int_{\nu}^{1-\nu} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^t dt \\ &\leq \int_{\nu}^{1-\nu} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^{1-\nu} dt. \end{aligned}$$

It follows that

$$\begin{aligned} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^{\nu} &\leq \frac{1}{1-2\nu} \int_{\nu}^{1-\nu} \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^t dt \\ &\leq \left(f(A)^{\frac{-1}{2}} f(B) f(A)^{\frac{-1}{2}} \right)^{1-\nu}. \end{aligned}$$

Multiplying the both sides of the above inequality by $f(A)^{\frac{1}{2}}$ gives us

$$f(A) \sharp_{\nu} f(B) \leq \frac{1}{1-2\nu} \int_{\nu}^{1-\nu} f(A) \sharp_t f(B) dt \leq f(A) \sharp_{1-\nu} f(B).$$

Also, we know that

$$\begin{aligned} \lim_{\nu \rightarrow \frac{1}{2}} f(A) \sharp_{\nu} f(B) &= \lim_{\nu \rightarrow \frac{1}{2}} \frac{1}{1-2\nu} \int_{\nu}^{1-\nu} f(A) \sharp_t f(B) dt \\ &= \lim_{\nu \rightarrow \frac{1}{2}} f(A) \sharp_{1-\nu} f(B) \\ &= f(A) \sharp f(B). \end{aligned}$$

Similarly, for $\nu \in (\frac{1}{2}, 1]$, by a same proof as above we get

$$f(A) \sharp_{1-\nu} f(B) \leq \frac{1}{2\nu-1} \int_{1-\nu}^{\nu} f(A) \sharp_t f(B) dt \leq f(A) \sharp_{\nu} f(B).$$

Hence, the proof will be completed. ■

By definition of geometrically convex function and (2.20) we have

$$\begin{aligned} f(A \sharp_{\nu} B) &\leq \frac{1}{1-2\nu} \int_{\nu}^{1-\nu} f(A \sharp_t B) dt \\ &\leq \frac{1}{1-2\nu} \int_{\nu}^{1-\nu} f(A) \sharp_t f(B) dt \\ &\leq f(A) \sharp_{1-\nu} f(B), \end{aligned}$$

for $v \in [0, \frac{1}{2})$.

We should mention here that by Hopital's rule we have

$$\lim_{v \rightarrow \frac{1}{2}} \frac{1}{1-2v} \int_v^{1-v} f(A\sharp_t B) dt = \lim_{v \rightarrow \frac{1}{2}} f(A\sharp_v B) = f(A\sharp B).$$

On the other hand, by definition of geometrically convex function and (2.21) we have

$$\begin{aligned} f(A\sharp_{1-v} B) &\leq \frac{1}{2v-1} \int_{1-v}^v f(A\sharp_t B) dt \\ &\leq \frac{1}{2v-1} \int_{1-v}^v f(A)\sharp_t f(B) dt \\ &\leq f(A)\sharp_v f(B) \end{aligned}$$

for $v \in (\frac{1}{2}, 1]$.

2.2.2 Examples of operator geometrically convex functions

In this section we represent some examples of operator geometrically convex functions. At the end, we derive a refinement of an norm inequality.

Remark 2.2.15 For operators $A, B \in \mathbb{B}(\mathcal{H})^{++}$, Ando proved in [1] that if Ψ is a positive linear map, then we have

$$\Psi(A\sharp_v B) \leq \Psi(A)\sharp_v \Psi(B), \quad v \in [0, 1].$$

The above inequality shows that we can find some examples of operator geometrically convex functions f when they are linear.

Example 2.2.16 According to Lemma 2.2.2, part (iii), we can easily check that the function $f(t) = t^{-1}$ is operator geometrically convex function on $\mathbb{B}(\mathcal{H})^{++}$.

Theorem 2.2.17 [3] Let A and B be two strictly positive operators. Then the block-matrix $\begin{bmatrix} A & X \\ X^* & B \end{bmatrix}$ is positive if and only if $A \geq XB^{-1}X^*$.

Definition 2.2.18 [1] Let ϕ be a map on C^* -algebra $\mathbb{B}(\mathcal{H})$. We say that ϕ is 2-positive if the 2×2 operator matrix $\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$ then we have $\begin{bmatrix} \phi(A) & \phi(B) \\ \phi(B^*) & \phi(C) \end{bmatrix} \geq 0$.

In [9], M. Lin gave an example of a 2-positive map over contraction operators (i.e., $\|A\| < 1$). He proved that

$$\phi(t) = (1-t)^{-1} \tag{2.22}$$

2.2. Operator geometrically convex functions

is 2-positive. In the following example, we show that this function is operator geometrically convex function.

Example 2.2.19 [13] Let A and B be two contraction operators in $\mathbb{B}(\mathcal{H})^{++}$. Then

$$\|A\sharp B\|^2 \leq \|A\|\|B\| < 1,$$

or in the other words, $A\sharp B$ is a contraction operator. Also, we have $A\sharp B = B\sharp A$ and

$$A = (A\sharp B)B^{-1}(A\sharp B),$$

so, according to Theorem 2.2.17 we can conclude that the positivity of operator matrix

$$\begin{bmatrix} A & A\sharp B \\ A\sharp B & B \end{bmatrix}.$$

Hence, by (2.22) we obtain

$$\begin{bmatrix} (I-A)^{-1} & (I-(A\sharp B))^{-1} \\ (I-(A\sharp B))^{-1} & (I-B)^{-1} \end{bmatrix}$$

is positive.

On the other hand, by Ando's characterization of the geometric mean if X is a Hermitian operator and

$$\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq 0.$$

Then $X \leq A\sharp B$.

So we conclude that $(I-(A\sharp B)^{-1}) \leq (I-A)^{-1}\sharp(I-B)^{-1}$. Therefore, the function $\phi(t) = (1-t)^{-1}$ is operator geometrically convex function.

Also, Lin proved that the function

$$\psi(t) = \frac{1+t}{1-t}$$

is 2-positive over contractions. By the same argument as above example we can say the above function is operator geometrically convex too.

Example 2.2.20 [13] In the proof of [6, Theorem 4.12], by applying Hölder-McCarthy inequality the authors showed the following inequalities

$$\begin{aligned} \langle A\sharp_{\alpha} Bx, x \rangle &= \left\langle \left(A^{\frac{-1}{2}} B A^{\frac{-1}{2}} \right)^{\alpha} A^{\frac{1}{2}} x, A^{\frac{1}{2}} x \right\rangle \\ &\leq \left\langle \left(A^{\frac{-1}{2}} B A^{\frac{-1}{2}} \right) A^{\frac{1}{2}} x, A^{\frac{1}{2}} x \right\rangle^{\alpha} \left\langle A^{\frac{1}{2}} x, A^{\frac{1}{2}} x \right\rangle^{1-\alpha} \\ &= \langle Ax, x \rangle^{1-\alpha} \langle Bx, x \rangle^{\alpha} \\ &= \langle Ax, x \rangle \sharp_{\alpha} \langle Bx, x \rangle \end{aligned}$$

Hermite-Hadamard inequality for operator geometrically convex functions

for $x \in \mathcal{H}$ and $\alpha \in [0, 1]$. By taking the supremum over unit vector x , we obtain that $f(x) = \|x\|$ is geometrically convex function for usual operator norms.

By the above example and Corollary 2.2.11, when $\|A\| \leq \|B\|$ we have

$$\|A\sharp B\| \leq \int_0^1 \|A\sharp_t B\| dt \leq \frac{1}{2}(\sqrt{\|A\|\|B\|} + \|B\|)$$

for $A, B \in \mathbb{B}(\mathcal{H})^{++}$, which is a norm inequality with the operator norm.

Bibliography

- [1] T. Ando, *Concavity of certain maps on positive definite matrices and applications to Hadamard products*, *Linear Algebra Appl.* 26 (1979) 203-241.
- [2] R. Bhatia, C. Davis, *A Cauchy-Schwarz inequality for operators with applications*, *Linear Algebra Appl.* 223 (1995) 119-129.
- [3] S. S. Dragomir, *Hermite-Hadamard's type inequalities for operator convex functions*, *Applied Mathematics and Computation* 218 (2011) 766-772.
- [4] T. Furuta, J. Mii Hot, J. Peari, Y. Seo, *Mond-Peari Method in Operator Inequalities. Inequalities for Bounded Selfadjoint Operators on a Hilbert Space*, *Element, Zagreb*, (2005).
- [5] I. Iscan, *Some new Hermite-Hadamard type inequalities for geometrically convex functions*, *Mathematics and Statistics*, 1(2) (2013) 86-91.
- [6] C. Jung, H. Lee, Y. Lim, T. Yamazaki, *Weighted geometric mean of n -operators with n -parameters*, *Linear Algebra Appl.* 432 (2010) 151-1530.
- [7] F. Kubo, T. Ando, *Means of positive linear operators*, *Math. Ann.* 246 (3) (1980) 205-224.
- [8] J. Lawson, Y. Lim, *Weighted means and Karcher equations of positive operators*, *Proc. Natk. Acad. Sci. USA*, 110 (39) (2013) 15626-15632.
- [9] M. Lin, *A Lewent type determinantal inequality*, *Taiwanese J. Math.* 17 (2013) 1303-1309.
- [10] M. Lin, *Remarks on two recent results of Audenaert*, *Linear Algebra Appl.* 489 (2016) 24-29.
- [11] C.P. Niculescu, *Convexity according to the geometric mean*, *Math Inequal. Appl.* 3(2) (2000) 155-167.
- [12] C.P. Niculescu, L.E. Persson, *Convex functions and their applications, A contemporary approach*, *CMS Books in Mathematics, vol. 23, Springer Verlag, New York*, (2006).

Bibliography

- [13] A. Taghavi, V. Darvish, T. Azimi Roushan. *Hermite-Hadamard Type Inequalities For Operator Geometrically Convex Functions II*. *Kragujevac Journal of Mathematics*, 45(2) (2021) 191-202.
- [14] A. Taghavi, V. Darvish, H. M. Nazari, S. S. Dragomir, *Hermite-Hadamard type inequalities for operator geometrically convex functions*, *Monatsh. Math.* 181 (2016) 187-203.

Refinements for some inequalities related to unitarily invariant norm

Another kind of convex functions, are log-convex functions. This chapter deals with these functions. By utilizing the results which we derive, we give some refinements on interpolation between the arithmetic-geometric mean and Cauchy-Schwarz matrix norm inequalities.

3.1 Some inequalities for arithmetic-geometric convex functions

Definition 3.1.1 A continuous function $f : \mathcal{I} \subset \mathbb{R} \rightarrow \mathbb{R}^+$ is said to be an AG-convex function (arithmetic-geometrically or log convex function) if

$$f(\lambda a + (1 - \lambda)b) \leq f(a)^\lambda f(b)^{1-\lambda}, \quad (3.1)$$

for $a, b \in \mathcal{I}$ and $\lambda \in [0, 1]$, i.e., $\log f$ is convex.

As an example, $f(t) = e^t$ is an arithmetic-geometric convex function.

According to Young inequality, we can write inequality (3.1) as following

$$f(\lambda a + (1 - \lambda)b) \leq f(a)^\lambda f(b)^{1-\lambda} \leq \lambda f(a) + (1 - \lambda)f(b),$$

it means that every arithmetic-geometric convex function is a convex function.

Remark 3.1.2 [7] The condition (3.1) can be written as following

$$f \circ \log(\exp(\lambda a) \exp((1 - \lambda)b)) \leq (f \circ \log(\exp(a)))^\lambda (f \circ \log(\exp(b)))^{1-\lambda}, \quad (3.2)$$

then we observe that $f : \mathcal{I} \rightarrow \mathbb{R}^+$ is AG-convex on \mathcal{I} if and only if $f \circ \log$ is G-convex (geometrically or multiplicative convex) on $\exp(\mathcal{I}) := \{\exp(x), x \in \mathcal{I}\}$.

If $\mathcal{I} = [a, b]$, then $\exp(\mathcal{I}) = [\exp(a), \exp(b)]$.

The authors of [10] represented Hermite-Hadamard type inequality for arithmetic-geometric convex functions. To complete the discussion we bring the proof here in a short way.

Theorem 3.1.3 [10] Let f be an AG-convex function defined on $[a, b]$. Then, we have

$$\begin{aligned}
 f\left(\frac{a+b}{2}\right) &\leq \sqrt{f\left(\frac{3a+b}{4}\right)f\left(\frac{a+3b}{4}\right)} \\
 &\leq \exp\left(\frac{1}{b-a}\int_a^b \log \circ f(u) du\right) \\
 &\leq \sqrt{f\left(\frac{a+b}{2}\right)} \cdot \sqrt[4]{f(a)} \cdot \sqrt[4]{f(b)} \\
 &\leq \sqrt{f(a)f(b)},
 \end{aligned} \tag{3.3}$$

Proof Let $f : [a, b] \rightarrow \mathbb{R}^+$ be an AG-convex function. Then by Remark 3.1.2, $f \circ \log$ is G-convex on $[\exp(a), \exp(b)]$. By applying (2.2), we obtain

$$\begin{aligned}
 f \circ \log(\sqrt{\exp(a)\exp(b)}) &\leq \sqrt{f \circ \log\left(\exp(a)^{\frac{3}{4}}\exp(b)^{\frac{1}{4}}\right) f \circ \log\left(\exp(a)^{\frac{1}{4}}\exp(b)^{\frac{3}{4}}\right)} \\
 &\leq \exp\left(\frac{1}{\log \exp(b) - \log \exp(a)} \int_{\exp(a)}^{\exp(b)} \frac{\log \circ f \circ \log(t)}{t} dt\right) \\
 &\leq \sqrt{f \circ \log \sqrt{\exp(a)\exp(b)}} \cdot \sqrt[4]{f \circ \log \exp(a)} \cdot \sqrt[4]{f \circ \log \exp(b)} \\
 &\leq \sqrt{f \circ \log \exp(a) f \circ \log \exp(b)}.
 \end{aligned}$$

Hence, we can write

$$\begin{aligned}
 f \circ \log\left(\exp\left(\frac{a+b}{2}\right)\right) &\leq \sqrt{\left(f \circ \log \exp\left(\frac{3a+b}{4}\right)\right) \left(f \circ \log \exp\left(\frac{a+3b}{4}\right)\right)} \\
 &\leq \exp\left(\frac{1}{\log \exp(b) - \log \exp(a)} \int_{\exp(a)}^{\exp(b)} \frac{\log \circ f \circ \log(t)}{t} dt\right) \\
 &\leq \sqrt{f \circ \log \exp\left(\frac{a+b}{2}\right)} \cdot \sqrt[4]{f \circ \log \exp(a)} \cdot \sqrt[4]{f \circ \log \exp(b)} \\
 &\leq \sqrt{f \circ \log \exp(a) f \circ \log \exp(b)}.
 \end{aligned}$$

Then, we have the desired result. ■

3.2 Inequalities related to unitarily invariant norms of operators

For $a, b \in \mathbb{C}$ arithmetic-geometric mean inequality is as following.

$$|ab| \leq \frac{|a|^2 + |b|^2}{2}. \quad (3.4)$$

The norm version of above inequality is

$$\| \|AB^*\| \| \leq \frac{1}{2} \| \|A^*A + B^*B\| \| \quad (3.5)$$

where $A, B \in M_n(\mathbb{C})$.

A generalization of the above inequality has been given in [4] that asserts that for $A, B, X \in M_n(\mathbb{C})$ we have

$$\| \|AXB^*\| \| \leq \frac{1}{2} \| \|A^*AX + XB^*B\| \|. \quad (3.6)$$

On the other hand, the Cauchy-Schwarz inequality for $a_i, b_i \in \mathbb{R}^+$ is

$$\sum_{i=1}^n a_i b_i \leq \left(\sum_{i=1}^n a_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n b_i^2 \right)^{\frac{1}{2}}. \quad (3.7)$$

The norm version of the Cauchy-Schwarz inequality for $A, B \in M_n(\mathbb{C})$ is as follows.

$$\| \|AB^*\| \|^2 \leq \| \|A^*A\| \| \| \|B^*B\| \|. \quad (3.8)$$

A generalization of the above inequality asserts that if $A, B, X \in M_n(\mathbb{C})$, we have

$$\| \|AXB^*\| \|^2 \leq \| \|A^*AX\| \| \| \|XB^*B\| \|. \quad (3.9)$$

We refer the readers to [3] for more results about the Cauchy-Schwarz inequality.

Recently, it has been proved in [2] that for all $A, B \in M_n(\mathbb{C})$ and all $\nu \in [0, 1]$, we have

$$\| \|AB^*\| \|^2 \leq \| \|\nu A^*A + (1-\nu)B^*B\| \| \| \|(1-\nu)A^*A + \nu B^*B\| \|. \quad (3.10)$$

The above inequality interpolates between arithmetic-geometric mean inequality (3.5) and Cauchy-Schwarz inequality (3.8) by letting $\nu = \frac{1}{2}$ and $\nu = 1$, respectively.

In [6], the author has proved inequality (3.10) in a different way.

A generalization of inequality (3.10) has been derived in [11] which asserts that if $A, B, X \in M_n(\mathbb{C})$ we have

$$\| \|AXB^*\| \|^2 \leq \| \|\nu A^*AX + (1-\nu)XB^*B\| \| \| \|(1-\nu)A^*AX + \nu XB^*B\| \|. \quad (3.11)$$

The above inequality interpolates between inequalities (3.6) and (3.9).

In the following, we give examples of arithmetic-geometric convex functions and then by using

Hermite-Hadamard inequalities for them, we derive an interpolation between the arithmetic-geometric mean and Cauchy-Schwarz matrix norm inequalities.

The author of [5] proved that if $A, B, X \in \mathbb{B}(\mathcal{H})$ such that A and B are positive operators, then for $0 \leq \nu \leq 1$ we have

$$\| \|A^\nu X B^{1-\nu}\| \| \|AX\| \| \|XB\| \|^{1-\nu}. \quad (3.12)$$

In the following we show the log-convexity of a function which the idea of proof comes from [8, Lemma 2.1].

For $T \in \mathbb{B}(\mathcal{H})$, when we consider $\| \| \cdot \| \|$, we are implicitly assuming that the operator T belongs to the norm ideal associated with $\| \| \cdot \| \|$.

Lemma 3.2.1 [9] *Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that A and B are positive operators. Then*

$$f(t) = \| \|A^t X B^{1-t}\| \|$$

is AG-convex on $[0, 1]$.

Proof *We should prove that $f(\alpha t + (1-\alpha)s) \leq f(t)^\alpha f(s)^{1-\alpha}$, for $\alpha \in (0, 1)$ and $s, t \in [0, 1]$.*

$$\begin{aligned} f(\alpha t + (1-\alpha)s) &= \| \|A^{\alpha t + (1-\alpha)s} X B^{1-(\alpha t + (1-\alpha)s)}\| \| \\ &= \| \|A^{\alpha(t-s)} A^s X B^{1-t} B^{(1-\alpha)(t-s)}\| \| \\ &\leq \| \|A^{t-s} A^s X B^{1-t}\| \|^\alpha \| \|A^s X B^{1-t} B^{t-s}\| \|^{1-\alpha} \quad \text{by (3.12)} \\ &= \| \|A^t X B^{1-t}\| \|^\alpha \| \|A^s X B^{1-s}\| \|^{1-\alpha} \\ &= f(t)^\alpha f(s)^{1-\alpha}. \end{aligned}$$

So, $f(t) = \| \|A^t X B^{1-t}\| \|$ is an AG-convex function. ■

Applying inequalities (3.3) to the function $f(t) = \| \|(A^ A)^t X (B^* B)^{1-t}\| \|$ on the interval $[\nu, 1-\nu]$ when $\nu \in [0, \frac{1}{2})$ and on the interval $[1-\nu, \nu]$ when $\nu \in (\frac{1}{2}, 1]$, we obtain the following theorem.*

Theorem 3.2.2 [9] *Let $A, B, X \in \mathbb{B}(\mathcal{H})$ and $\nu \in [0, 1]$. Then*

$$\begin{aligned} \| \|(A^* A)^{\frac{1}{2}} X (B^* B)^{\frac{1}{2}}\| \| &\leq \| \|(A^* A)^{\frac{1+2\nu}{4}} X (B^* B)^{\frac{3-2\nu}{4}}\| \|^\frac{1}{2} \| \|(A^* A)^{\frac{3-2\nu}{4}} X (B^* B)^{\frac{1+2\nu}{4}}\| \|^\frac{1}{2} \\ &\leq \exp\left(\frac{1}{1-2\nu} \int_\nu^{1-\nu} \log \| \|(A^* A)^u X (B^* B)^{1-u}\| \| du\right) \\ &\leq \| \|(A^* A)^{\frac{1}{2}} X (B^* B)^{\frac{1}{2}}\| \|^\frac{1}{2} \| \|(A^* A)^\nu X (B^* B)^{1-\nu}\| \|^\frac{1}{4} \| \|(A^* A)^{1-\nu} X (B^* B)^\nu\| \|^\frac{1}{4} \\ &\leq \| \|(A^* A)^\nu X (B^* B)^{1-\nu}\| \|^\frac{1}{2} \| \|(A^* A)^{1-\nu} X (B^* B)^\nu\| \|^\frac{1}{2}. \end{aligned}$$

3.2. Inequalities related to unitarily invariant norms of operators

Proof Let $\nu \in [0, \frac{1}{2})$, then by inequalities (3.3) for operator AG-convex function f , we have

$$\begin{aligned} f\left(\frac{\nu+1-\nu}{2}\right) &\leq \sqrt{f\left(\frac{3\nu+1-\nu}{4}\right)f\left(\frac{\nu+3(1-\nu)}{4}\right)} \\ &\leq \exp\left(\frac{1}{1-2\nu} \int_{\nu}^{1-\nu} \log \circ f(u) du\right) \\ &\leq \sqrt{f\left(\frac{\nu+1-\nu}{2}\right)} \cdot \sqrt[4]{f(\nu)} \cdot \sqrt[4]{f(1-\nu)} \\ &\leq \sqrt{f(\nu)f(1-\nu)}. \end{aligned}$$

Now by considering $f(t) = |||(A^*A)^t X(B^*B)^{1-t}|||$, since it is AG-convex according to Lemma 3.2.1, by replacing it in the above inequalities we conclude:

$$\begin{aligned} |||(A^*A)^{\frac{1}{2}} X(B^*B)^{\frac{1}{2}}||| &\leq |||(A^*A)^{\frac{1+2\nu}{4}} X(B^*B)^{\frac{3-2\nu}{4}}|||^{\frac{1}{2}} |||(A^*A)^{\frac{3-2\nu}{4}} X(B^*B)^{\frac{1+2\nu}{4}}|||^{\frac{1}{2}} \\ &\leq \exp\left(\frac{1}{1-2\nu} \int_{\nu}^{1-\nu} \log |||(A^*A)^u X(B^*B)^{1-u}||| du\right) \\ &\leq |||(A^*A)^{\frac{1}{2}} X(B^*B)^{\frac{1}{2}}|||^{\frac{1}{2}} |||(A^*A)^{\nu} X(B^*B)^{1-\nu}|||^{\frac{1}{4}} |||(A^*A)^{1-\nu} X(B^*B)^{\nu}|||^{\frac{1}{4}} \\ &\leq |||(A^*A)^{\nu} X(B^*B)^{1-\nu}|||^{\frac{1}{2}} |||(A^*A)^{1-\nu} X(B^*B)^{\nu}|||^{\frac{1}{2}}. \end{aligned}$$

Similarly, for $\nu \in (\frac{1}{2}, 1]$ we have

$$\begin{aligned} |||(A^*A)^{\frac{1}{2}} X(B^*B)^{\frac{1}{2}}||| &\leq |||(A^*A)^{\frac{1+2\nu}{4}} X(B^*B)^{\frac{3-2\nu}{4}}|||^{\frac{1}{2}} |||(A^*A)^{\frac{3-2\nu}{4}} X(B^*B)^{\frac{1+2\nu}{4}}|||^{\frac{1}{2}} \\ &\leq \exp\left(\frac{1}{2\nu-1} \int_{1-\nu}^{\nu} \log |||(A^*A)^u X(B^*B)^{1-u}||| du\right) \\ &\leq |||(A^*A)^{\frac{1}{2}} X(B^*B)^{\frac{1}{2}}|||^{\frac{1}{2}} |||(A^*A)^{\nu} X(B^*B)^{1-\nu}|||^{\frac{1}{4}} |||(A^*A)^{1-\nu} X(B^*B)^{\nu}|||^{\frac{1}{4}} \\ &\leq |||(A^*A)^{\nu} X(B^*B)^{1-\nu}|||^{\frac{1}{2}} |||(A^*A)^{1-\nu} X(B^*B)^{\nu}|||^{\frac{1}{2}}. \end{aligned}$$

Since

$$\lim_{\nu \rightarrow \frac{1}{2}} \exp\left(\frac{1}{1-2\nu} \int_{\nu}^{1-\nu} \log |||(A^*A)^u X(B^*B)^{1-u}||| du\right) = |||(A^*A)^{\frac{1}{2}} X(B^*B)^{\frac{1}{2}}|||.$$

This completes the proof. ■

Theorem 3.2.3 [9] Let A, B, X be operators in $\mathbb{B}(\mathcal{H})$. Then

$$\begin{aligned}
 |||(A^*A)^{\frac{1}{2}}X(B^*B)^{\frac{1}{2}}||| &\leq |||(A^*A)^{\frac{1}{4}}X(B^*B)^{\frac{3}{4}}|||^{\frac{1}{2}} |||(A^*A)^{\frac{3}{4}}X(B^*B)^{\frac{1}{4}}|||^{\frac{1}{2}} \\
 &\leq \exp\left(\int_0^1 \log |||(A^*A)^u X(B^*B)^{1-u}||| du\right) \\
 &\leq |||(A^*A)^{\frac{1}{2}}X(B^*B)^{\frac{1}{2}}|||^{\frac{1}{2}} |||(A^*A)X|||^{\frac{1}{4}} |||X(B^*B)|||^{\frac{1}{4}} \\
 &\leq |||(A^*A)X|||^{\frac{1}{2}} |||X(B^*B)|||^{\frac{1}{2}}.
 \end{aligned} \tag{3.13}$$

Proof By the use of inequalities (3.3) and Lemma 3.2.1 for $a = 0$ and $b = 1$ we get the result. ■

Remark 3.2.4 The authors of [1] proved the following inequality for $A, B, X \in M_n(\mathbb{C})$ and $\nu \in [0, 1]$

$$|||AXB^*|||^2 \leq |||(A^*A)^\nu X(B^*B)^{1-\nu}||| |||(A^*A)^{1-\nu} X(B^*B)^\nu|||, \tag{3.14}$$

for every unitarily invariant norm.

Let $\nu = \frac{1}{2}$, then we have

$$|||AXB^*|||^2 \leq |||(A^*A)^{\frac{1}{2}}X(B^*B)^{\frac{1}{2}}||| |||(A^*A)^{\frac{1}{2}}X(B^*B)^{\frac{1}{2}}|||. \tag{3.15}$$

By considering (3.15), the inequalities of Theorem 3.2.2 is a refinement of inequality (3.14) as the following.

$$\begin{aligned}
 |||AXB^*||| &\leq |||(A^*A)^{\frac{1+2\nu}{4}}X(B^*B)^{\frac{3-2\nu}{4}}|||^{\frac{1}{2}} |||(A^*A)^{\frac{3-2\nu}{4}}X(B^*B)^{\frac{1+2\nu}{4}}|||^{\frac{1}{2}} \\
 &\leq \exp\left(\frac{1}{1-2\nu} \int_\nu^{1-\nu} \log |||(A^*A)^u X(B^*B)^{1-u}||| du\right) \\
 &\leq |||(A^*A)^{\frac{1}{2}}X(B^*B)^{\frac{1}{2}}|||^{\frac{1}{2}} |||(A^*A)^\nu X(B^*B)^{1-\nu}|||^{\frac{1}{4}} |||(A^*A)^{1-\nu} X(B^*B)^\nu|||^{\frac{1}{4}} \\
 &\leq |||(A^*A)^\nu X(B^*B)^{1-\nu}|||^{\frac{1}{2}} |||(A^*A)^{1-\nu} X(B^*B)^\nu|||^{\frac{1}{2}}.
 \end{aligned} \tag{3.16}$$

Remark 3.2.5 We recall inequality (3.10) which is due to K. Audenaert [2] and it states that that for $A, B \in M_n(\mathbb{C})$ and $\nu \in [0, 1]$, we have

$$|||AB^*|||^2 \leq |||\nu A^*A + (1-\nu)B^*B||| |||(1-\nu)A^*A + \nu B^*B|||. \tag{3.17}$$

3.2. Inequalities related to unitarily invariant norms of operators

By knowing that $|||(A^*A)^\nu(B^*B)^{1-\nu}||| \leq |||\nu A^*A + (1-\nu)B^*B|||$ holds for all A, B and $\nu \in [0, 1]$, inequalities (3.16) for $X = I$ gives a refinement for inequality (3.17) as following

$$\begin{aligned}
 |||AB^*||| &\leq |||(A^*A)^{\frac{1+2\nu}{4}}(B^*B)^{\frac{3-2\nu}{4}}|||^{\frac{1}{2}} |||(A^*A)^{\frac{3-2\nu}{4}}(B^*B)^{\frac{1+2\nu}{4}}|||^{\frac{1}{2}} \\
 &\leq \exp\left(\frac{1}{1-2\nu} \int_{\nu}^{1-\nu} \log |||(A^*A)^u(B^*B)^{1-u}||| du\right) \\
 &\leq |||(A^*A)^{\frac{1}{2}}(B^*B)^{\frac{1}{2}}|||^{\frac{1}{2}} |||(A^*A)^\nu(B^*B)^{1-\nu}|||^{\frac{1}{4}} |||(A^*A)^{1-\nu}(B^*B)^\nu|||^{\frac{1}{4}} \\
 &\leq |||(A^*A)^\nu(B^*B)^{1-\nu}|||^{\frac{1}{2}} |||(A^*A)^{1-\nu}(B^*B)^\nu|||^{\frac{1}{2}} \\
 &\leq |||\nu A^*A + (1-\nu)B^*B|||^{\frac{1}{2}} |||(1-\nu)A^*A + \nu B^*B|||^{\frac{1}{2}}.
 \end{aligned}$$

From now on, we try to give a refinement of the following inequality

$$|||A^\nu XB^\nu||| \leq |||X|||^{1-\nu} |||AXB|||^\nu \quad (3.18)$$

for $0 \leq \nu \leq 1$ and $A, B, X \in \mathbb{B}(\mathcal{H})$ such that A and B are positive operators. This inequality has been proved in [5]. Our refinement is a special case when $\nu = \frac{1}{2}$.

Lemma 3.2.6 [9] Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that A and B are positive operators. Then

$$f(s) = |||A^s XB^s|||$$

is AG-convex on $[0, 1]$.

Proof Similar to Lemma 3.2.1, we should prove that

$$f(\alpha t + (1-\alpha)s) \leq f(t)^\alpha f(s)^{1-\alpha}$$

for $\alpha \in (0, 1)$ and $s, t \in [0, 1]$.

$$\begin{aligned}
 f(\alpha t + (1-\alpha)s) &= |||A^{\alpha t + (1-\alpha)s} XB^{\alpha t + (1-\alpha)s}||| \\
 &= |||A^{\alpha(t-s)} A^s X B^t B^{(1-\alpha)(s-t)}||| \\
 &\leq |||A^{t-s} A^s X B^t|||^\alpha |||A^s X B^t B^{s-t}|||^{1-\alpha} \quad \text{by (3.12)} \\
 &= |||A^t X B^t|||^\alpha |||A^s X B^s|||^{1-\alpha} \\
 &= f(t)^\alpha f(s)^{1-\alpha}.
 \end{aligned}$$

So, $f(s) = |||A^s XB^s|||$ is an AG-convex function. ■

By applying inequalities (3.3) to the function $f(t) = |||A^t XB^t|||$ on the interval $[\nu, 1-\nu]$ when $\nu \in [0, \frac{1}{2})$ and on the interval $[1-\nu, \nu]$ when $\nu \in (\frac{1}{2}, 1]$, we obtain the following theorem.

Theorem 3.2.7 [9] Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that A and B be positive operators. Then

$$\begin{aligned}
 |||A^{\frac{1}{2}}XB^{\frac{1}{2}}||| &\leq |||A^{\frac{1+2\nu}{4}}XB^{\frac{1+2\nu}{4}}|||^{\frac{1}{2}}|||A^{\frac{3-2\nu}{4}}XB^{\frac{3-2\nu}{4}}|||^{\frac{1}{2}} \\
 &\leq \exp\left(\frac{1}{1-2\nu}\int_{\nu}^{1-\nu}\log|||A^uXB^u|||du\right) \\
 &\leq |||A^{\frac{1}{2}}XB^{\frac{1}{2}}|||^{\frac{1}{2}}|||A^{\nu}XB^{\nu}|||^{\frac{1}{4}}|||A^{1-\nu}XB^{1-\nu}|||^{\frac{1}{4}} \\
 &\leq |||A^{\nu}XB^{\nu}|||^{\frac{1}{2}}|||A^{1-\nu}XB^{1-\nu}|||^{\frac{1}{2}}.
 \end{aligned}$$

Proof Utilizing inequalities (3.3) and Lemma 3.2.6, the proof is similiar to Theorem 3.2.2. ■

Let $\nu = 0$ in the above theorem. Then, we have the following result.

Corollary 3.2.8 Let A, B, X be operators in $\mathbb{B}(\mathcal{H})$ and A and B be positive operators. Then

$$\begin{aligned}
 |||A^{\frac{1}{2}}XB^{\frac{1}{2}}||| &\leq |||A^{\frac{1}{4}}XB^{\frac{1}{4}}|||^{\frac{1}{2}}|||A^{\frac{3}{4}}XB^{\frac{3}{4}}|||^{\frac{1}{2}} \\
 &\leq \exp\left(\int_0^1\log|||A^uXB^u|||du\right) \\
 &\leq |||A^{\frac{1}{2}}XB^{\frac{1}{2}}|||^{\frac{1}{2}}|||X|||^{\frac{1}{4}}|||AXB|||^{\frac{1}{4}} \\
 &\leq |||X|||^{\frac{1}{2}}|||AXB|||^{\frac{1}{2}},
 \end{aligned}$$

which is a refinement for inequality (3.18) for the especial case $\nu = \frac{1}{2}$.

Bibliography

- [1] M. Al-khlyleh, F. Kittaneh, *Interpolating inequalities related to a recent result of Audenaert*, *Linear and Multilinear Algebra*, 65(5) (2017) 922-929 .
- [2] K. Audenaert, *Interpolating between the arithmetic-geometric mean and Cauchy-Schwarz matrix norm inequalities*, *Oper. Matrices*. 9 (2015) 475-479.
- [3] R. Bhatia, C. Davis, *A Cauchy-Schwarz inequality for operators with applications*, *Linear Algebra Appl.* 22 (1995) 119-129.
- [4] F. Kittaneh, *A note on the arithmetic-geometric mean inequality for matrices*, *Linear Algebra Appl.* 171 (1992) 1-8.
- [5] F. Kittaneh, *Notes on some inequalities for Hilbert Space Operators*, *Publ. Res. Inst. Math. Sci.* 24 (2) (1988) 283-293.
- [6] M. Lin, *Remarks on two recent results of Audenaert*, *Linear Algebra Appl.* 489 (2016) 24-29.
- [7] C. P. Niculescu, *Convexity according to the geometric mean*, *Math. Inequal. Appl.* 3 (2) (2000) 155-167.
- [8] M. Sababheh, *Interpolated inequalities for unitarily invariant norms*, *Linear Algebra Appl.* 475 (2015) 240-250.
- [9] A. Taghavi, T. Azimi Roushan, V. Darvish, *Some Refinements For The Arithmetic-Geometric Mean and Cauchy-Schwarz Matrix Norm Interpolating Inequalities*. *Bulletin of the Iranian Mathematical Society*, 44(4) (2018) 927-936.
- [10] A. Taghavi, V. Darvish, H. M. Nazari, S. S. Dragomir, *Hermite-Hadamard type inequalities for operator geometrically convex functions*, *Monatsh. Math.* 181 (2016) 187-203.
- [11] L. Zou, Y. Jiang, *A note on interpolation between the arithmetic-geometric mean and Cauchy-Schwarz matrix norm inequalities*, *J. Math. Inequal.* 10 (2016) 1119-1122.

Extension of some inequalities to the numerical radius version

As we mentioned in Chapter 1, the numerical radius is a weakly unitarily invariant norm which is equivalent to the operator norm. Actually

$$\frac{1}{2}\|A\| \leq w(A) \leq \|A\|. \quad (4.1)$$

Improvements and generalizations of inequalities (4.1) have been presented in recent years. For example, Kittaneh in [12] proved that if $A \in \mathbb{B}(\mathcal{H})$, then

$$w(A) \leq \frac{1}{2}(\|A\| + \|A^2\|^{\frac{1}{2}}),$$

and

$$\frac{1}{4}\| |A|^2 + |A^*|^2 \| \leq w^2(A) \leq \frac{1}{2}\| |A|^2 + |A^*|^2 \|.$$

In [4] Dragomir also proved the inequality below, which in fact is a refinement for the second inequality (4.1):

$$w^2(A) \leq \frac{1}{2}(w(A^2) + \|A\|^2),$$

for every $A \in \mathbb{B}(\mathcal{H})$.

Among the objects in the field of numerical radius, which are very much considered, it is important to find an answer to questions such as, can one prove the unitarily invariant norm inequalities for numerical radius and vice versa? The inequalities that have been proved for the spectral radius have a numerical radius version or not? Or by reducing and weakening the conditions whether the proved inequalities again hold or not. Also, such inequalities can be generalized or improved?

In the next section, we will discuss some of these questions in more detail.

4.1 Numerical radius inequalities via convexity

In this section, using the convexity of certain numerical radius functions we can extend some unitarily invariant norms inequalities, such as Heinz and Young inequalities, to the context of numerical radius. However, due to weak unitary invariance of the numerical radius, these extensions will be weaker than the corresponding unitarily invariant norms versions.

In [3] it has been proved that for $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive operators and $0 \leq t \leq 1$

$$2\| \|A^{\frac{1}{2}}XB^{\frac{1}{2}}\| \| \leq \| \|A^tXB^{1-t} + A^{1-t}XB^t\| \| \leq \| \|AX + XB\| \| . \quad (4.2)$$

This pair of inequalities is usually referred to as Heinz inequality: in which the middle term are the Heinz means which interpolate between the geometric and arithmetic means.

On the other hand, the following Hölder-type inequality has been presented for the same parameters in [14]:

$$\| \|A^tXB^t\| \| \leq \| \|X\| \|^{1-t} \| \|AXB\| \| ^t. \quad (4.3)$$

As an application of inequality (4.3), it was proved in [14] the following inequality

$$\| \|A^tXB^{1-t}\| \| \leq \| \|AX\| \| ^t \| \|XB\| \|^{1-t}, \quad (4.4)$$

which implies the following Young-type inequality

$$\| \|A^tXB^{1-t}\| \| \leq t\| \|AX\| \| + (1-t)\| \|XB\| \| . \quad (4.5)$$

A stronger version than (4.5) would be

$$\| \|A^tXB^{1-t}\| \| \leq \| \|tAX + (1-t)XB\| \| . \quad (4.6)$$

Unfortunately, (4.6) does not hold for arbitrary unitarily invariant norm. However, it is true for the Hilbert-Schmidt norm, see [7].

A weaker version of (4.6) by letting $X = I$ is true, [2].

The main purpose of this section is to study inequalities similar to those in above, for the numerical radius.

In [19], utilizing the following lemma, it was given an example as a counterexample showing that the inequalities (4.2) and (4.5) are not exactly true for the numerical radius:

Lemma 4.1.1 [8] Suppose $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \in M_2$ and $a\bar{c}$ is real. Then

$$w(A) = \frac{1}{2}(|a+c| + \sqrt{|b|^2 + |a-c|^2}).$$

Example 4.1.2 Let $A = I_2$, $B = \text{diag}(\frac{4+\sqrt{12}}{2}, 1)$, $v = \frac{1}{2}$ and $X = \begin{bmatrix} \frac{1}{4+\sqrt{12}} & 3 \\ 0 & -2 \end{bmatrix}$. Then $AXB = \begin{bmatrix} \frac{1}{2} & 3 \\ 0 & -2 \end{bmatrix}$ and $A^2X + XB^2 = \begin{bmatrix} 2 & 6 \\ 0 & -4 \end{bmatrix}$. Now by Lemma 4.1.1 we have

$$w(AXB) \simeq 2.7025 > 2.6213 \simeq \frac{1}{2}w(A^2X + XB^2).$$

Also $A^2X = X$ and $XB^2 = \begin{bmatrix} \frac{2+\sqrt{3}}{2} & 3 \\ 0 & -2 \end{bmatrix}$, hence

$$w(AXB) \simeq 2.702, \quad w(A^2X) \simeq 2.77 \quad \text{and} \quad w(XB^2) \simeq 2.514.$$

From these numbers and by easy computation one can easily see that

$$w(AXB) > \frac{w(A^2X) + w(XB^2)}{2}.$$

The entrywise product of two matrices A and B is called their Schur (or Hadamard) product and denoted by $A \circ B$. With this multiplication M_n becomes a commutative algebra and the matrix J with all entries equal to one is its unit.

Proposition 4.1.3 [6] Let $A = (a_{ij})_{i,j=1}^n$ is a $n \times n$ positive matrix and B is arbitrary $n \times n$ matrix. Then $w(A \circ B) \leq (\max_i a_{ii}) w(B)$.

According to the above proposition, the following lemma is represented in [17].

Lemma 4.1.4 Let $A \geq 0$ and $X \in M_n$. Then

$$w(AXA) \leq \frac{1}{2}w(A^2X + XA^2).$$

Proof Suppose $A = \text{diag}(\lambda_i)$ with $\lambda_i \geq 0$. Then $AXA = Y \circ (A^2X + XA^2)$, which $Y = (\lambda_i \frac{1}{\lambda_i^2 + \lambda_j^2} \lambda_j)_{ij}$. It is clear $Y \geq 0$. By Proposition 4.1.3, we get

$$\begin{aligned} w(AXA) &\leq \max_i y_{ii} w(A^2X + XA^2) \\ &= \frac{1}{2}w(A^2X + XA^2). \end{aligned}$$

This complete the proof when A is diagonal. If A is not diagonal, then by using the spectral decomposition $A = U \text{diag}(\lambda_i) U^*$ and considering weak unitary invariance of the numerical radius, the result be implied. ■

Extension of some inequalities to the numerical radius version

Now we are ready to present the main result.

Theorem 4.1.5 [17] For $A > 0$ and $X \in M_n$ the function

$$f(t) = w(A^t X A^{1-t} + A^{1-t} X A^t)$$

is convex on \mathbb{R} , and attains its minimum at $t = \frac{1}{2}$.

Proof Since f is continuous on \mathbb{R} , it suffices to show that f is mid-convex. So, let $t, s \in \mathbb{R}$. Then

$$\begin{aligned} f\left(\frac{t+s}{2}\right) &= w\left(A^{\frac{t+s}{2}} X A^{1-\frac{t+s}{2}} + A^{1-\frac{t+s}{2}} X A^{\frac{t+s}{2}}\right) \\ &= w\left[A^{\frac{t+s}{2}} (A^s X A^{1-t} + A^{1-t} X A^s) A^{\frac{t-s}{2}}\right] \\ &\leq \frac{1}{2} w\left[A^{t-s} (A^s X A^{1-t} + A^{1-t} X A^s) + (A^s X A^{1-t} + A^{1-t} X A^s) A^{t-s}\right] \\ &= \frac{1}{2} w\left(A^t X A^{1-t} + A^{1-s} X A^s + A^s X A^{1-s} + A^{1-t} X A^t\right) \\ &\leq \frac{1}{2} \left[w\left(A^t X A^{1-t} + A^{1-t} X A^t\right) + w\left(A^s X A^{1-s} + A^{1-s} X A^s\right)\right] \\ &= \frac{f(t) + f(s)}{2}, \end{aligned}$$

which the first inequality implied by Lemma 4.1.4 and the second inequality is due to $w(\cdot)$ is a norm. Hence, the convexity of f is proved.

Since f is convex and symmetric about $t = \frac{1}{2}$, it follows that it decreases on $(-\infty, \frac{1}{2}]$ and increases on $[\frac{1}{2}, \infty)$, so it attains its minimum at $t = \frac{1}{2}$. This completes the proof. ■

From the above theorem, w -version of Heinz inequality, however in the weaker form for the case $A = B$, can be stated as follows:

Corollary 4.1.6 [17] Let $A \geq 0$ and $X \in M_n$. Then

$$w\left(A^{\frac{1}{2}} X A^{\frac{1}{2}}\right) \leq \frac{1}{2} w\left(A^t X A^{1-t} + A^{1-t} X A^t\right) \leq w\left(A X + X A\right),$$

for $0 \leq t \leq 1$.

Proof Since f attains its minimum at $t = \frac{1}{2}$, it is decreasing for $t \leq \frac{1}{2}$ and increasing for $t \geq \frac{1}{2}$ the result is immediate. ■

On the other hand, a reverse of the above inequality can be found as follows.

Corollary 4.1.7 [17] Let $A \geq 0$ and $X \in M_n$. Then for $t \notin [0, 1]$,

$$w(A^t X A^{1-t} + A^{1-t} X A^t) \geq w(AX + XA).$$

Proof By the monotonicity property of f we have $f(t) \geq f(1) = f(0)$ when $t \notin [0, 1]$. This implies the result. ■

Now we present another result which is a Young-Type inequality for numerical radius.

Theorem 4.1.8 [17] Let $A > 0$ and $X \in M_n$. Then

$$g(t) = w(A^t X A^{1-t})$$

is a convex function on \mathbb{R} .

Proof Again, it suffices to show that f is a mid-convex function. If $t, s \in \mathbb{R}$, then by Lemma 4.1.4 we have

$$\begin{aligned} g\left(\frac{t+s}{2}\right) &= w\left(A^{\frac{t+s}{2}} X A^{1-\frac{t+s}{2}}\right) \\ &= w\left[A^{\frac{t-s}{2}} (A^s X A^{1-t}) A^{\frac{t-s}{2}}\right] \\ &\leq \frac{1}{2} w[A^{t-s} (A^s X A^{1-t}) + (A^s X A^{1-t}) A^{t-s}] \\ &\leq \frac{1}{2} [w(A^t X A^{1-t}) + w(A^s X A^{1-s})] \\ &= \frac{g(t) + g(s)}{2}. \end{aligned}$$

■

Corollary 4.1.9 [17] Suppose $A > 0$ and $X \in M_n$. If $0 \leq t \leq 1$, then

$$w(A^t X A^{1-t}) \leq t w(AX) + (1-t) w(XA).$$

On the other hand, if $t \notin [0, 1]$, then

$$w(A^t X A^{1-t}) \geq t w(AX) + (1-t) w(XA).$$

Proof For convex function $g(t) = w(A^t X A^{1-t})$, we have

$$g(t) = g((1-t) \cdot 0 + t \cdot 1) \leq (1-t)g(0) + tg(1), \quad 0 \leq t \leq 1,$$

which implies the first inequality. We can also get the second inequality by knowing that for $t \notin [0, 1]$, we have $g(t) \geq (1-t)g(0) + tg(1)$. ■

Extension of some inequalities to the numerical radius version

In [16], the following one-term refinement for a convex function $y = y(t)$ on $[0, 1]$ is given by the formula

$$y(t) + 2 \min\{t, 1-t\}(y(1) + y(0) - 2y(\frac{1}{2})) \leq (1-t)y(0) + ty(1); \quad 0 \leq t \leq 1.$$

Applying this refinement to the function $f(t) = w(A^t X A^{1-t} + A^{1-t} X A^t)$ implies

$$\begin{aligned} w(A^t X A^{1-t} + A^{1-t} X A^t) + 4 \min\{t, 1-t\}(w(AX + XA) - w(A^{\frac{1}{2}} X A^{\frac{1}{2}})) \\ \leq w(AX + XA), \quad 0 \leq t \leq 1, \end{aligned}$$

which refines the Heinz inequality.

It is notable that in dealing with unitarily invariant norms, we have the stronger result that the function $f(t) = \|A^t X B^t\|$ is convex for $A, B > 0$. But this is not true with the numerical radius instead of unitarily invariant norm. In fact, mappings such as $t \rightarrow w(A^t X B^t)$, $t \rightarrow w(A^t X B^{1-t})$ and $t \rightarrow w(A^t X B^{1-t} + A^{1-t} X B^t)$ are not convex, otherwise they are in contradiction with Example 4.1.2.

The question posed by Sababheh in [17] is, in the inequality

$$w(A^t X B^{1-t}) \leq c w^t(AX) w^{1-t}(XB), \quad 0 \leq t \leq 1$$

where $A, B \geq 0$ and X is an arbitrary matrix in $M_n(\mathbb{C})$, what is the smallest positive constant c ? Noting the inequalities (4.1) and applying Hölder inequality, for $0 \leq t \leq 1$,

$$w(A^t X B^{1-t}) \leq \|A^t X B^{1-t}\| \leq \|AX\|^t \|XB\|^{1-t} \leq 2 w^t(AX) w^{1-t}(XB).$$

In the other words

$$w(A^t X B^{1-t}) \leq 2 w^t(AX) w^{1-t}(XB), \quad (4.7)$$

on the other hand, by Example 4.1.2 we know that

$$w(A^t X B^{1-t}) > w^t(AX) w^{1-t}(XB). \quad (4.8)$$

So, according to inequalities (4.7) and (4.8) we must have $1 < c \leq 2$. But the exact amount has not been fixed yet.

4.2 Some upper bounds for the Berezin number of an operator

In this section, we obtain some Berezin number inequalities based on the definition of Berezin symbol.

Before presenting the results, we introduce the Berezin number.

A functional Hilbert space is the Hilbert space of complex-valued functions on some set Ω such

4.2. Some upper bounds for the Berezin number of an operator

that the evaluation functional $\varphi_\lambda = f(\lambda)$, $\lambda \in \Omega$ are continuous on \mathcal{H} . Then by the Riesz representation theorem for each $\lambda \in \Omega$, there exists a unique function $k_\lambda \in \mathcal{H}$ such that $f(\lambda) = \langle f, k_\lambda \rangle$ for all $f \in \mathcal{H}$. The family $\{k_\lambda : \lambda \in \Omega\}$ is called the reproducing kernel of the space \mathcal{H} . For A a bounded linear operator on \mathcal{H} , the Berezin symbol of A is the function \tilde{A} on Ω defined by

$$\tilde{A}(\lambda) = \langle A\hat{k}_\lambda(z), \hat{k}_\lambda(z) \rangle,$$

where $\hat{k}_\lambda := \frac{k_\lambda}{\|k_\lambda\|}$ is the normalized reproducing kernel of the space \mathcal{H} [9, 10, 11]. Berezin set and Berezin number of operator A are defined respectively by

$$\mathbf{Ber}(A) := \{\tilde{A}(\lambda) : \lambda \in \Omega\} \text{ and } \mathbf{ber}(A) := \sup\{|\tilde{A}(\lambda)| : \lambda \in \Omega\}.$$

It is clear that the Berezin symbol \tilde{A} is the bounded function on Ω whose value lies in the numerical range of the operator A and hence for any $A \in \mathbb{B}(\mathcal{H})$,

$$\mathbf{Ber}(A) \subseteq W(A) \text{ and } \mathbf{ber}(A) \leq w(A).$$

The Berezin number of an operator A satisfies the following properties:

(i) $\mathbf{ber}(\alpha A) = |\alpha| \mathbf{ber}(A)$ for all $\alpha \in \mathbb{C}$,

(ii) $\mathbf{ber}(A + B) \leq \mathbf{ber}(A) + \mathbf{ber}(B)$.

In the following we obtain some upper bounds for the Berezin number of the geometric mean of A and B . Whatsmore, we establish some inequalities involving generalization of Berezin number inequalities.

To prove our Berezin number inequalities, we need the following well-known results.

For $a, b > 0$ and $0 \leq \nu \leq 1$, the Young inequality says that

$$a^\nu b^{1-\nu} \leq \nu a + (1 - \nu)b, \tag{4.9}$$

which was mentioned in the first chapter.

Recently Kittaneh and Manasrah in [15] refined inequality (4.9) as following

$$a^\nu b^{1-\nu} + r_0(\sqrt{a} - \sqrt{b})^2 \leq \nu a + (1 - \nu)b, \tag{4.10}$$

where $r_0 = \min\{\nu, 1 - \nu\}$.

Furthermore, in [1] they generalized inequality (4.10) in the following form.

$$(a^\nu b^{1-\nu})^m + r_0^m (a^{\frac{m}{2}} - b^{\frac{m}{2}})^2 \leq (\nu a + (1 - \nu)b)^m, \tag{4.11}$$

for $m = 1, 2, 3, \dots$.

Extension of some inequalities to the numerical radius version

From the spectral theorem for positive operators and Jensen inequality we have:

Lemma 4.2.1 [13] Let A be a positive operator in $\mathbb{B}(\mathcal{H})$ and let x be any unit vector in \mathcal{H} . Then

1. $\langle Ax, x \rangle^r \leq \langle A^r x, x \rangle$ for $r \geq 1$,
2. $\langle A^r x, x \rangle \leq \langle Ax, x \rangle^r$ for $r \leq 1$.

Dragomir in [5] obtained an useful extension for four operators of the Schwarz inequality as following.

Theorem 4.2.2 Let $A, B, C, D \in \mathbb{B}(\mathcal{H})$. Then for $x, y \in \mathcal{H}$ we have the inequality

$$|\langle DCBAx, y \rangle|^2 \leq \langle A^* |B|^2 Ax, x \rangle \langle D |C^*|^2 D^* y, y \rangle. \quad (4.12)$$

From now on, our means of r_0 and R_0 , are $\min\{v, 1-v\}$ and $\max\{v, 1-v\}$, respectively.

4.2.1 Main results

Now we are in a position to present our first result.

Theorem 4.2.3 [21] Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that $A, B > 0$ and $v \in [0, 1]$. Then for all $r \geq 2m$ ($m = 1, 2, 3, \dots$), and $\alpha \geq 0$

$$\mathbf{ber}^r((A\sharp_{\alpha}B)X) \leq \mathbf{ber}\left(v(X^*AX)^{\frac{r}{2mv}} + (1-v)(A\sharp_{2\alpha}B)^{\frac{r}{2m(1-v)}}\right)^m - r_0^m \inf_{\lambda \in \Omega} \zeta(\hat{k}_{\lambda}), \quad (4.13)$$

where $\zeta(\hat{k}_{\lambda}) = \left(\langle (X^*AX)^{\frac{r}{2m}} \hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{m}{2v}} - \langle (A\sharp_{2\alpha}B)^{\frac{r}{2m}} \hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{m}{2(1-v)}}\right)^2$.

Proof Let \hat{k}_{λ} be the normalized reproducing kernel of $\mathcal{H}(\Omega)$, then

$$\begin{aligned} |\langle (A\sharp_{\alpha}B)X\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle|^r &= |\langle A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}X\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle|^r \\ &\quad \text{By Theorem 4.2.2} \\ &\leq \langle X^*AX\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{r}{2}} \langle A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{2\alpha}A^{\frac{1}{2}}\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{r}{2}} \\ &= \left(\langle X^*AX\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{r}{2m}} \langle (A\sharp_{2\alpha}B)\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{r}{2m}}\right)^m \\ &\leq \left(\langle (X^*AX)^{\frac{r}{2m}} \hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle \langle (A\sharp_{2\alpha}B)^{\frac{r}{2m}} \hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle\right)^m \quad \text{By Lemma 4.2.1(1)}. \end{aligned}$$

4.2. Some upper bounds for the Berezin number of an operator

Now, by refinement of Young inequality (4.11) we have

$$\begin{aligned}
& \left(\langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle \right)^m \\
& \leq \left(\nu \langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{1}{\nu}} + (1-\nu) \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{1}{1-\nu}} \right)^m \\
& \quad - r_0^m \left(\langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2\nu}} - \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2(1-\nu)}} \right)^2 \\
& \leq \left(\nu \langle (X^* AX)^{\frac{r}{2m\nu}} \hat{k}_\lambda, \hat{k}_\lambda \rangle + (1-\nu) \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m(1-\nu)}} \hat{k}_\lambda, \hat{k}_\lambda \rangle \right)^m \\
& \quad - r_0^m \left(\langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2\nu}} - \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2(1-\nu)}} \right)^2 \quad \text{By Lemma 4.2.1 (2)} \\
& = \left\langle \left(\nu (X^* AX)^{\frac{r}{2m\nu}} + (1-\nu) (A\sharp_{2\alpha} B)^{\frac{r}{2m(1-\nu)}} \right) \hat{k}_\lambda, \hat{k}_\lambda \right\rangle^m \\
& \quad - r_0^m \left(\langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2\nu}} - \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2(1-\nu)}} \right)^2 \\
& \leq \mathbf{ber} \left(\nu (X^* AX)^{\frac{r}{2m\nu}} + (1-\nu) (A\sharp_{2\alpha} B)^{\frac{r}{2m(1-\nu)}} \right)^m \\
& \quad - r_0^m \left(\langle (X^* AX)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2\nu}} - \langle (A\sharp_{2\alpha} B)^{\frac{r}{2m}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{m}{2(1-\nu)}} \right)^2
\end{aligned}$$

Now, by taking supremum over $\lambda \in \Omega$, we get the desired inequality. \blacksquare

Choosing $m = 1$ in the proof of Theorem 4.2.3 we have:

Corollary 4.2.4 [21] Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that $A, B > 0$ and $\nu \in [0, 1]$. Then for all $r \geq 2R_0$

$$\mathbf{ber}^r((A\sharp_\alpha B)X) \leq \mathbf{ber} \left(\nu (X^* AX)^{\frac{r}{2\nu}} + (1-\nu) (A\sharp_{2\alpha} B)^{\frac{r}{2(1-\nu)}} \right) - r_0 \inf_{\lambda \in \Omega} \zeta(\hat{k}_\lambda), \quad (4.14)$$

where $\zeta(\hat{k}_\lambda) = (\langle (X^* AX)^{\frac{r}{4\nu}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{4\nu}} - \langle (A\sharp_{2\alpha} B)^{\frac{r}{4(1-\nu)}} \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{4(1-\nu)}})^2$.

By letting $\alpha = \frac{1}{2}$ and $m = 1$ in the proof of Theorem 4.2.3, since $A\sharp B = B\sharp A$ we obtain the following corollary which was proved earlier in [20] for the numerical radius in (p, q) -version.

Corollary 4.2.5 [21] Let $A, B, X \in \mathbb{B}(\mathcal{H})$ such that $A, B > 0$ and $\nu \in [0, 1]$. Then for all $r \geq 2R_0$

$$\mathbf{ber}^r((A\sharp B)X) \leq \mathbf{ber} \left(\nu A^{\frac{r}{2\nu}} + (1-\nu) (X^* BX)^{\frac{r}{2(1-\nu)}} \right) - r_0 \inf_{\lambda \in \Omega} \zeta(\hat{k}_\lambda), \quad (4.15)$$

where $\zeta(\hat{k}_\lambda) = (\langle A \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{4\nu}} - \langle X^* BX \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{4(1-\nu)}})^2$.

Remark 4.2.6 [21] Note that, if we set $X = I$, $r = 2$ and $\nu = \frac{1}{2}$, in (4.15), then we have

$$\mathbf{ber}^2(A\sharp B) \leq \mathbf{ber} \left(\frac{A^2 + B^2}{2} \right) - \frac{1}{2} \inf_{\lambda \in \Omega} \zeta(\hat{k}_\lambda), \quad (4.16)$$

where $\zeta(\hat{k}_\lambda) = \langle (A - B) \hat{k}_\lambda, \hat{k}_\lambda \rangle^2$. Actually, (4.16) is an operator Berezin number version for arithmetic-geometric mean.

Extension of some inequalities to the numerical radius version

The next result reads as follows.

Theorem 4.2.7 [21] Let A, B be positive definite operators in $\mathbb{B}(\mathcal{H})$ and $\nu \in [0, 1]$. Then for $\alpha \in [0, 1]$ and all $r \geq R_0/\alpha$

$$\mathbf{ber}^r(A\sharp_{\alpha}B) \leq \mathbf{ber}\left(\nu A^{\frac{(1-\alpha)r}{\nu}} + (1-\nu)B^{\frac{\alpha r}{1-\nu}}\right) - r_0 \inf_{\lambda \in \Omega} \zeta(\hat{k}_{\lambda}), \quad (4.17)$$

where $\zeta(\hat{k}_{\lambda}) = (\langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{2\nu}} - \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{2(1-\nu)}})^2$.

Proof If \hat{k}_{λ} is the normalized reproducing kernel of $\mathcal{H}(\Omega)$, then

$$\begin{aligned} \langle (A\sharp_{\alpha}B)\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^r &= \langle A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^r \\ &= \langle (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}\hat{k}_{\lambda}, A^{\frac{1}{2}}\hat{k}_{\lambda} \rangle^r \\ &\leq \|A^{\frac{1}{2}}\hat{k}_{\lambda}\|^{(2-2\alpha)r} \langle (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}\hat{k}_{\lambda}, A^{\frac{1}{2}}\hat{k}_{\lambda} \rangle^{\alpha r} \quad \text{By Lemma 4.2.1(2)} \\ &= \langle A^{\frac{1}{2}}\hat{k}_{\lambda}, A^{\frac{1}{2}}\hat{k}_{\lambda} \rangle^{(1-\alpha)r} \langle (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^{\alpha}A^{\frac{1}{2}}\hat{k}_{\lambda}, A^{\frac{1}{2}}\hat{k}_{\lambda} \rangle^{\alpha r} \\ &= \langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{(1-\alpha)r} \cdot \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\alpha r} \\ &\leq \nu \langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{\nu}} + (1-\nu) \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{1-\nu}} \\ &\quad - r_0 \left(\langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{2\nu}} - \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{2(1-\nu)}} \right)^2 \quad \text{By inequality (4.10)} \\ &\leq \nu \langle A^{\frac{(1-\alpha)r}{\nu}}\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle + (1-\nu) \langle B^{\frac{\alpha r}{1-\nu}}\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle \\ &\quad - r_0 \left(\langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{2\nu}} - \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{2(1-\nu)}} \right)^2 \quad \text{By Lemma 4.2.1(1)} \\ &= \left\langle \left(\nu A^{\frac{(1-\alpha)r}{\nu}} + (1-\nu)B^{\frac{\alpha r}{1-\nu}} \right) \hat{k}_{\lambda}, \hat{k}_{\lambda} \right\rangle \\ &\quad - r_0 \left(\langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{2\nu}} - \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{2(1-\nu)}} \right)^2 \\ &\leq \mathbf{ber}\left(\nu A^{\frac{(1-\alpha)r}{\nu}} + (1-\nu)B^{\frac{\alpha r}{1-\nu}}\right) \\ &\quad - r_0 \left(\langle A\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{(1-\alpha)r}{2\nu}} - \langle B\hat{k}_{\lambda}, \hat{k}_{\lambda} \rangle^{\frac{\alpha r}{2(1-\nu)}} \right)^2. \end{aligned}$$

Now, by taking supremum over $\lambda \in \Omega$, we get the inequality. \blacksquare

Remark 4.2.8 If we put $\alpha = \frac{1}{2}$, $r = 2$ and $\nu = \frac{1}{2}$ in (4.17), we get the inequality in (4.16).

Finally, we end this section by the following results.

Theorem 4.2.9 [21] Let $A, B \in \mathbb{B}(\mathcal{H})$ be positive definite operators and $\alpha \in [0, 1]$, then

$$\mathbf{ber}(A\sharp_{\alpha}B) \leq \mathbf{ber}^{1-\alpha}(A)\mathbf{ber}^{\alpha}(B).$$

In particular,

$$\mathbf{ber}(A\sharp B) \leq \sqrt{\mathbf{ber}(A)\mathbf{ber}(B)}.$$

4.2. Some upper bounds for the Berezin number of an operator

Proof let \hat{k}_λ be the normalized reproducing kernel of $\mathcal{H}(\Omega)$, then

$$\begin{aligned} \langle (A\#_\alpha B)\hat{k}_\lambda, \hat{k}_\lambda \rangle &= \langle A^{\frac{1}{2}}(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha A^{\frac{1}{2}}\hat{k}_\lambda, \hat{k}_\lambda \rangle \\ &= \langle (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})^\alpha A^{\frac{1}{2}}\hat{k}_\lambda, A^{\frac{1}{2}}\hat{k}_\lambda \rangle \\ &\leq \langle (A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}\hat{k}_\lambda, A^{\frac{1}{2}}\hat{k}_\lambda \rangle^\alpha \langle A^{\frac{1}{2}}\hat{k}_\lambda, A^{\frac{1}{2}}\hat{k}_\lambda \rangle^{(1-\alpha)} \\ &= \langle A\hat{k}_\lambda, \hat{k}_\lambda \rangle^{(1-\alpha)} \cdot \langle B\hat{k}_\lambda, \hat{k}_\lambda \rangle^\alpha. \end{aligned}$$

Now, by taking supremum over $\lambda \in \Omega$, we get the first inequality. In particular, for $\alpha = \frac{1}{2}$ we obtain the second one. ■

Corollary 4.2.10 [21] Let $A, B \in \mathbb{B}(\mathcal{H})$ be positive definite operators which commute with each other and $\alpha \in [0, 1]$, then

$$\mathbf{ber}(A^{1-\alpha}B^\alpha) \leq \mathbf{ber}^{1-\alpha}(A)\mathbf{ber}^\alpha(B).$$

In particular, if $\alpha = \frac{1}{2}$, then

$$\mathbf{ber}(\sqrt{AB}) \leq \sqrt{\mathbf{ber}(A)\mathbf{ber}(B)}.$$

4.2.2 Additional results

To prove our results in this section, the following basic lemmas are also required.

Lemma 4.2.11 [13] Let A be an operator in $\mathbb{B}(\mathcal{H})$, and f, g be nonnegative functions on $[0, \infty)$ which are continuous and satisfy the relation $f(t)g(t) = t$ for all $t \in [0, \infty)$. Then for all x, y in \mathcal{H} ,

$$|\langle Ax, y \rangle| \leq \|f(|A|)x\| \|g(|A^*|)y\|. \quad (4.18)$$

Lemma 4.2.12 [18] Let a_i be positive real numbers, ($i = 1, 2, \dots, n$). Then

$$\left(\sum_{i=1}^n a_i \right)^r \leq n^{r-1} \sum_{i=1}^n a_i^r \quad \forall r \geq 1. \quad (4.19)$$

The following result is proved in [18], for the numerical radius.

Theorem 4.2.13 [21] Let $A_i, B_i, X_i \in \mathbb{B}(\mathcal{H})$ ($i = 1, 2, \dots, n$), and let f and g be nonnegative continuous functions on $[0, \infty)$ which satisfy the relation $f(t)g(t) = t$ for all $t \in [0, \infty)$. Then

$$\mathbf{ber}^r \left(\sum_{i=1}^n A_i^* X_i B_i \right) \leq \frac{n^{r-1}}{2} \mathbf{ber} \left(\sum_{i=1}^n ([A_i^* g^2(|X_i^*|)A_i]^r + [B_i^* f^2(|X_i|)B_i]^r) \right) \quad (4.20)$$

for all $r \geq 1$.

Extension of some inequalities to the numerical radius version

The proof is similiar to that in [18], but we bring here with more details. If \hat{k}_λ is the normalized reproducing kernel of $\mathcal{H}(\Omega)$, then

$$\begin{aligned}
& \left| \left\langle \left(\sum_{i=1}^n A_i^* X_i B_i \right) \hat{k}_\lambda, \hat{k}_\lambda \right\rangle \right|^r \\
&= \left| \sum_{i=1}^n \langle A_i^* X_i B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle \right|^r \\
&\leq \left(\sum_{i=1}^n |\langle A_i^* X_i B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle| \right)^r \\
&= \left(\sum_{i=1}^n |\langle X_i B_i \hat{k}_\lambda, A_i \hat{k}_\lambda \rangle| \right)^r \\
&\leq \left(\sum_{i=1}^n \langle f^2(|X_i|) B_i \hat{k}_\lambda, B_i \hat{k}_\lambda \rangle^{\frac{1}{2}} \langle g^2(|X_i^*|) A_i \hat{k}_\lambda, A_i \hat{k}_\lambda \rangle^{\frac{1}{2}} \right)^r \quad \text{By (4.2.11)} \\
&\leq n^{r-1} \sum_{i=1}^n \langle B_i^* f^2(|X_i|) B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{2}} \langle A_i^* g^2(|X_i^*|) A_i \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{2}} \quad \text{By (4.2.12)} \\
&\leq n^{r-1} \sum_{i=1}^n \langle (B_i^* f^2(|X_i|) B_i)^r \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{1}{2}} \langle (A_i^* g^2(|X_i^*|) A_i)^r \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{1}{2}} \quad \text{By Lemma 4.2.1} \\
&\leq \frac{n^{r-1}}{2} \sum_{i=1}^n \left(\langle [B_i^* f^2(|X_i|) B_i]^r \hat{k}_\lambda, \hat{k}_\lambda \rangle + \langle [A_i^* g^2(|X_i^*|) A_i]^r \hat{k}_\lambda, \hat{k}_\lambda \rangle \right) \quad \text{By (4.9)} \\
&= \frac{n^{r-1}}{2} \left\langle \sum_{i=1}^n \left([B_i^* f^2(|X_i|) B_i]^r + [A_i^* g^2(|X_i^*|) A_i]^r \right) \hat{k}_\lambda, \hat{k}_\lambda \right\rangle \\
&\leq \frac{n^{r-1}}{2} \mathbf{ber} \left(\sum_{i=1}^n \left([A_i^* g^2(|X_i^*|) A_i]^r + [B_i^* f^2(|X_i|) B_i]^r \right) \right)
\end{aligned}$$

Now, by taking supremum over $\lambda \in \Omega$, we get the desired inequality. \blacksquare

If we take $f(t) = t^\alpha$ and $g(t) = t^{1-\alpha}$, $\alpha \in (0, 1)$, in inequality (4.20), we get the following inequality.

Corollary 4.2.14 [21] Let $A_i, B_i, X_i \in \mathbb{B}(\mathcal{H})$ ($i = 1, 2, \dots, n$), $0 < \alpha < 1$. Then

$$\mathbf{ber}^r \left(\sum_{i=1}^n A_i^* X_i B_i \right) \leq \frac{n^{r-1}}{2} \mathbf{ber} \left(\sum_{i=1}^n \left([A_i^* |X_i^*|^{2(1-\alpha)} A_i]^r + [B_i^* |X_i|^{2\alpha} B_i]^r \right) \right) \quad (4.21)$$

for $r \geq 1$.

Inequality (4.21) includes some special cases as follows.

Corollary 4.2.15 [21] Let $A, B, X \in \mathbb{B}(\mathcal{H})$. Then

- $\mathbf{ber}^r(A) \leq \frac{1}{2} \mathbf{ber}(|A|^r + |A^*|^r) \quad \forall r \geq 1$,
- $\mathbf{ber}(A^* B) \leq \frac{1}{2} \mathbf{ber}(A^* A + B^* B)$,

4.2. Some upper bounds for the Berezin number of an operator

- $\mathbf{ber}(A^*XB) \leq \frac{1}{2}\mathbf{ber}(A^*|X^*|A+B^*|X|B)$.

Now, we want to generalize inequality (4.20) in the following form.

Theorem 4.2.16 [21] Let $A_i, B_i, X_i \in \mathbb{B}(\mathcal{H})$ ($i = 1, 2, \dots, n$), and let f and g be nonnegative functions on $[0, \infty)$ which are continuous and satisfy the relation $f(t)g(t) = t$ for all $t \in [0, \infty)$. Then for $\nu \in [0, 1]$ and $r \geq 2R_0$

$$\mathbf{ber}^r \left(\sum_{i=1}^n A_i^* X_i B_i \right) \leq n^{r-1} \mathbf{ber} \left(\sum_{i=1}^n \nu (B_i^* f^2(|X_i|) B_i)^{\frac{r}{2\nu}} + (1-\nu) (A_i^* g^2(|X_i^*|) A_i)^{\frac{r}{2(1-\nu)}} \right). \quad (4.22)$$

Proof let \hat{k}_λ be the normalized reproducing kernel of $\mathcal{H}(\Omega)$, then

$$\begin{aligned} & \left| \left\langle \left(\sum_{i=1}^n A_i^* X_i B_i \right) \hat{k}_\lambda, \hat{k}_\lambda \right\rangle \right|^r \\ &= \left| \sum_{i=1}^n \langle A_i^* X_i B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle \right|^r \\ &\leq \left(\sum_{i=1}^n |\langle A_i^* X_i B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle| \right)^r \\ &= \left(\sum_{i=1}^n |\langle X_i B_i \hat{k}_\lambda, A_i \hat{k}_\lambda \rangle| \right)^r \\ &\leq \left(\sum_{i=1}^n \langle f^2(|X_i|) B_i \hat{k}_\lambda, B_i \hat{k}_\lambda \rangle^{\frac{1}{2}} \langle g^2(|X_i^*|) A_i \hat{k}_\lambda, A_i \hat{k}_\lambda \rangle^{\frac{1}{2}} \right)^r \quad \text{By (4.2.11)} \\ &\leq n^{r-1} \sum_{i=1}^n \langle B_i^* f^2(|X_i|) B_i \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{2}} \langle A_i^* g^2(|X_i^*|) A_i \hat{k}_\lambda, \hat{k}_\lambda \rangle^{\frac{r}{2}} \quad \text{By (4.2.12)} \\ &\quad \text{By inequality (4.9) and Lemma 4.2.1} \\ &\leq n^{r-1} \sum_{i=1}^n \left(\nu \langle (B_i^* f^2(|X_i|) B_i)^{\frac{r}{2\nu}} \hat{k}_\lambda, \hat{k}_\lambda \rangle + (1-\nu) \langle (A_i^* g^2(|X_i^*|) A_i)^{\frac{r}{2(1-\nu)}} \hat{k}_\lambda, \hat{k}_\lambda \rangle \right) \\ &= n^{r-1} \left\langle \sum_{i=1}^n \left(\nu (B_i^* f^2(|X_i|) B_i)^{\frac{r}{2\nu}} + (1-\nu) (A_i^* g^2(|X_i^*|) A_i)^{\frac{r}{2(1-\nu)}} \right) \hat{k}_\lambda, \hat{k}_\lambda \right\rangle \\ &\leq n^{r-1} \mathbf{ber} \left(\sum_{i=1}^n \left(\nu (B_i^* f^2(|X_i|) B_i)^{\frac{r}{2\nu}} + (1-\nu) (A_i^* g^2(|X_i^*|) A_i)^{\frac{r}{2(1-\nu)}} \right) \right) \end{aligned}$$

Now, the result follows by taking the supremum over $\lambda \in \Omega$. ■

By letting $A_i = B_i = I$ ($i = 1, 2, \dots, n$), and $f(t) = t^\alpha$ and $g(t) = t^{1-\alpha}$, $\alpha \in (0, 1)$, in inequality (4.22), we obtain the following inequalities.

Corollary 4.2.17 [21] Let $X_i \in \mathbb{B}(\mathcal{H})$ ($i = 1, 2, \dots, n$) and $0 < \alpha < 1$. Then for $\nu \in [0, 1]$ and $r \geq R_0 \alpha$

$$\mathbf{ber}^r \left(\sum_{i=1}^n X_i \right) \leq n^{r-1} \mathbf{ber} \left(\sum_{i=1}^n \nu |X_i|^{\frac{r\alpha}{\nu}} + (1-\nu) |X_i^*|^{\frac{r(1-\alpha)}{1-\nu}} \right). \quad (4.23)$$

Extension of some inequalities to the numerical radius version

In particular, if $X_1 = X_2 = \dots = X_n = X$, then

$$\mathbf{ber}^r(X) \leq \mathbf{ber} \left(\nu |X|^{\frac{ar}{\nu}} + (1-\nu) |X^*|^{\frac{(1-a)r}{1-\nu}} \right). \quad (4.24)$$

As special cases of (4.22), (4.23) and (4.24), we present the following inequalities.

- $\mathbf{ber}^r(A) \leq \mathbf{ber} \left(\nu |A|^{\frac{r}{2\nu}} + (1-\nu) |A^*|^{\frac{r}{2(1-\nu)}} \right),$
- $\mathbf{ber}^r(A^*B) \leq \mathbf{ber} \left(\nu |B|^{\frac{r}{\nu}} + (1-\nu) |A|^{\frac{r}{1-\nu}} \right),$
- $\mathbf{ber}^r(A^*XB) \leq \mathbf{ber} \left(\nu (B^*|X|B)^{\frac{r}{2\nu}} + (1-\nu) (A^*|X^*|A)^{\frac{r}{2(1-\nu)}} \right).$

Bibliography

- [1] Y. Al-Manasrah, F. Kittaneh, A generalization of two refined Young inequalities. *Positivity* 19 (2015) 757-768.
- [2] T. Ando, Matrix Young inequalities, *oper. Theory Adv. Appl.* 75 (1995) 33-38.
- [3] R. Bhatia, Interpolating the arithmetic-geometric mean inequality and its operator version, *Linear Algebra Appl.* 413 (2006) 355-363.
- [4] S.S. Dragomir, Some inequalities for the norm and the numerical radius of linear operators in Hilbert spaces, *Tamkang J. Math.* 39 (1) (2008) 1-7.
- [5] S.S. Dragomir, Some inequalities generalizing Kato's and Furuta's results, *Filomat* 28(1) (2014) 179-195.
- [6] H.L. Gau, P.Y. Wu, Numerical radius of Hadamard product of matrices, *Linear Algebra Appl.* 504 (2016) 292-308.
- [7] O. Hirzallah, F. Kittaneh, Matrix Young inequalities for the Hilbert-Schmidt norm, *Linear Algebra Appl.* 308 (2000) 77-84.
- [8] C.R. Johnson, I. Spitkovsky and S. Gottlieb, Inequalities involving the numerical radius, *Linear and Multilinear Algebra*, (37) (1994) 13-24.
- [9] M.T. Karaev, On the Berezin symbol and invertibility of operators on the functional Hilbert Spaces, *J. Funct. Anal.* 238 (2006) 181-192.
- [10] M.T. Karaev, S. Saltan, Some results on Berezin symbols, *Complex Var. Theory Appl.* 50 (3) (2005) 185-193.
- [11] M.T. Karaev, Reproducing kernels and Berezin symbols techniques in various questions of operator theory, *Complex Anal. Oper. Theory* 7 (4) (2013) 983-1018.
- [12] F. Kittaneh, A numerical radius inequality and an estimate for the numerical radius of the Frobenius companion matrix. *Studia Math.* 158 (2003) 11-17.

Bibliography

- [13] F. Kittaneh, *Notes on some inequalities for Hilbert Space Operators*, *Publ. Res. Inst. Math. Sci.* 24 (2) (1988) 283-293.
- [14] F. Kittaneh, *Norm inequalities for fractional powers of positive operators*, *Lett. Math. Phys.* 27 (1993) 279-285.
- [15] F. Kittaneh, Y. Manasrah, *Improved Young and Heinz inequalities for matrices*, *J. Math. Anal. Appl.* 361 (2010) 262-269.
- [16] M. Sababheh, *Interpolated inequalities for unitarily invariant norms*, *Linear Algebra Appl.* 475 (2015) 240-250.
- [17] M. Sababheh, *Numerical radius inequalities via convexity*, *Linear Algebra and its Application*, 549 (2018) 67-78.
- [18] KH. Shebrawi, H. Albadawi, *Numerical radius and operator norm inequalities*, *J. of Inequal. and Appl.* (2009) 1-11.
- [19] A. Sheikhhosseini, *An arithmetic-geometric mean inequality related to numerical radius of matrices*, *Konuralp J. Math.* 5 (2017) 85-91.
- [20] A. Sheikhhosseini, *A numerical radius version of the arithmetic-geometric mean of operators*, *Filomat* 30(8) (2016) 2139-2145.
- [21] A. Taghavi, T. Azimi Roushan, V. Darvish, *Some upper bounds for the Berezin number of Hilbert space operators*, *Filomat* 33(14) (2019) 4353-4360.

Approximation of differentiable convex functions

We begin this chapter with giving two equivalent definitions of a polytope and then state some well-known fundamental theorems and properties of such a geometric object. In the third section, we define the notion of generalized barycentric coordinates with respect to an arbitrary set of points in \mathbb{R}^d , or equivalently, with respect to a (convex) polytope. In the sequel, we deal with the discussion of the approximation of differentiable convex functions in the concept of real functions, and then we use the powerful tools of the Double Operator Integral (DOI) and continuous functional calculus to extend the results into the operator differentiable convex functions. Finally, we apply the obtained results for the Hermite-Hadamard inequality.

5.1 Convex polytopes

Convex polytopes are fundamental geometric objects. In the following, we give two different versions of the definition of a polytope. The two versions are mathematically, but not algorithmically, equivalent. The proof of equivalence between the two concepts is nontrivial, see [25].

Definition 5.1.1 (polytope) A V -polytope is the convex hull of a finite set of points in some \mathbb{R}^d . An H -polyhedron is an intersection of finitely many closed halfspaces in some \mathbb{R}^d . An H -polytope is an H -polyhedron that is bounded in the sense that it does not contain a ray $\{x + ty : t \geq 0\}$ for any $y \neq 0$.

Now we present a basic version of the representation theorem for polytopes.

Theorem 5.1.2 (Main theorem for polytopes)[25] A subset $P \subset \mathbb{R}^d$ is the convex hull of a finite point set (a V -polytope)

$$P = \text{conv}(V) \text{ for some } V \subset \mathbb{R}^d$$

if and only if it is a bounded intersection of halfspaces (an H -polytope)

$$P = P(A, z) = \{x \in \mathbb{R}^d : Ax \leq z\} \text{ for some } A \in \mathbb{R}^{m \times d}, z \in \mathbb{R}^m.$$

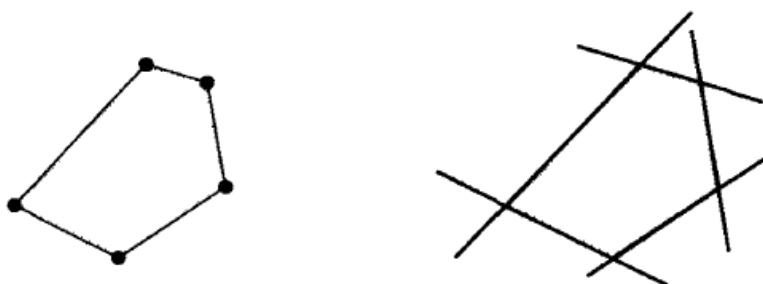


Figure 5.1 – The left figure shows a pentagon constructed as a V -polytope as the convex hull of five points; the right figure shows the same pentagon as an H -polytope.

In the above theorem, $Ax \leq z$ is namely $\langle a_1, x \rangle \leq z_1, \dots, \langle a_m, x \rangle \leq z_m$, where a_1, \dots, a_m are the rows of A , and z_1, \dots, z_m are the components of z .

A compact convex set $C \subset \mathbb{R}^d$ is a polytope provided $\text{ext} C$ (the set of all extreme points of C) is a finite set. From the results of [9, Section 2.4 and Theorem 2.3.4] it follows that polytopes may equivalently be defined as convex hulls of finite sets. For a polytope C , it is customary to call the points of $\text{ext} C$ as vertices. We denote them by $\text{vert} C$. The dimension of a polytope is the dimension of its affine hull. A d -polytope is a polytope of dimension d in some \mathbb{R}^e ($e \geq d$). For examples, zero-dimensional polytopes are points, one-dimensional polytopes are line segments, two-dimensional polytopes are called polygons. A polygon with n vertices is called an n -gon. Convexity here requires that the interior angles (at the vertices) are all smaller than π . The following illustrations show a convex 6-gon, or hexagon, also, the tetrahedron which is a familiar geometric object (a 3-dimensional polytope) in \mathbb{R}^3 . Similarly, its d -dimensional generalization forms the first (and simplest) infinite family of higher-dimensional polytopes we want to consider.

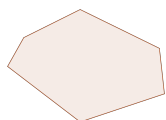
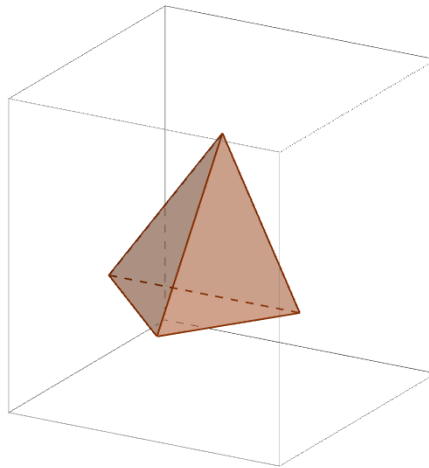


Figure 5.2 – 6-gon or hexagon.

Figure 5.3 – The tetrahedron in \mathbb{R}^3 .

Theorem 5.1.3 [25] For any compact subset M of \mathbb{R}^d , the convex hull $\text{conv}(M)$ is again compact.

Since any finite set is compact, the following result is immediate.

Corollary 5.1.4 Any convex polytope P in \mathbb{R}^d is a compact set.

In the following proposition we give some simple but basic facts about polytopes and their vertices, see [25, Proposition 2.2].

Proposition 5.1.5 The following statements hold:

- (i) Every polytope P is the convex hull of its vertices: $P = \text{conv}(\text{vert}(P))$,
- (ii) If a polytope P can be written as the convex hull of a finite point set, then the set contains all the vertices of the polytope: $P = \text{conv}(V)$ implies that $\text{vert}(P) \subseteq V$.

5.2 Barycentric coordinates

One of the important concepts related to the concept of polytopes, is the notion of barycentric coordinates, which were first introduced by August Ferdinand Möbius (1790-1816) in his book *the barycentric calculus* [12]. These coordinates are useful for simply representing a point in a triangle as a convex combination of its vertices, and frequently occur in computer graphics, modelling geometry triangular meshes, terrain modelling and the finite element method.

For simplices, barycentric coordinates are a very common tool in many computations. Basically, they are defined as follows: let $X_d = \{\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_d\}$ be a set of $d + 1$ affinely independent points

in \mathbb{R}^d , the simplex T with the set of vertices \mathbf{X}_d is the convex hull of \mathbf{X}_d , (e.g., a triangle in 2D or a tetrahedron in 3D). Let $A_i(\mathbf{x})$ be the signed volume (or area) of the subsimplex of T created with the vertex \mathbf{v}_i replaced by \mathbf{x} . Then the barycentric coordinate functions $\{\lambda_0, \lambda_1, \dots, \lambda_d\}$ of the simplex T with respect to its vertices are uniquely defined by

$$\lambda_i(\mathbf{x}) = \frac{A_i(\mathbf{x})}{\text{vol}(T)} \quad (5.1)$$

where $\text{vol}(T)$ will mean the volume measure of T . It is easily seen that each point \mathbf{x} of T has a (unique) representation, that $\mathbf{x} = \sum_{i=0}^d \lambda_i(\mathbf{x}) \mathbf{v}_i$ and the barycentric coordinates $\{\lambda_0, \lambda_1, \dots, \lambda_d\}$ are nonnegative affine functions on T . The uniqueness of this representation allows the weights $\lambda_i(\mathbf{x})$ to be interpreted as an alternative set of coordinates for point \mathbf{x} , the so-called barycentric coordinates. Note that a d -simplex is a special polytope given as the convex hull of $d+1$ vertices in d dimensions, each pair of which is joined by an edge. For $n > d$, which is the case of interest in this chapter, the linear constraints form an under-determined system.

In our study we need to deal with polytope in higher dimension too, thus we need to generalize these coordinates to any polytope in \mathbb{R}^d .

5.3 Generalized barycentric coordinates on polytopes

For a (convex) polytope $P \subseteq \mathbb{R}^d$ we will use generalized barycentric coordinates (they are often called generalized barycentric coordinates to distinguish them from the original barycentric coordinates, which were only defined with respect to simplices). While barycentric coordinates are unique for simplices, there are many possible solutions for polygons with more sides. Usual Barycentric coordinates are natural coordinates for meshes, and have many useful applications, ranging from Gouraud and Phong shading, rendering of quadrilaterals, image warping, mesh deformation and finite element applications, see, e.g., [13, 23]. From now on, let $\Omega \subset \mathbb{R}^d$ be a compact convex polytope of positive measure, and let $\mathbf{X} := \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ be a finite subset that includes the vertices of Ω . Thus, the convex hull of \mathbf{X} must be equal to Ω . We wish to construct one coordinate function $\lambda_i(\mathbf{x})$ per point \mathbf{x}_i for all $\mathbf{x} \in \Omega$ and we will consider it in the next chapter. These functions are called barycentric coordinates with respect to $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n$ (or Ω) if they satisfy three properties. First, the coordinate functions are nonnegative on Ω ,

$$\lambda_i(\mathbf{x}) \geq 0, \text{ for all } \mathbf{x} \in \Omega. \quad (5.2)$$

Secondly, the functions form a partition of unity, which means that the equation

$$\sum_{i=0}^n \lambda_i(\mathbf{x}) = 1, \text{ for all } \mathbf{x} \in \Omega. \quad (5.3)$$

5.4. Approximation of a class of convex functions

Finally, the functions act as coordinates in that, given a value of \mathbf{x} , weighting each point \mathbf{x}_i by $\lambda_i(\mathbf{x})$ returns back \mathbf{x} i.e.,

$$\mathbf{x} = \sum_{i=0}^n \lambda_i(\mathbf{x}) \mathbf{x}_i \text{ for all } \mathbf{x} \in \Omega. \quad (5.4)$$

The last property is also sometimes referred to as linear precision since the coordinate functions can reproduce linear functions. For most potential applications, it is also preferable that these coordinate functions are as smooth as possible. Constructing the barycentric coordinates of a point \mathbf{x} with respect to some given points in a polytope Ω is often not a trivial task. The first result on the existence of barycentric coordinates for more general types of polytopes was due to Kalman:

Theorem 5.3.1 [10] *Let P be a polytope in \mathbb{R}^d , $\{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$ its vertices. Then there exist non-negative real-valued continuous functions on P ,*

$$\lambda = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$$

such that

$$\mathbf{x} = \sum_{i=0}^n \lambda_i(\mathbf{x}) \mathbf{x}_i \text{ and } \sum_{i=0}^n \lambda_i(\mathbf{x}) = 1 \text{ for each } \mathbf{x} \in P. \quad (5.5)$$

Note also that Equations (5.5) can be rewritten in the following general way:

$$\sum_{i=0}^n \lambda_i(\mathbf{x}) (\mathbf{x} - \mathbf{x}_i) = \mathbf{0}, \quad (5.6)$$

which obviously implies

$$\sum_{i=0}^n \int_P \lambda_i(\mathbf{x}) (\mathbf{x} - \mathbf{x}_i) d\mathbf{x} = \mathbf{0}. \quad (5.7)$$

5.4 Approximation of a class of convex functions

In this section, after expressing the definitions and necessary preconditions, we investigate some of the results which derived by Guessab in [7].

Suppose that $X_n = \{\mathbf{x}_i\}_{i=0}^n$ is a finite set of pairwise distinct points in \mathbb{R}^d , and $P = \text{conv}(X_n)$ be a polytope of the points in X_n . In [7] the author approximated an unknown scalar-valued continuous convex function $f : P \rightarrow \mathbb{R}$ from given function values $f(x_0), f(x_1), \dots, f(x_n)$ sampled at X_n , by considering a weighted average of the function values at data points of the following form

$$B_n[f](x) = \sum_{i=0}^n \lambda_i(x) f(x_i), \quad (5.8)$$

Approximation of differentiable convex functions

which the function $\lambda_i : P \rightarrow \mathbb{R}$, $i = 0, 1, \dots, n$ are barycentric coordinates and B_n is referred as barycentric approximation.

By decreasing the dimension to $d = 1$ we can simplify $P := \{x_0, x_1, \dots, x_n\}$ as a partition of an interval $\mathcal{I} := [a, b]$ such that $a = x_0 < x_1 < \dots < x_n = b$, and approximate $f : \mathcal{I} \rightarrow \mathbb{R}$ in the following form

$$B_n[f](x) = \sum_{i=0}^n \lambda_i(x) f(x_i), \quad (x \in [a, b])$$

where

$$\lambda_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}}, & \text{if } x_{i-1} \leq x \leq x_i; \\ \frac{x_{i+1}-x}{x_{i+1}-x_i}, & \text{if } x_i \leq x \leq x_{i+1}; \\ 0, & \text{for all other } x. \end{cases}$$

Here, it was considered $x_{-1} := a$ and $x_{n+1} := b$.

Definition 5.4.1 Let C is an open subset of \mathbb{R}^d . A differentiable function $f : C \rightarrow \mathbb{R}$ has a Lipschitz-continuous gradient with constant $l \geq 0$ if and only if

$$\|\nabla f(x) - \nabla f(y)\| \leq l \|x - y\|, \quad \forall x, y \in C.$$

Definition 5.4.2 The mapping $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is said to be co-coercive with constant c if for all $x, y \in \mathbb{R}^d$

$$\langle F(x) - F(y), x - y \rangle \geq c \|F(x) - F(y)\|^2.$$

The following theorem offers simple ways to characterize differentiable convex function with Lipschitz-continuous gradient.

Theorem 5.4.3 [14] Let f be a differentiable convex function defined on \mathbb{R}^d . Then the following properties are equivalent.

1. ∇f is Lipschitz-continuous with constant l .
2. The function f satisfies for all $x, y \in \mathbb{R}^d$

$$f(x) + \langle \nabla f(x), y - x \rangle \leq f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{l}{2} \|y - x\|^2.$$

3. The function f satisfies for all $x, y \in \mathbb{R}^d$

$$\frac{1}{2l} \|\nabla f(x) - \nabla f(y)\|^2 + f(x) + \langle \nabla f(x), y - x \rangle \leq f(y).$$

5.4. Approximation of a class of convex functions

4. ∇f is co-coercive mapping with constant $\frac{1}{l}$. That is

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{1}{l} \|\nabla f(x) - \nabla f(y)\|^2, \quad \forall x, y \in \mathbb{R}^d.$$

We know from Theorem 5.4.3 part (4), that if $f \in C^{1,1}(P)$ with $l_f > 0$ and, in addition, f be convex. Then ∇f satisfies the following property

$$\frac{1}{l_f} \|\nabla f(y) - \nabla f(x)\|^2 \leq \langle \nabla f(y) - \nabla f(x), y - x \rangle, \quad \forall x, y \in P, \quad (5.9)$$

where $C^{1,1}(P)$ denote the subclass of all functions f which continuously differentiable on P with Lipschitz-continuous gradient, i.e., there exists a constant l_f , which can not be replaced by smaller one, such that

$$\|\nabla f(x) - \nabla f(y)\| \leq l_f \|x - y\|, \quad (\forall x, y \in P).$$

For a convex function $f \in C^{1,1}(P)$, the exclusive symbol

$$E_n[f](x) := E_n[f, \lambda](x) = \sum_{i=0}^n \lambda_i(x) f(x_i) - f(x) \quad (5.10)$$

was reserved to show the error which is imposed by barycentric approximation of f . By inequality (5.9) we have the following result (see [6, Proposition 2.2]).

Proposition 5.4.4 If $f \in C^{1,1}(P)$ with Lipschitz constant $l_f > 0$, then the functions

$$g_{\pm} := \frac{l_f}{2} \|\cdot\|^2 \pm f$$

are both convex and belong to $C^{1,1}(P)$. If in addition f is convex, then $l_{g_{-}} \leq l_f$.

Proof The proof is similiar to that in [7], but we bring here with more details.

We need to show that the function g_{\pm} also belong to $C^{1,1}(P)$. Indeed, they are obviously differentiable and it is easy to check that

$$\|\nabla g_{\pm}(y) - \nabla g_{\pm}(x)\| = \|l_f(y - x) \pm (\nabla f(y) - \nabla f(x))\|,$$

which by triangle inequality it implies

$$\|\nabla g_{\pm}(y) - \nabla g_{\pm}(x)\| \leq l_f \|y - x\| + \|\nabla f(y) - \nabla f(x)\| \leq 2l_f \|y - x\|,$$

Hence, we have $l_{g_{\pm}} \leq 2l_f$. Moreover, since $f \in C^{1,1}(P)$, then by the Cauchy-Schwartz inequality

Approximation of differentiable convex functions

we obtain

$$\begin{aligned} \pm \langle \nabla f(y) - \nabla f(x), y - x \rangle &\leq |\langle \nabla f(y) - \nabla f(x), y - x \rangle| \\ &\leq \|\nabla f(y) - \nabla f(x)\| \|y - x\| \\ &\leq l_f \|y - x\| \|y - x\|, \end{aligned}$$

then

$$\pm \langle \nabla f(y) - \nabla f(x), y - x \rangle \leq l_f \|y - x\|^2$$

and so

$$l_f \|y - x\|^2 \pm \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

From this, it is immediate

$$\langle \nabla g_{\pm}(y) - \nabla g_{\pm}(x), y - x \rangle = l_f \|y - x\|^2 \pm \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0,$$

which according to theorem 1.1.6 means that g_{\pm} are convex.

What remains to be shown is that if f is in addition convex, we have $l_{g_-} \leq l_f$. Since function g_- has a Lipschitz-continuous gradient, then due to convexity of f and inequality (5.9) we get

$$\begin{aligned} \|\nabla g_-(y) - \nabla g_-(x)\|^2 &\leq l_{g_-} \langle \nabla g_-(y) - \nabla g_-(x), y - x \rangle \\ &= l_{g_-} \langle l_f(y - x) - (\nabla f(y) - \nabla f(x)), y - x \rangle \\ &= l_{g_-} \cdot l_f \|y - x\|^2 - l_{g_-} \langle \nabla f(y) - \nabla f(x), y - x \rangle \\ &\leq l_{g_-} \cdot l_f \|y - x\|^2. \end{aligned}$$

So,

$$\|\nabla g_-(y) - \nabla g_-(x)\| \leq \sqrt{l_{g_-} \cdot l_f} \|y - x\|.$$

Since l_{g_-} is the smallest possible Lipschitz constant, we have

$$l_{g_-} \|y - x\| \leq \sqrt{l_{g_-} \cdot l_f} \|y - x\|,$$

this allows us to conclude that $l_{g_-} \leq l_f$, and this completes the proof. ■

Utilizing Proposition (5.4.4), in [7] Guessab stated and proved an elegant characterization of all upper approximation operators, which we bring here with more details.

Theorem 5.4.5 [7] Let $T : C^1(P) \rightarrow C(P)$ be a linear operator. The following statements are equivalent:

(i) For every convex function $g \in C^{1,1}(P)$, we have

$$g(x) \leq T[g](x), \quad (x \in P). \quad (5.11)$$

(ii) For every $f \in C^{1,1}(P)$ with a Lipschitz constant l_f , we have

$$|f(x) - T[f](x)| \leq \frac{l_f}{2} (T[\|\cdot\|^2](x) - \|x\|^2). \quad (5.12)$$

Equality is attained for all functions of the form

$$f(x) = a(x) + c\|x\|^2, \quad (5.13)$$

where $c \in \mathbb{R}$ and $a(\cdot)$ is any affine function.

Proof Let $f \in C^{1,1}(P)$ with a Lipschitz constant l_f and suppose that (i) holds. Define the two following functions

$$g_{\pm} := \frac{l_f}{2} \|\cdot\|^2 \pm f.$$

Due to Proposition (5.4.4), we know that both of these functions are convex and belong to $C^{1,1}(P)$. Therefore, since T is linear, statement (i) implies

$$\begin{aligned} T\left[\frac{l_f}{2} \|\cdot\|^2 \pm f\right] &\geq \frac{l_f}{2} \|\cdot\|^2 \pm f \\ \Rightarrow T\left[\frac{l_f}{2} \|\cdot\|^2\right] \pm T[f] &\geq \frac{l_f}{2} \|\cdot\|^2 \pm f \\ \Rightarrow \frac{l_f}{2} T[\|\cdot\|^2] \pm T[f] &\geq \frac{l_f}{2} \|\cdot\|^2 \pm f, \end{aligned}$$

which gives the error estimate in statement (ii). Let $x \in P$, so there exists $y \in \mathbb{R}^d$ and a real constant c such that $a(x) = \langle x, y \rangle + c$. Hence, the case of equality is easily verified.

Conversely, let $g \in C^{1,1}(P)$ be a convex function, and suppose that statement (ii) holds. Let the function f is given by

$$f := \frac{l_g}{2} \|\cdot\|^2 - g.$$

And set $E := T - I$, where I is the identity on $C^{1,1}(P)$. Utilizing Proposition 5.4.4 again. we have $f \in C^{1,1}(P)$ with $l_f \leq l_g$. Now, the error estimate (ii), applied to f , implies that

$$E\left[\frac{l_g}{2} \|\cdot\|^2 - g\right] \leq \frac{l_f}{2} E[\|\cdot\|^2] \leq \frac{l_g}{2} E[\|\cdot\|^2],$$

therefore

$$\frac{l_g}{2} E[\|\cdot\|^2] - E[g] \leq \frac{l_g}{2} E[\|\cdot\|^2].$$

Approximation of differentiable convex functions

This shows that $E[g] \geq 0$, and the proof be completed. ■

According to Jensen inequality and Theorem 5.4.5, the following result has been derived.

Corollary 5.4.6 [7] Let B_n is barycentric approximation (5.8). Then for all $f \in C^{1,1}(P)$ with Lipschitz constant l_f we have

$$|f(x) - B_n[f](x)| \leq \frac{l_f}{2} (B_n[\|\cdot\|^2](x) - \|x\|^2). \quad (5.14)$$

Equality is attained for all functions of the form (5.13).

According to (5.14), error estimate of convex function f in $C^{1,1}(P)$ is as follows.

$$0 \leq E_n[f](x) \leq \frac{l_f}{2} E_n[\|\cdot\|^2](x), \quad (x \in P). \quad (5.15)$$

For deriving a practical error estimate which led to computationally attractive barycentric approximation, in [7], it was derived a novel upper bound which is somewhat poorer than (5.15) and is formulated in terms of the smallest enclosing ball $\text{SEB}(P)$ containing P

$$\text{SEB}(P) := \{x \in P \subset \mathbb{R}^d : \|x - \mathbf{c}^{\text{seb}}\| \leq \mathbf{r}^{\text{seb}}\}.$$

Actually by this formulation, the error will be as follows.

$$0 \leq E_n[f](x) \leq \frac{l_f}{2} ((\mathbf{r}^{\text{seb}})^2 - \|x - \mathbf{c}^{\text{seb}}\|^2), \quad (5.16)$$

for $x \in P$ and convex function $f \in C^{1,1}(P)$.

5.5 On a generalization of an approximation operator defined by A. Guessab

As mentioned at the beginning of this chapter, in this section we prove the numerical results of the previous section in a completely different way for the operator version.

Suppose $\{e_i\}_{i \in I}$ is the orthonormal basis for the complex Hilbert space \mathcal{H} , we say $A \in \mathbb{B}(\mathcal{H})$ is a trace class operator if

$$\|A\|_1 := \sum_{i \in I} \langle |A| e_i, e_i \rangle < \infty.$$

The definition of $\|A\|_1$ does not depend on the choice of the orthonormal basis $\{e_i\}_{i \in I}$.

We denote by $S_1(\mathcal{H})$ the set of all trace class operators in $\mathbb{B}(\mathcal{H})$.

Theorem 5.5.1 The following properties hold.

5.5. On a generalization of an approximation operator defined by A. Guessab

- For every $A \in S_1(\mathcal{H})$

$$\|A\|_1 = \|A^*\|_1.$$

- $S_1(\mathcal{H})$ is an ideal in $\mathbb{B}(\mathcal{H})$, i.e.,

$$\mathbb{B}(\mathcal{H})S_1(\mathcal{H})\mathbb{B}(\mathcal{H}) \subseteq S_1(\mathcal{H}).$$

- $(S_1(\mathcal{H}), \|\cdot\|_1)$ is a Banach space.
- The following isometry property holds.

$$S_1(\mathcal{H})^* \cong \mathbb{B}(\mathcal{H}),$$

where $S_1(\mathcal{H})^*$ is the dual space of $S_1(\mathcal{H})$.

For function f on \mathbb{R} the following equivalence is well-known:

- (a) There exists a positive number C such that

$$\|f(A) - f(B)\| \leq C\|A - B\|,$$

for all bounded self-adjoint operators A and B ;

- (b) There exists a positive number C such that

$$\|f(A) - f(B)\|_1 \leq C\|A - B\|_1,$$

for all bounded self-adjoint operators A and B with $A - B$ belongs to trace class $S_1(\mathcal{H})$.

Note that the minimal value of the constant C is the same in (a) and (b). For more information see [13].

Functions satisfying (a) are called operator Lipschitz functions. We denote by $OL(\mathbb{R})$ the space of operator Lipschitz function on \mathbb{R} . For $f \in OL(\mathbb{R})$, we define $\|f\|_{OL(\mathbb{R})}$ as following:

$$\begin{aligned} \|f\|_{OL(\mathbb{R})} &= \sup \left\{ \frac{\|f(A) - f(B)\|}{\|A - B\|} : A \text{ and } B \text{ are self-adjoint, } A - B \text{ is bounded} \right\} \\ &= \sup \left\{ \frac{\|f(A) - f(B)\|_1}{\|A - B\|_1} : A \text{ and } B \text{ are self-adjoint, } A - B \in S_1 \right\}. \end{aligned}$$

We consider the notation $OC^{1,1}(\mathcal{I})$ for the subclass of all functions which are operator continuously differentiable on an interval \mathcal{I} with operator Lipschitz continuous gradients, i.e., there exist constant $\|f\|_{OL}$ such that

$$\|\nabla f(B) - \nabla f(A)\| \leq \|f\|_{OL}\|B - A\|,$$

Approximation of differentiable convex functions

for all $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} .

Double Operator Integral (DOI) is an useful tool in many problems of spectral theory of self-adjoint and unitary operators in the Hilbert space. To prove the main results of this section this tool is required.

In formal expression, DOI are objects of the form

$$\Phi = \int_{\chi} \int_{\Upsilon} \phi(\lambda, \mu) dE(\lambda) T dF(\mu),$$

where $(\chi, E(\cdot))$ and $(\Upsilon, F(\cdot))$ are two spaces with spectral measure. The values of the measure $E(\cdot)$ are orthogonal projections in a separable Hilbert space \mathfrak{H} , and similiar for the measure $F(\cdot)$ in the Hilbert space \mathfrak{B} . The scalar-valued function $\phi(\lambda, \mu)$ (the symbol of the DOI) is defined on $\chi \times \Upsilon$. Finally, T is a linear bounded operator acting from \mathfrak{B} to \mathfrak{H} (notation $T \in B(\mathfrak{B}, \mathfrak{H})$). It is clear that under any reasonable definition, the result Φ of integration is also an operator acting from \mathfrak{B} to \mathfrak{H} . If

$$\int_{\chi} \int_{\Upsilon} \phi(\lambda, \mu) dE(\lambda) T dF(\mu) \in S_1$$

for every $T \in S_1(\mathcal{H})$, we say that ϕ is a Schur multiplier of S_1 associated with the spectral measures E and F . We denote by $\mathfrak{m}(E, F)$ the space of all Schur multipliers of S_1 with respect to E and F . The norm $\|\phi\|_{\mathfrak{m}(E, F)}$ of ϕ in the space $\mathfrak{m}(E, F)$ is, by definition, the norm of the linear transformer

$$T \rightarrow \int_{\chi} \int_{\Upsilon} \phi(\lambda, \mu) dE(\lambda) T dF(\mu)$$

on the class $S_1(\mathcal{H})$.

If $\phi \in \mathfrak{m}(E, F)$, one can define by duality Double Operator Integrals of the form

$$\int_{\chi} \int_{\Upsilon} \phi(\lambda, \mu) dE(\lambda) T dF(\mu),$$

for an arbitrary bounded linear operator T . For more information we can refer the reader to the fruitful and valuable papers [2, 11].

We also remind here that according to [1, Section 3.3], we know that a continuous function f on \mathbb{R} is operator Lipschitz if and only if it is differentiable every where and the divided difference

$$\mathfrak{D}f(x, y) := \frac{f(x) - f(y)}{x - y}, \quad x, y \in \mathbb{R},$$

be a Schur multiplier.

Consider $x \in \mathcal{H}$, $\|x\| = 1$ and arbitrary self-adjoint operators A and B with spectra in \mathcal{I} . For $t \in [0, 1]$, define the function $\varphi_{x, A, B} : [0, 1] \rightarrow \mathbb{R}$ as $\varphi_{x, A, B}(t) = \langle f((1-t)A + tB)x, x \rangle$. In [4], Dragomir proved that if f be an operator convex function, then $\varphi_{x, A, B}$ is a convex function:

5.5. On a generalization of an approximation operator defined by A. Guessab

Since f is operator convex, then for any $t_1, t_2 \in [0, 1]$ and $\lambda \in [0, 1]$, we have

$$\begin{aligned}
 & \varphi_{x,A,B}(\lambda t_1 + (1 - \lambda) t_2) \\
 &= \langle f((1 - (\lambda t_1 + (1 - \lambda) t_2))A + (\lambda t_1 + (1 - \lambda) t_2)B)x, x \rangle \\
 &= \langle f(\lambda[(1 - t_1)A + t_1B] + (1 - \lambda)[(1 - t_2)A + t_2B])x, x \rangle \\
 &\leq \lambda \langle f([(1 - t_1)A + t_1B])x, x \rangle + (1 - \lambda) \langle f([(1 - t_2)A + t_2B])x, x \rangle \\
 &= \lambda \varphi_{x,A,B}(t_1) + (1 - \lambda) \varphi_{x,A,B}(t_2),
 \end{aligned}$$

which yields the convexity of function $\varphi_{x,A,B}$ on $[0, 1]$.

Now, suppose that $f \in OC^{1,1}(\mathcal{A})$, we show that $\varphi \in C^{1,1}([0, 1])$.

By considering DOI with the symbol

$$\psi_f(x, y) = \frac{f(x) - f(y)}{x - y}$$

as following:

$$T_{\psi_f}(A - B) = \int_{\sigma(A)} \int_{\sigma(B)} \psi_f(x, y) dE^A(x)(A - B) dE^B(y),$$

Daletskii-Krien formula [3], yields $T_{\psi_f}(A - B) = f(A) - f(B)$. So, the property of being operator gradient Lipschitz is equivalent to the boundedness of the corresponding DOI on $\mathbb{B}(\mathcal{H})$. By duality, this boundedness implies the boundedness of the same DOI on the predual, the trace class $S_1(\mathcal{H})$. As we are concerning the behavior of the DOI on rank one operators, the boundedness of DOI on the trace class gets the result.

Now, by utilizing Theorem 5.4.5 for $\varphi_{x,A,B}$ instead of f and g , we can extend it to the operator version as following:

Theorem 5.5.2 [15] Let $T : C^1(\mathcal{A}) \rightarrow C(\mathcal{A})$ be a linear operator. The following statements are equivalent:

(i) For every operator convex function $g \in OC^{1,1}(\mathcal{A})$, we have

$$g(A) \leq T[g](A), \quad (\text{for } A \in \mathbb{B}_{sa}(\mathcal{A}) \text{ with spectra in } \mathcal{A}).$$

(ii) For every $f \in OC^{1,1}(\mathcal{A})$, with a Lipschitz constant $\|f\|_{OL}$, we have

$$\|f(A) - T[f](A)\| \leq \frac{\|f\|_{OL}}{2} (T[(\cdot)^2](x) - x^2), \quad (x \in [0, 1])$$

Equality occurs for all functions in the form of

$$f(A) = (a(x) + cx^2)I, \tag{5.17}$$

where $c \in \mathbb{R}$ and $a(\cdot)$ is any affine function.

Approximation of differentiable convex functions

Proof Suppose that statement (i) holds, so for $\mathcal{I} = [0, 1]$ and all operator convex functions $g \in OC^{1,1}([0, 1])$ we have

$$g(A) \leq T[g](A), \quad (\text{for } A \in \mathbb{B}_{sa}(\mathcal{H}) \text{ with spectra in } \mathcal{I}).$$

As a result

$$g(xI) \leq T[g](xI), \quad (x \in \sigma(A)),$$

for all operator convex functions $g \in OC^{1,1}([0, 1])$. So, it is also valid for all convex functions $g \in C^{1,1}([0, 1])$.

Now by Theorem 5.4.5, we have

$$|f(x) - T[f](x)| \leq \frac{l_f}{2} (T[(\cdot)^2](x) - x^2),$$

for all $f \in C^{1,1}([0, 1])$. Since for all $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} , and $f \in OC^{1,1}(\mathcal{I})$ the function $\varphi := \varphi_{x,A,B} \in C^{1,1}([0, 1])$, we have

$$|\varphi(t) - T[\varphi](t)| \leq \frac{\|f\|_{OL}}{2} (T[(\cdot)^2](t) - t^2), \quad (t \in [0, 1]).$$

By definition of φ , we get

$$\begin{aligned} & |\langle f((1-t)A + tB)x, x \rangle - T[\langle f((1-t)A + tB)x, x \rangle]| \\ & \leq \frac{\|f\|_{OL}}{2} (T[(\cdot)^2](t) - t^2), \quad (t \in [0, 1]), \end{aligned}$$

by the continuity of T , it follows

$$\begin{aligned} & |\langle (f((1-t)A + tB) - T[f((1-t)A + tB)])x, x \rangle| \\ & \leq \frac{\|f\|_{OL}}{2} (T[(\cdot)^2](t) - t^2), \quad (t \in [0, 1]). \end{aligned}$$

Now by taking supremum over $x \in \mathcal{H}$ with $\|x\| = 1$, and by considering $A := (1-t)A + tB$ we get

$$\|f(A) - T[f](A)\| \leq \frac{\|f\|_{OL}}{2} (T[(\cdot)^2](t) - t^2), \quad (t \in [0, 1]),$$

which yields the error estimate in statement (ii). For the equality case, from (??) we know that for

$$\varphi(t) = a(t) + ct^2, \quad t \in [0, 1],$$

equality in Theorem (5.4.5) part (ii) will also be valid. Hence, by definition of φ we have

$$\begin{aligned} \langle f((1-t)A + tB)x, x \rangle & = a(t) + ct^2 \\ & = \langle a(t) + ct^2x, x \rangle, \quad (\|x\| = 1). \end{aligned}$$

5.5. On a generalization of an approximation operator defined by A. Guessab

Now, by letting $A := (1 - t)A + tB$ we get (5.25).

Conversely, suppose that (ii) holds. So it is also valid for $\mathcal{I} = [0, 1]$ and all $f \in C^{1,1}([0, 1])$, as follows:

$$|f(x) - T[f](x)| \leq \frac{l_f}{2} (T[(\cdot)^2](x) - x^2), \quad x \in [0, 1].$$

By Theorem (5.4.5), we have

$$g(x) \leq T[g](x),$$

for all convex function $g \in C^{1,1}([0, 1])$. By replacing $\varphi := \varphi_{x,A,B}$ instead of g , we obtain

$$\varphi(t) \leq T[\varphi](t), \quad (t \in [0, 1]),$$

which results

$$\langle f((1 - t)A + tB)x, x \rangle \leq T[\langle f((1 - t)A + tB)x, x \rangle],$$

where $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} . By letting $A := (1 - t)A + tB$

$$\langle f(A)x, x \rangle \leq T[\langle f(A)x, x \rangle],$$

since T is linear, we have

$$\langle f(A)x, x \rangle \leq \langle T[f(A)]x, x \rangle,$$

which yields that for all operator convex functions in $OC^{1,1}(\mathcal{I})$

$$f(A) \leq T[f](A),$$

and this completes the proof. ■

Corollary 5.5.3 Let B_n be the barycentric approximation. Then for every function $f \in OC^{1,1}(\mathcal{I})$ with a Lipschitz constant $\|f\|_{OL}$, we have

$$\|f(A) - B_n[f](A)\| \leq \frac{\|f\|_{OL}}{2} (B_n[(\cdot)^2](x) - x^2),$$

for $A \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{I} and $x \in [0, 1]$.

Proof Since $f \in OC^{1,1}(\mathcal{I})$, we can consider inequality (5.14) for function $\varphi_{x,A,B}$ on $[0, 1]$ as following:

$$|\varphi_{x,A,B}(t) - B_n[\varphi_{x,A,B}](t)| \leq \frac{\|f\|_{OL}}{2} (B_n[(\cdot)^2](t) - t^2), \quad t \in [0, 1].$$

By definition of $\varphi_{x,A,B}$ we have

$$\begin{aligned} & |\langle f((1 - t)A + tB)x, x \rangle - \sum_{i=0}^n \lambda_i(t) \langle f((1 - t_i)A + t_i B)x, x \rangle| \\ & \leq \frac{\|f\|_{OL}}{2} \left(\sum_{i=0}^n \lambda_i(t) t_i^2 - t^2 \right), \end{aligned} \tag{5.18}$$

Approximation of differentiable convex functions

for $t, t_i \in [0, 1]$, and λ_i as barycentric coordinates.

Inequality (5.18) yields

$$\begin{aligned} & \langle (f((1-t)A + tB) - \sum_{i=0}^n \lambda_i(t) f((1-t_i)A + t_i B))x, x \rangle \\ & \leq \frac{\|f\|_{OL}}{2} \left(\sum_{i=0}^n \lambda_i(t) t_i^2 - t^2 \right), \end{aligned}$$

by letting $A := (1-t)A + tB$ and $A_i := (1-t_i)A + t_i B$, and taking supremum over $x \in \mathcal{H}$ with $\|x\| = 1$, we get

$$\|f(A) - \sum_{i=0}^n \lambda_i(t) f(A_i)\| \leq \frac{\|f\|_{OL}}{2} \left(\sum_{i=0}^n \lambda_i(t) t_i^2 - t^2 \right),$$

which is equivalent to

$$\|f(A) - B_n[f](A)\| \leq \frac{\|f\|_{OL}}{2} (B_n[(\cdot)^2](t) - t^2).$$

■

Now, by considering the inequality (5.15) for function $\varphi_{x,A,B}$ with the operator convex function f in $OC^{1,1}(\mathcal{I})$ we deduce the following result.

Theorem 5.5.4 [15] For every operator convex function $f \in OC^{1,1}(\mathcal{I})$ with Lipschitz constant $\|f\|_{OL}$, we have

$$0 \leq E_n[f](A) \leq \left(\frac{\|f\|_{OL}}{2} E_n[(\cdot)^2](x) \right) I, \quad (5.19)$$

for $A \in \mathbb{B}_{sa}(\mathcal{H})$ with $\sigma(A) \subset \mathcal{I}$ and $x \in [0, 1]$.

Proof As the proof of previous result, by using inequality (5.15) for $\varphi_{x,A,B}$ we obtain

$$0 \leq E_n[\varphi_{x,A,B}](t) \leq \frac{\|f\|_{OL}}{2} E_n[(\cdot)^2](t), \quad t \in [0, 1].$$

By definition of $E_n[\cdot]$, we have

$$0 \leq \sum_{i=0}^n \lambda_i(t) \varphi_{x,A,B}(t_i) - \varphi_{x,A,B}(t) \leq \frac{\|f\|_{OL}}{2} \left[\sum_{i=0}^n \lambda_i(t) t_i^2 - t^2 \right], \quad t \in [0, 1],$$

which results

$$\begin{aligned} & 0 \leq \sum_{i=0}^n \lambda_i(t) \langle f((1-t_i)A + t_i B)x, x \rangle - \langle f((1-t)A + tB)x, x \rangle \\ & \leq \frac{\|f\|_{OL}}{2} \left[\sum_{i=0}^n \lambda_i(t) t_i^2 - t^2 \right], \end{aligned} \quad (5.20)$$

5.5. On a generalization of an approximation operator defined by A. Guessab

by letting $A := (1 - t)A + tB$ and $A_i := (1 - t_i)A + t_iB$, we obtain

$$0 \leq E_n[f](A) \leq \left(\frac{\|f\|_{OL}}{2} E_n[(\cdot)^2](t)\right)I, \quad t \in [0, 1],$$

which yields the result. ■

Utilizing the continuous functional calculus we can state the operator version of (5.16) as follows.

Theorem 5.5.5 For every operator convex function $f \in OC^{1,1}(\mathcal{J})$ with a constant $\|f\|_{OL}$, we have

$$0 \leq E_n[f](A) \leq \frac{\|f\|_{OL}}{2} \left(\frac{1}{4} - \left|x - \frac{1}{2}\right|^2\right)I, \quad (5.21)$$

for $A \in \mathbb{B}_{sa}(\mathcal{H})$ with $\sigma(A) \subset \mathcal{J}$ and $x \in [0, 1]$.

Proof let f be an operator convex function belongs to $OC^{1,1}(\mathcal{J})$, by replacing $\varphi_{x,A,B}$ in (5.16), we have

$$0 \leq E_n[\varphi](t) \leq \frac{\|f\|_{OL}}{2} \left(\frac{1}{4} - \left|t - \frac{1}{2}\right|^2\right)I, \quad t \in [0, 1],$$

equivalently

$$0 \leq \sum_{i=0}^n \lambda_i(t) \varphi(t_i) - \varphi(t) \leq \frac{\|f\|_{OL}}{2} \left(\frac{1}{4} - \left|t - \frac{1}{2}\right|^2\right)I, \quad t \in [0, 1]. \quad (5.22)$$

Now by definition of φ and using the methods which used in the proof of the previous theorem, we get the result. ■

It is notable that the bound which is represented in (5.21) for error of barycentric approximation does not depend on the choice of barycentric coordinates.

Example 5.5.6 Let $f \in OC^{1,1}(\mathcal{J})$ be an operator convex function, A and B be two self-adjoint operators with spectra in \mathcal{J} , and consider $t_0 = 0$, $t_1 = \frac{1}{4}$, $t_2 = \frac{3}{4}$ and $t_3 = 1$ as a partition of interval $[0, 1]$. For $t = \frac{1}{2}$, we obtain barycentric coordinates as following:

$$\lambda_0\left(\frac{1}{2}\right) = 0, \quad \lambda_1\left(\frac{1}{2}\right) = \frac{1}{2}, \quad \lambda_2\left(\frac{1}{2}\right) = \frac{1}{2}, \quad \lambda_3\left(\frac{1}{2}\right) = 0.$$

By replacing the quantities of t , t_i and $\lambda_i(t)$ ($i = 0, 1, 2, 3$) in (5.20), and by easy computations we obtain

$$0 \leq \frac{1}{2} \left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right) \right] - f\left(\frac{A+B}{2}\right) \leq \frac{\|f\|_{OL}}{32} I. \quad (5.23)$$

Approximation of differentiable convex functions

In the other hand, according to (5.22), by using the same barycentric coordinates as above, we get

$$\begin{aligned} 0 &\leq \frac{1}{2} \left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right) \right] - f\left(\frac{A+B}{2}\right) \\ &\leq \frac{\|f\|_{OL}}{2} \left(\frac{1}{4} - \left|\frac{1}{2} - \frac{1}{2}\right|^2\right) I = \frac{\|f\|_{OL}}{8} I, \end{aligned} \quad (5.24)$$

as we mentioned before, this estimation is not in terms of barycentric coordinates, but it is poorer than upper bound in (5.23), ($\frac{1}{8} > \frac{1}{32}$).

In [8, Section 7.1] has been proved a theorem which is a characterization of any integration formula in \mathbb{R}^d , which admits an upper or a lower Hermite-Hadamard inequality. For $d = 1$ we can state it as follows.

Theorem 5.5.7 Let $f : [a, b] \rightarrow \mathbb{R}$ is a convex function, α_i ($i = 0, 1, \dots, n$) are positive real numbers and x_i as a partition of $[a, b]$. Define the integration formula via

$$\int_a^b f(x) dx = \sum_{i=0}^n \alpha_i f(x_i) + E(f),$$

and let $\sigma \in \{-1, 1\}$.

(i) For every operator convex function $g \in OC^{1,1}(\mathcal{S})$, we have

$$\sigma E[g] \geq 0,$$

(ii) For every $f \in OC^{1,1}(\mathcal{S})$, with a Lipschitz constant l_f , we have

$$|E[f]| \leq \sigma E[(\cdot)^2] \frac{l_f}{2}.$$

Equality occurs for all functions in the form of

$$f(x) = a(x) + cx^2, \quad (5.25)$$

where $c \in \mathbb{R}$ and $a(\cdot)$ is any affine function.

According to above theorem and Theorem 5.5.2 we can state:

Theorem 5.5.8 Let A and B be two self-adjoint operators with spectra in \mathcal{S} . Define the integration formula via

$$\int_0^1 f((1-t)A + tB) dt = \sum_{i=0}^n \alpha_i f((1-t_i)A + t_iB) + E[f],$$

5.5. On a generalization of an approximation operator defined by A. Guessab

where $\alpha_i, i = 0, 1, \dots, n$ are positive real numbers and for $i = 0, 1, \dots, n, t_i \in [0, 1]$ is a partition of interval $[0, 1]$, and let $\sigma \in \{-1, 1\}$. Then, the following two statements are equivalent:

(i) For every operator convex function $g \in OC^{1,1}(\mathcal{S})$, we have

$$\sigma E[g] \geq 0.$$

(ii) For every $f \in OC^{1,1}(\mathcal{S})$, we have

$$\|E[f]\| \leq \sigma E[(\cdot)^2] \frac{\|f\|_{OL}}{2}.$$

Equality valids for all functions in the form of

$$f((1-t)A + tB) = (a(t) + ct^2)I,$$

where $c \in \mathbb{R}$ and $a(\cdot)$ is any affine function.

According to Theorem 2.3, we have the refinement of Hermite-Hadamard inequality for operator convex function $f : \mathcal{S} \rightarrow \mathbb{R}$ as following:

$$\begin{aligned} f\left(\frac{A+B}{2}\right) &\leq \frac{1}{2} \left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right) \right] \\ &\leq \int_0^1 f((1-t)A + tB) dt \\ &\leq \frac{f(A) + f(B)}{2}, \end{aligned} \tag{5.26}$$

where $A, B \in \mathbb{B}_{sa}(\mathcal{H})$ with spectra in \mathcal{S} .

Now consider $f \in OC^{1,1}(\mathcal{S})$ be an operator convex function, so we have

$$\int_0^1 f((1-t)A + tB) dt = \frac{1}{2} \left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right) \right] + E[f],$$

by inequalities (5.26) we know that $E[f] \geq 0$, so by Theorem (5.5.8) statement (i) and Example (5.5.6), we get

$$0 \leq \int_0^1 f((1-t)A + tB) dt - \frac{1}{2} \left[f\left(\frac{3A+B}{4}\right) + f\left(\frac{A+3B}{4}\right) \right] \leq \frac{\|f\|_{OL}}{32} I. \tag{5.27}$$

Namely, by inequalities (5.27) we obtained a bound according to barycentric coordinates, for the difference between the left and right hand side of the second inequality in (5.26).

As an example, we can refer the function $f(t) = t^r$ on $(0, \infty)$ which is operator convex function if either $1 \leq r \leq 2$ or $-1 \leq r \leq 0$. This function belongs to $OC^{1,1}(\mathbb{R})$, so, for positive operators A

Approximation of differentiable convex functions

and B with spectra in $\mathcal{I} = [0, 1]$, according to inequalities (5.27) we have

$$\begin{aligned} o &\leq \int_0^1 ((1-t)A + tB)^r dt - \frac{1}{2} \left[\left(\frac{3A+B}{4}\right)^r + \left(\frac{A+3B}{4}\right)^r \right] \\ &\leq \frac{r(r-1)}{32} I. \end{aligned}$$

Bibliography

- [1] A.B. Aleksandrov, V.V. Peller, *Operator Lipschitz functions*, *Russian Mathematical Survey* 71 (2016) 605-624.
- [2] M.SH. Birman, M.Z. Solomyak, *Lectures on double operator integrals*, *Institut Mittag-Leffler*, rep. 20, (2002/2003).
- [3] 18. Yu.L. Daleetskii, S.G. Krein, *Integration and differentiation of functions of Hermitian operators and applications to the theory of perturbations*, (Russian) *Voronezh. Gos. Univ. Trudy Sem. Funkcional. Anal.* 11 (1956) 81-105.
- [4] S. S. Dragomir, *Hermite-Hadamard's type inequalities for operator convex functions*, *Applied Mathematics and Computation* 218 (2011) 766-772.
- [5] J. Ferrera, *An Introduction to Nonsmooth Analysis*, *Academic Press*, (2014).
- [6] A. Guessab, *Approximation of differentiable convex functions on arbitrary convex polytopes*, *Applied Mathematics and Computation* 240 (2014) 326-338.
- [7] A. Guessab, *Generalized barycentric coordinates and approximations of convex functions on arbitrary convex polytopes*, *Comput Math Appl.* 66(6) (2013) 11201136.
- [8] A. Guessab, B. Semisalov, *Numerical integration using integrals over hyperplane sections of simplices in a triangulation of a polytope*, *BIT Numer Math*, 58 (2018) 613-660.
- [9] B. Grünbaum, *Convex polytopes, second edition*, *Springer-verlag, New york, Inc.* (2003).
- [10] J. A. Kalman, *Continuity and convexity of projections and barycentric coordinates in convex polyhedra*, *Pacific J. of Math.* 2 (1961) 1017-1022.
- [11] E. Kissin, D. Potapov, V. Shulman, F. Sukochev, *Operator smoothness in Schatten norms for functions of several variables: Lipschitz conditions, differentiability and unbounded derivations*, *Proc. Lond. Math. Soc.* (3) 105 (2012) 661-702.

Bibliography

- [12] A.F. Möbius, *barycentrische Calcul*, Johann Ambrosius Barth, Leipzig, (1827).
- [13] V.V. Peller, *The Lifshits-Krein trace formula and operator Lipschitz functions*, *Proc. Amer. Math. Soc.*144(12) (2016) 5207-5215.
- [14] R.T. Rockafeller, *Convex Analysis*, Princeton University Press, (1970).
- [15] A. Taghavi, T. Azimi Roushan, *On a generalization of an approximation operator defined by A. Guessab*, *Linear and Multilinear Algebra* 67(11) (2019) 1-13.
- [16] G.M. Ziegler, *Lectures on Polytopes*, Springer Verlag, New York, (1995).

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

This chapter is based on our paper [9], which constitutes the progression of recent works [6, 7, 8, 11] where focused on the study of some classes of multi-dimensional numerical integration in the context of the classical notion of convexity. Here, our objective is to extend the results given there for strongly convex functions. Indeed, we study and construct a family of multi-dimensional numerical integration formulas (cubature formulas), which approximate all strongly convex functions from above. We call them strongly negative definite cubature formulas (or for brevity snd-formulas). We attempt to quantify their sharp approximation errors when using continuously differentiable functions with Lipschitz continuous gradients. We show that the error estimates based on such cubature formulas are always controlled by the Lipschitz constants of the gradients and the error associated with using the quadratic function. Moreover, assuming the integrand is itself strongly convex, we establish sharp upper as well as lower refined bounds for their error estimates. Based on the concepts of barycentric coordinates with respect to an arbitrary polytope P , we provide a necessary and sufficient condition for the existence of a class of snd-formulas on P : It consists of checking that such coordinates exist on P . Then, the Delaunay triangulation is used as a convenient partition of the integration domain for constructing the best piecewise snd-formulas in L^1 metric. Finally, we present numerical examples illustrating the proposed method.

6.1 Introduction, motivation and terminology

To describe our problem of integration from a numerical standpoint more precisely, let $\Omega \subset \mathbb{R}^d$ be a nonempty compact convex set and let $f : \Omega \rightarrow \mathbb{R}$ be a given function. We sometimes know beforehand that the function f satisfies various known structural and regularity properties. For example, it may be known that f has some additional kind of convexity, therefore we would wish to use this information in order to get most appropriate methods for numerical integra-

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

tion of f . To get a better approximation of the integral of our function, we try to approximate it using cubature formulas, which approximate the integral of all strongly convex functions from above. The strongly convex functions are widely applied in economic theory (see [24]), and are also central to optimization theory (see [18]). Indeed, in the framework of function minimization, this convexity notion has important and well-known implications. As we will see, the key advantage of using cubature formulas of such kind is that their associated approximation errors can always be controlled by the error associated with using the quadratic function. Hence, if we want a more accurate approximation of the integral of our function, we need to find a better approximation of the integral of the quadratic function.

To appreciate the problem more clearly, let us start by describing briefly a specific one-dimensional example, since its simplicity helps us better understand all the necessary steps through very simple explicit computations. Assume that μ is a fixed nonnegative real number. In one-dimensional numerical integration, say on an interval $[a, b]$, a simple way of approximating the integral of a given real μ -strongly convex function $f : [a, b] \rightarrow \mathbb{R}$ is first to choose a partition $P := \{x_0, x_1, \dots, x_n\}$ of the interval $[a, b]$, such that $a = x_0 < x_1 < \dots < x_n = b$, and then to apply the classical local trapezoidal quadrature rule $T_i(f) = \frac{f(x_{i-1}) + f(x_i)}{2}$ on each subinterval $I_i := [x_{i-1}, x_i]$, $i = 1, 2, \dots, n$, and summing up the results. Among its many important properties, this rule satisfies the well-known Hermite-Hadamard inequality, which ensures an upper estimate for the exact value of the integral of any convex function:

$$\frac{1}{x_i - x_{i-1}} \int_{x_{i-1}}^{x_i} f(t) dt \leq T_i(f), \quad (i = 1, 2, \dots, n), \quad (6.1)$$

where the sign of equality being achieved if f is an affine function. Recall that the local trapezoidal rule $T_i(f)$ could be obtained by integrating the barycentric approximation operator:

$$B_i[f](x) := \lambda_{i-1}(x)f(x_{i-1}) + \lambda_i(x)f(x_i), \quad (x \in I_i),$$

where $\lambda_{i-1}(x)$ and $\lambda_i(x)$ are the barycentric coordinates of x with respect to I_i , that are defined as:

$$\lambda_{i-1}(x) := \frac{x - x_i}{x_{i-1} - x_i}, \quad \lambda_i(x) := \frac{x - x_{i-1}}{x_i - x_{i-1}}, \quad (x \in I_i).$$

Observe that B_i is a first order barycentric polynomial interpolating f at two points, x_{i-1} and x_i and that the weights λ_{i-1}, λ_i can be expressed as

$$\lambda_{i-1}(x) = \frac{1}{\text{length}(I_i)} \begin{vmatrix} 1 & x \\ 1 & x_i \end{vmatrix}, \quad \lambda_i(x) = \frac{1}{\text{length}(I_i)} \begin{vmatrix} 1 & x_{i-1} \\ 1 & x \end{vmatrix}.$$

Rearranging terms, it is clear that these weights are nonnegative on I_i and moreover they satisfy

$$\lambda_{i-1}(x) + \lambda_i(x) = 1, \quad x = \lambda_{i-1}(x)x_{i-1} + \lambda_i(x)x_i, \quad (x \in I_i). \quad (6.2)$$

The trapezoidal rule is the simplest, most well-known and widely used quadrature rule. The reason for this popularity lies in the large number of useful theoretical and computational properties of this rule. It actually served as basic ingredients for constructing more accurate and adaptive formulas. For this reason, this rule together with its fundamental inequality (6.1) has been an effective starting point for several subsequent investigations, see [4, 10]. Furthermore, in the local error analysis of the rule $T_i(f)$:

$$ET_i(f) := T_i(f) - \frac{1}{x_i - x_{i-1}} \int_{x_{i-1}}^{x_i} f(t) dt,$$

estimate of (6.1), is a very useful tool. Indeed, let $(\cdot)^2$ denote the square function $t \rightarrow t^2$, and assume that the first derivative of f is a Lipschitz function with a Lipschitz constant $L(f')$ in $[a, b]$ (or $f \in C^{1,1}[a, b]$), then Hermite-Hadamard inequality implies the following upper local estimation:

$$|ET_i(f)| \leq \frac{ET_i((\cdot)^2)}{2} L(f') \quad (6.3)$$

$$= \frac{T_i\left(\left(\cdot - \frac{x_{i-1} + x_i}{2}\right)^2\right)}{3} L(f') \quad (6.4)$$

$$= \frac{(x_i - x_{i-1})^2}{12} L(f'), \quad (6.5)$$

where equality is attained for all quadratic functions. In addition, if f is μ -strongly convex then the following lower local estimation also holds for all $i = 1, 2, \dots, n$,

$$ET_i(f) \geq \frac{ET_i((\cdot)^2)}{2} \mu \quad (6.6)$$

$$= \frac{T_i\left(\left(\cdot - \frac{x_{i-1} + x_i}{2}\right)^2\right)}{3} \mu \quad (6.7)$$

$$= \frac{(x_i - x_{i-1})^2}{12} \mu. \quad (6.8)$$

We did not find any reference to such result. However, the above mentioned estimates can be derived as an immediate consequence of our multivariate general results, see Remark 6.4.5. Estimates (6.3) and (6.6) say that for the trapezoidal rule, we can always control its approximation error by the Lipschitz constants of the first derivative, the parameter (of the strong convexity) and the error associated with using the quadratic function. It should also be noted

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

that equalities in (6.3) and (6.6) are satisfied for all μ -strongly convex functions of the form

$$f(x) = a(x) + \frac{\mu}{2}x^2, \quad (6.9)$$

where $a(\cdot)$ is any affine function. Therefore, in this sense, the error estimates (6.3) and (6.6) are sharp for the class of μ -strongly convex functions having Lipschitz continuous first derivatives. This provides the starting point of this chapter study. Indeed, the contributions of this chapter are two-fold: first, we would like to consider the general multivariate variable case. More precisely, we deal with the problem of approximation of the integral of multivariate functions by snd-formulas, that is, those which approximate from above all strongly convex functions with Lipschitz-continuous gradients. Geometrically, if a function f belongs to such class, then its gradient ∇f cannot change too quickly and it cannot change too slowly either. Functions satisfying these conditions are widely used in the optimization literature, we refer to Nesterov's book [18].

Secondly, under the assumption of strong convexity and the standard Lipschitz continuity of the gradient, we prove some results that pertain to sharp estimates of the error arising from such approximations.

Hence, the questions that arise, as a natural consequence of the estimates (6.3) and (6.6), are the following:

- Can we extend the one-dimensional approach to construct a natural multivariate version of the trapezoidal quadrature rule in any polytope?
- Can the approximation errors for such cubature formulas satisfy similar lower and upper bounds in the multi-dimensional case?

We will answer these questions positively by defining and studying a class of snd-formulas on an arbitrary polytope to approximate the integral of a function by piecewise cubature formulas. Our extensions are derived in a natural way by using the generalized barycentric coordinates, which turn out to be appropriate to the more general multivariate setting. In particular, we will show how the Delaunay triangulation can be used as a convenient partition of the integration domain for constructing the best piecewise snd-formulas in L^1 metric.

Now, we present a very useful tool in our study which is triangulations of a point set and the Delaunay triangulation.

Definition 6.1.1 (triangulation of a point set) Let S be a finite set of points in the plane. A triangulation of S is a simplicial complex \mathbb{T} such that S is the set of vertices in \mathbb{T} , and the union of all the simplices in \mathbb{T} is the convex hull of S , that is $|\mathbb{T}| = \text{conv}(S)$.

A simplicial complex is a collection of simplices that intersect only in mutual faces (The convex hull of any nonempty subset of the $n + 1$ points that define an n -simplex is called a face of the simplex). i.e., any face of a simplex from simplicial complex \mathbb{T} is also in \mathbb{T} and the intersection of any two simplices $\sigma_1, \sigma_2 \in \mathbb{T}$ is either \emptyset or a face of both σ_1 and σ_2 .

6.1. Introduction, motivation and terminology

Definition 6.1.2 (triangulation of a point set in \mathbb{R}^d) The definition (6.1.1) defines a triangulation of a set of points to be simplicial complex whose vertices are the points and whose union is the convex hull of the points. With no change, the definition holds in any finite dimension d , i.e., a simplex in \mathbb{R}^d is a d -dimensional simplex (d -simplex), which is defined by its $(d + 1)$ vertices, and a triangulation of a set of points in \mathbb{R}^d is a simplicial decomposition of the convex hull of the point set where the vertices of the triangles are contained in the point set.

Every finite point set in \mathbb{R}^d has a triangulation, see [21, Section 2.1]. One of the famous and optimal triangulation is the Delaunay triangulation which is a geometric structure that engineers have used for meshes since mesh generation was in its infancy. The Delaunay triangulation of a point set S , introduced by Boris Nikolavich Delaunay in 1934, is characterized by the empty circumdisk property: no point in S lies in the interior of any triangles circumscribing disk; see [21, Definition 1.17].

Delaunay triangulations can be generalized easily to higher dimensions (\mathbb{R}^d).

Definition 6.1.3 (The Delaunay triangulation in \mathbb{R}^d) Let S be a finite point set in \mathbb{R}^d , and let k be the dimension of its affine hull. A Delaunay triangulation of S , $\text{Del}(S)$ is a triangulation of S in which every k -simplex is Delaunay and therefore, every simplex is Delaunay. i.e., the Delaunay triangulation of a set of points in \mathbb{R}^d is defined to be the triangulation such that the circumsphere of every triangle in the triangulation contains no point from the set in its interior.

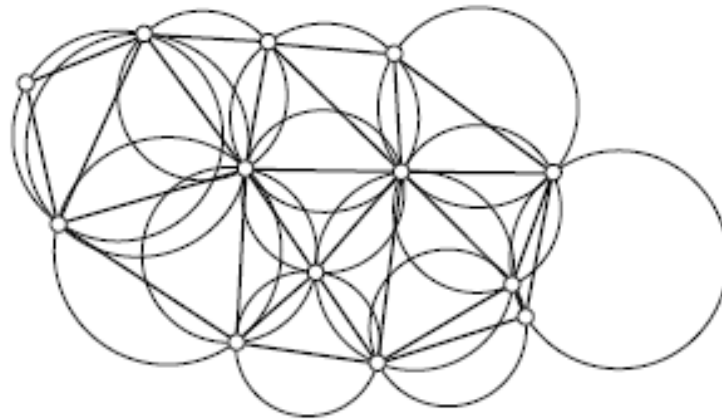


Figure 6.1 – Every triangle in a Delaunay triangulation has an empty open circumdisk.

Such a triangulation exists for every point set in \mathbb{R}^d , see [21, Section 2.2].

In the forthcoming sections of this chapter, first we establish two general characterization results (see Lemma 6.2.1 and Lemma 6.2.2). These general results provide two equivalent conditions for a linear functional to be negative in the set of convex functions. We then use them to establish a first characterization of the approximation error of our class of cubature formulas. In

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

the sequel, we provide a necessary and sufficient condition for the existence of the *snd*-formulas. It consists of checking the existence of a set of these coordinates. Section 6.4 uses the generalized barycentric coordinates to construct a multivariate version of the classical trapezoidal rule in arbitrary higher-dimensional polytopes. As a result we get explicit lower and upper bounds for the approximation error when using continuously differentiable functions with Lipschitz continuous gradients. Indeed, analogously to the one-dimensional estimates (6.3) and (6.6), we offer sharp error estimates which only depend on the parameter of the strong convexity, the Lipschitz constants of the gradients and the error associated with using the quadratic function. In the following, by using the Delaunay triangulation as a partition of a polytope, we present an explicit construction of our sharp cubature schemes. Finally, we provide a numerical example to illustrate the efficiency of this approach.

6.2 General setting

Our main results in this section first concern two characterization results of any negative linear functional in the set of convex functions, which hold in a general framework and will be repeatedly applied in the sequel.

Let Ω be a subset of \mathbb{R}^d . As usual, we mean by Ω° the interior of Ω . We say that Ω is measurable if it has a finite Lebesgue measure, which we denote by $|\Omega|$. For measurable Ω , the class $L^1(\Omega)$ comprises all Lebesgue integrable functions $f : \Omega \rightarrow \mathbb{R}$. A property holds almost everywhere (abbreviated by *a.e.*) on Ω if it holds on Ω except for a set of measure zero. Furthermore, we denote by $C(\Omega)$ the class of all real-valued continuous functions on Ω and by $C^k(\Omega)$, where $k \in \mathbb{N}$, the subclass of all functions which are k times continuously differentiable. It is convenient to agree that $C^0(\Omega) = C(\Omega)$.

Now we state a first characterization result of linear functionals, which are negative in the set of convex functions. It is shown that in order to prove such property for the given functional E , it suffices to check that E is negative in a subset of strongly convex functions with a given fixed strong convexity parameter.

Lemma 6.2.1 *Let $\Omega \subset \mathbb{R}^d$ be a compact convex set. Let μ be a positive fixed real number and let E be a linear functional defined on $C(\Omega)$. Then, the following conditions are equivalent:*

(i) *For every convex function $f \in C(\Omega)$, we have*

$$E(f) \leq 0.$$

(ii) *For every μ -strongly convex function $f \in C(\Omega)$, we have*

$$E(f) \leq 0.$$

Proof (i) implies (ii) is the trivial part of the proof. Indeed, assume that (i) holds. Let f be μ -strongly convex function. Set $g := f - \frac{\mu}{2} \|\cdot\|^2$. By definition, g is therefore convex. Hence,

applying property (i), it follows, by linearity of E

$$E(f) \leq \frac{\mu}{2} E(\|\cdot\|^2).$$

Since $\|\cdot\|^2$ is convex, then again by (i) we have $E(\|\cdot\|^2) \leq 0$. This shows that (ii) holds.

Now, assume that (ii) holds. Let ε be a positive real number and let f be a convex function. Define the function g by

$$g := f + \frac{\varepsilon}{2} \|\cdot\|^2.$$

Noting that

$$\frac{\mu}{\varepsilon} f = \frac{\mu}{\varepsilon} g - \frac{\mu}{2} \|\cdot\|^2$$

and since $\frac{\mu}{\varepsilon} f$ is convex, then by the definition of strong convexity $\frac{\mu}{\varepsilon} g$ is μ -strongly convex. Hence, by (ii) we can conclude that

$$E\left(\frac{\mu}{\varepsilon} g\right) \leq 0.$$

Thus it follows that

$$E(g) \leq 0,$$

or equivalently, by virtue of the linearity of E ,

$$E(f) \leq -\frac{\varepsilon}{2} E(\|\cdot\|^2).$$

In view of the fact that this inequality holds for all $\varepsilon > 0$, then by letting $\varepsilon \downarrow 0$, it follows that

$$E(f) \leq 0.$$

Hence, the desired statement (i) is valid and thus means that these two statements are equivalent. ■

If in addition, the functions belong to $C^{1,1}(\Omega)$, then our second characterization result is given in the following:

Lemma 6.2.2 Let $\Omega \subset \mathbb{R}^d$ be a compact convex set. Let $E : C^k(\Omega) \rightarrow \mathbb{R}$, where $k \in \{0, 1\}$, be a linear functional and let μ a positive real number. Then, the two following statements are equivalent:

(i) For every μ -strongly convex function $g \in C^{1,1}(\Omega)$, we have

$$E[g] \leq 0. \tag{6.10}$$

(ii) For every $f \in C^{1,1}(\Omega)$ with $L(\nabla f)$ -Lipschitz gradient, we have

$$|E[f]| \leq -E[\|\cdot\|^2] \cdot \frac{L(\nabla f)}{2}. \tag{6.11}$$

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

Equality is attained for all functions of the form

$$f(\mathbf{x}) := a(\mathbf{x}) + c\|\cdot\|^2, \quad (6.12)$$

where $c \in \mathbb{R}$ and $a(\cdot)$ is any affine function.

Proof First we prove (i) implies (ii). Let f be any function from $C^{1,1}(\Omega)$ with $L(\nabla f)$ -Lipschitz gradient. Define the following two functions

$$g_{\pm} := \|\cdot\|^2 \frac{L(\nabla f)}{2} \pm f.$$

Then, according to Proposition 5.4.4, we know that both of these functions belong to $C^{1,1}(\Omega)$ and are also convex. Hence, by (i) and Lemma 6.2.1, it follows that the functions g_- and g_+ satisfy

$$E[g_{\pm}] \leq 0.$$

Then, by linearity of E and a simple manipulation we find that

$$E[\|\cdot\|^2] \frac{L(\nabla f)}{2} \leq E[f] \leq -E[\|\cdot\|^2] \frac{L(\nabla f)}{2}.$$

This is equivalent to (6.11) and shows that property (ii) also holds.

Now, let us assume that (ii) holds. Then, we deduce that

$$E[\|\cdot\|^2] \leq 0, \quad (6.13)$$

Let $g \in C^{1,1}(\Omega)$ be any μ -strongly convex function and set

$$f := \frac{L(\nabla g)}{2} \|\cdot\|^2 - g.$$

Then, according to Proposition 5.4.4, we have

$$f \in C^{1,1}(\Omega) \text{ and } L(\nabla f) \leq L(\nabla g). \quad (6.14)$$

Since

$$g = \frac{L(\nabla g)}{2} \|\cdot\|^2 - f,$$

it can be written as follows

$$g = \left(\|\cdot\|^2 \frac{L(\nabla f)}{2} - f \right) + \|\cdot\|^2 \left(\frac{L(\nabla g)}{2} - \frac{L(\nabla f)}{2} \right),$$

we therefore obtain

$$E[g] = E \left[\|\cdot\|^2 \frac{L(\nabla f)}{2} - f \right] + E[\|\cdot\|^2] \left(\frac{L(\nabla g)}{2} - \frac{L(\nabla f)}{2} \right).$$

Finally, by combining (ii), (6.13) and (6.14) we can conclude that (i) is valid. For the statement on the occurrence of equality, it is enough to note that a linear functional E satisfying (6.10) for all convex functions must vanish for affine functions. ■

Lemma 6.2.2 extends the result given in Theorem 5.4.5 for convex functions to the case of strongly convex functions.

We now define our new general class of cubature formulas, which we formulate as follows:

Definition 6.2.3 Let $\Omega \subset \mathbb{R}^d$ be a compact set and let μ be a positive real number. For n points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in \Omega$, called nodes, and associated positive numbers A_1, A_2, \dots, A_n , we say that

$$\{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}, \quad (6.15)$$

defines the μ -strongly negative definite cubature formula

$$\int_{\Omega} f(\mathbf{x}) d\mathbf{x} = \sum_{i=1}^n A_i f(\mathbf{x}_i) + E[f], \quad (6.16)$$

if the approximation error E satisfies

$$E[f] \leq 0, \quad (6.17)$$

for all μ -strongly convex functions $f \in C(\Omega)$.

We say that (6.16) is a μ snd-formula for short. We also call (6.15) a μ snd-system, which is said to be of length n if the points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ are distinct. Let us mention that any μ snd-cubature formula approximates the exact value of the integral of a μ -strongly convex function from above. This means that the approximation error for such cubature formulas is negative on the set of μ -strongly convex functions.

Remark 6.2.4 Note that a μ snd-cubature formula as specified in Definition 6.2.3 is always of order two. In fact, by Lemma 6.2.2 inequality (6.11) the functional E vanishes for affine functions and so the order is at least two. However, if the order were greater than two, then (6.11) would imply that $E[f] = 0$ for all $f \in C^{1,1}(\Omega)$. Recall that, in the univariate case, a quadrature rule is snd-formula if and only if its second Peano kernel is greater than zero or less than zero, respectively; see [1, Chap.II.4] or [3, Chap. 4.3].

In the theory of inequalities, inequality (6.17), with E defined by (6.16) and valid for all μ -strongly convex functions, has also been called upper Hermite-Hadamard inequality.

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

Remark 6.2.5 Let $f \in C^{1,1}(\Omega)$ be μ -strongly convex function, then by Theorem 1.1.14 we have

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|^2 \quad (6.18)$$

$$\|\nabla f(x) - \nabla f(y)\| \leq L(\nabla f) \|x - y\|. \quad (6.19)$$

Using Cauchy-Schwarz inequality on the first term of (6.18), $\mu \|x - y\| \leq \|\nabla f(x) - \nabla f(y)\|$, then according to (6.19) we conclude that $\mu \leq L(\nabla f)$. Hence, μ should be less than or equal to $L(\nabla f)$ if the function is both μ -strongly convex and its gradient is $L(\nabla f)$ -Lipschitz continuous.

We now present a characterization of our class of cubature formulas in terms of their associated error functionals. Indeed, we show that for functions in $C^{1,1}(\Omega)$, the error estimates based on such cubature formulas are always controlled by the Lipschitz constants of the gradients, the strong convexity parameter and the error associated with using the quadratic function. This result is a direct consequence of Lemmas 6.2.1 and 6.2.2.

Theorem 6.2.6 Let $\Omega \subset \mathbb{R}^d$ be a compact convex set. A cubature formula (6.16) is μ -strongly snd-formula if and only if for all μ -strongly convex functions $f \in C^{1,1}(\Omega)$, its error functional satisfies

$$\frac{L(\nabla f)}{2} E[\|\cdot\|^2] \leq E[f] \leq \frac{\mu}{2} E[\|\cdot\|^2]. \quad (6.20)$$

In (6.20), equality is attained for all functions of the form

$$f(\mathbf{x}) := a(\mathbf{x}) + \frac{\mu}{2} \|\cdot\|^2,$$

where $a(\cdot)$ is any affine function.

Proof The error lower bound is a direct consequence of Lemma 6.2.2. So it remains to check that the upper bound holds, too. Assume that cubature formula (6.16) be a snd-formula. So, for every strongly convex function f with fixed parameter μ , $E[f] \leq 0$. Hence, by Lemma 6.2.1 for every convex function g , $E[g] \leq 0$. Let $g = f - \frac{\mu}{2} \|\cdot\|^2$, since g is convex so

$$E[g] = E\left[f - \frac{\mu}{2} \|\cdot\|^2\right] \leq 0.$$

By the linearity of E we have

$$E[f] - \frac{\mu}{2} E[\|\cdot\|^2] \leq 0,$$

which yields

$$E[f] \leq \frac{\mu}{2} E[\|\cdot\|^2].$$

Conversely, if the error functional satisfies in the inequalities (6.20), since for all μ -strongly convex functions $f \in C^{1,1}(\Omega)$ we have $\mu \leq l(\nabla f)$, so according to Lemma 6.2.2 we conclude that

6.3. Characterization of snd-cubature formulas in terms of the existence of a set of barycentric coordinates

$E[f] \leq 0$. It means that cubature formula (6.16) is a snd-formula. The equality case can be verified easely.

6.3 Characterization of snd-cubature formulas in terms of the existence of a set of barycentric coordinates

Let $\Omega \subset \mathbb{R}^d$ be a compact convex polytope of positive measure, and let $\mathbf{X} := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ be a finite subset that includes the vertices of Ω . Thus, the convex hull of \mathbf{X} must be equal to Ω .

The cubature formulae considered in [11] is introduced as follows:

Definition 6.3.1 For n points $\mathbf{X} := \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \Omega$, called nodes, and associated positive numbers A_1, A_2, \dots, A_n , we say that

$$\{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}, \quad (6.21)$$

defines the definite cubature formula

$$\int_{\Omega} f(\mathbf{x}) d\mathbf{x} = \sum_{i=1}^n A_i f(\mathbf{x}_i) + R_n[f], \quad (6.22)$$

if there exists a $\sigma \in \{-1, 1\}$ such that $\sigma R_n[f] \geq 0$ for all convex functions $f \in C(\Omega)$.

Then it was characterized the definite cubature formula as follows:

Theorem 6.3.2 [11] A set $\{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}$ defines a negative definite cubature formula on Ω if and only if there exists a partition of unity $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$ on Ω such that

$$\mathbf{x} = \sum_{i=1}^n \varphi_i(\mathbf{x}) \mathbf{x}_i \quad (\text{a.e. on } \Omega), \quad (6.23)$$

and

$$A_i = \int_{\Omega} \varphi_i(\mathbf{x}) d\mathbf{x} \quad (i = 1, 2, \dots, n). \quad (6.24)$$

Now, we utilize the above theorem to provide a necessary and sufficient condition for the existence of the snd-formulas. It consists of checking the existence of a set of barycentric coordinates.

Theorem 6.3.3 A set $\mathfrak{a} = \{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}$ defines a μ snd-cubature formula on Ω if and only if there exists a set of barycentric coordinates $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ on Ω such that

$$\mathbf{x} = \sum_{i=1}^n \lambda_i(\mathbf{x}) \mathbf{x}_i \quad (\text{a.e. on } \Omega), \quad (6.25)$$

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

and

$$A_i = \int_{\Omega} \lambda_i(\mathbf{x}) d\mathbf{x} \quad (i = 1, 2, \dots, n). \quad (6.26)$$

Proof Let $\{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}$ defines a μ snd-cubature formula on Ω . Then, according to the definition the error functional E satisfies

$$E[f] \leq 0, \quad (6.27)$$

for any μ -strongly convex function f . We deduce then by Lemma 6.2.1 that, for every convex function $g \in C(\Omega)$, we have

$$E[g] \leq 0. \quad (6.28)$$

This means that the estimate

$$\int_{\Omega} g(\mathbf{x}) d\mathbf{x} \leq \sum_{i=1}^n A_i g(\mathbf{x}_i),$$

holds for every convex function $g \in C(\Omega)$. Hence by Theorem 6.3.2, there exists a set of barycentric coordinates $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ on Ω , which satisfies the required conditions (6.25) and (6.26). Conversely, assume that there exists a set of barycentric coordinates $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ on Ω , such that conditions (6.25) and (6.26) hold. Let f be convex on Ω . Then, since f is convex, by Jensen's inequality it follows from (6.25) that

$$f(\mathbf{x}) \leq \sum_{i=1}^n \lambda_i(\mathbf{x}) f(\mathbf{x}_i).$$

Integrating both sides over Ω and using (6.26), we obtain the inequality

$$E[f] := \int_{\Omega} f(\mathbf{x}) d\mathbf{x} - \sum_{i=1}^n A_i f(\mathbf{x}_i) \leq 0.$$

Since the above inequality holds for every convex function, then according to Lemma 6.2.1 we also have, for every μ -strongly convex function,

$$E[f] \leq 0. \quad (6.29)$$

This shows that $\{(A_i, \mathbf{x}_i) : i = 1, 2, \dots, n\}$ defines a μ snd-cubature formula on Ω . ■

6.4 Integral Approximation using barycentric coordinates

Many of useful properties of the classical trapezoidal quadrature rule (6.1) on the interval $[a, b]$ can be carried over directly to the d -dimensional hypercube $\prod_{i=1}^d [a_i, b_i]$ by using tensor products of d copies of this latter. Non-tensorial constructions of the trapezoidal cubature formula are rare in the case of an arbitrary polytope. In general, leaving the tensor-product setting

6.4. Integral Approximation using barycentric coordinates

causes a lot of difficulties in theoretical as well as in computational aspects. From the theoretical point of view it gets harder to find a suitable set of barycentric coordinates needed for their constructions as we did for the one-dimensional case. An example of a non-tensorial construction on surplices with the derivation of an efficient computational scheme for the trapezoidal cubature formulas can be found in [10]. Using generalized barycentric coordinates, this section shows how the simple univariate trapezoidal rule (6.1) can be extended to arbitrary higher-dimensional polytopes. To this end, let $X_m = \{\mathbf{x}_i\}_{i=0}^m$ be a given finite set of pairwise distinct points in $\Omega \subset \mathbb{R}^d$, with $\Omega = \text{conv}(X_m)$ denoting the convex hull of the point set X_m . We are interested in approximating the integral of an unknown function $f : \Omega \rightarrow \mathbb{R}$ from given function values $f(\mathbf{y}_0), f(\mathbf{y}_1), \dots, f(\mathbf{y}_n)$ where $Y_n := \{\mathbf{y}_i\}_{i=0}^n \subset \Omega$. In order to obtain a simple and stable global approximation of the integral of f on Ω , we may consider a μ cubature formula of the following form:

$$I_n[f] := \sum_{i=0}^n A_i f(\mathbf{y}_i). \quad (6.30)$$

Theorem 6.3.3 tells us that there exists a set of barycentric coordinates $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ on Ω such that

$$\mathbf{x} = \sum_{i=0}^n \lambda_i(\mathbf{x}) \mathbf{y}_i \quad (\text{a.e. on } \Omega), \quad (6.31)$$

and

$$A_i = \int_{\Omega} \lambda_i(\mathbf{x}) d\mathbf{x} \quad (i = 0, 1, \dots, n). \quad (6.32)$$

For any function $f \in C^{1,1}(\Omega)$, the functional

$$E_n[f] := E_n[f, \boldsymbol{\lambda}] = I_n[f] - \int_{\Omega} f(\mathbf{x}) d\mathbf{x}, \quad (6.33)$$

will be reserved exclusively to denote the incurred approximation error between the integral of f and its approximation $I_n[f]$.

We now give a simple expression of the error $E_n[\|\cdot\|^2]$ in terms of the barycentric coordinates $\{\lambda_0, \dots, \lambda_n\}$.

Lemma 6.4.1 *The error $E_n[\|\cdot\|^2]$ when approximating the integral of the quadratic function $\|\cdot\|^2$ by $I_n[\|\cdot\|^2]$ can be expressed as:*

$$E_n[\|\cdot\|^2](\mathbf{x}) = \sum_{i=0}^n \int_{\Omega} \lambda_i(\mathbf{x}) \|\mathbf{x} - \mathbf{y}_i\|^2 d\mathbf{x}. \quad (6.34)$$

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

Proof For $f(\mathbf{x}) = \|\mathbf{x}\|^2$, we find by a simple calculation that

$$f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y}_i - \mathbf{x} \rangle = \|\mathbf{y}_i\|^2 - \|\mathbf{x} - \mathbf{y}_i\|^2.$$

Hence, multiplying on each side by λ_i , summing up with respect to i from 0 to n , using the linear precision property of barycentric coordinates and rearranging, we get the desired result and completes the proof of the Lemma. ■

The following Lemma shows that if the cubature formula I_n approximates every strongly convex function from above, then it generates a sharp lower bound for the error of any strongly convex function.

Lemma 6.4.2 Let μ be a positive real number. If the parycentric coordinate approximation functional I_n approximates every μ -strongly convex function from above then for every μ -strongly convex function f , it holds

$$\frac{\mu}{2} \sum_{i=0}^n \int_{\Omega} \lambda_i(\mathbf{x}) \|\mathbf{x} - \mathbf{y}_i\|^2 d\mathbf{x} \leq I_n[f] - \int_{\Omega} f(\mathbf{x}) d\mathbf{x}. \quad (6.35)$$

Equality in (6.35) is attained for all functions of the form

$$f(\mathbf{x}) = a(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x}\|^2, \quad (6.36)$$

where $a(\cdot)$ is any affine function.

Proof Let us fix f as a μ -strongly convex function on C , then for any $x, y \in C$

$$f(ty + (1-t)x) \leq tf(y) + (1-t)f(x) - \frac{\mu}{2}t(1-t)\|x - y\|^2.$$

Hence

$$tf(y) + (1-t)f(x) \geq f(ty + (1-t)x) + \frac{\mu}{2}t(1-t)\|x - y\|^2,$$

which yields

$$tf(y) + f(x) - tf(x) \geq f(ty + (1-t)x) + \frac{\mu}{2}t(1-t)\|x - y\|^2.$$

This can be expressed as

$$t(f(y) - f(x)) \geq f(ty + (1-t)x) - f(x) + \frac{\mu}{2}t(1-t)\|x - y\|^2.$$

Dividing the above inequality by t gives

$$f(y) - f(x) \geq \frac{f(ty + (1-t)x) - f(x)}{t} + \frac{\mu}{2}(1-t)\|x - y\|^2.$$

6.4. Integral Approximation using barycentric coordinates

Therefore, letting $t \rightarrow 0$ yields

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{x} - \mathbf{y}\|^2.$$

By letting \mathbf{y}_i instead of \mathbf{y} we have

$$f(\mathbf{y}_i) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y}_i - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{x} - \mathbf{y}_i\|^2.$$

Multiplying on each side by λ_i , summing up with respect to i from 0 to n and integrating each term, we get the desired result and completes the proof of the Lemma. The case of equality is easily verified. ■

The following Lemma gives an upper bound for the absolute value of the error of any function possessing Lipschitz continuous gradient:

Lemma 6.4.3 The following error estimate holds for every function $f \in C^{1,1}(\Omega)$:

$$\left| I_n[f] - \int_{\Omega} f(\mathbf{x}) \, d\mathbf{x} \right| \leq \frac{L(\nabla f)}{2} \sum_{i=0}^n \int_{\Omega} \lambda_i(\mathbf{x}) \|\mathbf{x} - \mathbf{y}_i\|^2. \quad (6.37)$$

Equality in (6.37) is attained for all functions of the form

$$f(\mathbf{x}) = a(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x}\|^2, \quad (6.38)$$

where $a(\cdot)$ is any affine function.

Proof This Lemma is an immediate consequence of Theorem 6.2.6 and Lemma 6.4.1. The case of equality is easily verified. ■

Now everything is set for giving an upper bound and a lower bound for the approximation error estimate $E_n[f] = I_n[f] - \int_{\Omega} f(\mathbf{x}) \, d\mathbf{x}$ of any μ -strongly convex function f , having Lipschitz continuous gradient.

Theorem 6.4.4 Let μ be a positive real number. Then, for every μ -strongly convex function $f \in C^{1,1}(\Omega)$ and any $\mathbf{x} \in \Omega$, it holds:

$$\frac{\mu}{2} \sum_{i=0}^n \int_{\Omega} \lambda_i(\mathbf{x}) \|\mathbf{x} - \mathbf{y}_i\|^2 \leq I_n[f] - \int_{\Omega} f(\mathbf{x}) \, d\mathbf{x} \leq \frac{L(\nabla f)}{2} \sum_{i=0}^n \int_{\Omega} \lambda_i(\mathbf{x}) \|\mathbf{x} - \mathbf{y}_i\|^2. \quad (6.39)$$

Equality in (6.39) is attained for all functions of the form

$$f(\mathbf{x}) = a(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x}\|^2, \quad (6.40)$$

where $a(\cdot)$ is any affine function.

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

Proof This is an immediate consequence of Lemmas 6.4.1, 6.4.2, 6.4.3 and Theorem 6.2.6. The case of equality is easily verified. ■

Remark 6.4.5 In the univariate case, a simple inspection of the error estimates (6.39) reveals that (6.39) is nicely reduced to the simple form given in (6.3) and (6.6).

6.5 Practical Construction of snd-Cubature Formulas

We now turn to a practical construction of snd-cubature formulas. To this end, let us first consider the case where Ω is a non-degenerate simplex in \mathbb{R}^d with $\mathbf{x}_i, i = 1, 2, \dots, d+1$, being the set of its vertices. Then each $\mathbf{x} \in \Omega$ has a unique representation as a convex combination

$$\mathbf{x} = \sum_{i=1}^{d+1} \lambda_i(\mathbf{x}) \mathbf{x}_i, \quad (6.41)$$

where λ_i is the restriction to Ω of the affine function that attains the value 1 at \mathbf{x}_i and is zero at all the other vertices of Ω . The value $\lambda_i(\mathbf{x})$ is the barycentric coordinate of \mathbf{x} with respect to \mathbf{x}_i . Then, if f is convex, by Jensen's inequality it follows from (6.41) that

$$f(\mathbf{x}) \leq \sum_{i=1}^{d+1} \lambda_i(\mathbf{x}) f(\mathbf{x}_i).$$

Integrating both sides over Ω and using the fact that $\int_{\Omega} \lambda_i(\mathbf{x}) d\mathbf{x} = \frac{|\Omega|}{d+1}, i = 1, 2, \dots, d+1$, (it is well-known that the integral of an affine function f over a simplex Ω is the arithmetic mean of the values of f at the vertices of Ω times $|\Omega|$), we deduce that

$$\int_{\Omega} f(\mathbf{x}) d\mathbf{x} \leq Q^{\text{TraR}}(f), \quad (6.42)$$

$$Q^{\text{TraR}}(f) := \sum_{i=1}^{d+1} \frac{|\Omega|}{d+1} f(\mathbf{x}_i). \quad (6.43)$$

Consequently, by Lemma 6.2.1, the set of barycentric coordinates $\lambda_1, \lambda_2, \dots, \lambda_{d+1}$ produces the snd-system

$$\left\{ \left(\frac{|\Omega|}{d+1}, \mathbf{x}_i \right) : i = 1, 2, \dots, d+1 \right\}.$$

It is the only snd-system on Ω which has no other nodes than the vertices.

Now let $\mathbf{X} = \{\mathbf{x}_i \in \mathbb{R}^d, i = 1, 2, \dots, n\}$ be an arbitrary set of points of \mathbb{R}^d . The previous approach can be generalized when $\Omega = \text{conv}(\mathbf{X})$ is an arbitrary polytope in \mathbb{R}^d . A triangulation \mathcal{T} of Ω with respect to \mathbf{X} is a decomposition of Ω into d -dimensional simplices such that \mathbf{X} is the set of all their vertices, and the intersection of any two simplices consists of a common lower-

dimensional simplex or is empty. As we mentioned in section 6.1, triangulations of compact convex polytopes exist¹. Indeed, given any finite set \mathbf{X} of points that do not all lie on a hyperplane, Chen and Xu [2, p. 301] describe a lifting-and-projection procedure which results in a triangulation of the convex hull of \mathbf{X} with respect to \mathbf{X} . For an explicit statement on the existence of triangulations with a proof based on an algorithmic method, see [12, Theorem 3, part a).

Now let $\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_l$ be the simplices of \mathbb{T} , and let N_i be the set of all integers j such that \mathbf{x}_i is a vertex of \mathbf{S}_j . If $\mathbf{x} \in \mathbf{S}_j$ and $j \in N_i$, then we denote by $\lambda_{ij}(\mathbf{x})$ the barycentric coordinate of \mathbf{x} with respect to \mathbf{x}_i for the simplex \mathbf{S}_j . It is easily verified that if $\mathbf{x} \in \mathbf{S}_j \cap \mathbf{S}_k$, then $\lambda_{ij}(\mathbf{x}) = \lambda_{ik}(\mathbf{x})$ if $j, k \in N_i$ and $\lambda_{ij}(\mathbf{x}) = 0$ if $j \notin N_i$. Therefore, setting

$$\phi_i(\mathbf{x}) := \begin{cases} \lambda_{ij}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathbf{S}_j \text{ and } j \in N_i \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \dots, n$, we obtain a well-defined barycentric coordinates $\phi_1, \phi_2, \dots, \phi_n$. This obviously produces the snd-formula

$$\int_{\Omega} f(\mathbf{x}) d\mathbf{x} = Q^{\text{tra}}(f) + E[f], \quad (6.44)$$

where

$$Q^{\text{tra}}(f) = \sum_{i=1}^n \left(\sum_{j \in N_i} \frac{|\mathbf{S}_j|}{d+1} \right) f(\mathbf{x}_i). \quad (6.45)$$

Let $\mathbb{T}(\Omega)$ be any triangulation of the point set X_n . Then $\boldsymbol{\lambda}^{\mathbb{T}(\Omega)} := \{\lambda_i^{\mathbb{T}(\Omega)}\}_{i=0}^n$ denotes the set of barycentric coordinates associated with each \mathbf{x}_i of X_n . Now we list the basic properties of $\boldsymbol{\lambda}^{\mathbb{T}(\Omega)}$ which are particularly relevant to us:

- (1) They are well-defined, piecewise linear and nonnegative real-valued continuous functions.
- (2) The function $\lambda_i^{\mathbb{T}(\Omega)}$ satisfies the delta property, which equals 1 at \mathbf{x}_i and 0 at all other points in $X_n \setminus \{\mathbf{x}_i\}$, that is, $\lambda_i^{\mathbb{T}(\Omega)}(\mathbf{x}_j) = \delta_{ij}$ (δ is the Kronecker delta).

We denote by

$$E_n^{\mathbb{T}(\Omega)}[f](\mathbf{x}) := \sum_{i=0}^n \lambda_i^{\mathbb{T}(\Omega)}(\mathbf{x}) f(\mathbf{x}_i) - f(\mathbf{x}). \quad (6.46)$$

As regard the error estimates (6.46), in [6] it was established that every set of barycentric coordinates generated by a Delaunay triangulation is optimal, in the sense that for all possible

¹It seems that in dimension $d = 3$ the existence was already known to mathematicians like Euler and Dirichlet.

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

barycentric coordinates, Delaunay triangulation provides the minimal barycentric approximation error $E_n[\|\cdot\|^2]$:

Theorem 6.5.1 *Let $\mathbb{T}(\Omega)$ be a triangulation of the point set \mathbf{X}_n . Then the following statements are equivalent.*

- (i) $\mathbb{T}(\Omega)$ is a Delaunay triangulation.
- (ii) For any set of barycentric coordinates $\lambda = \{\lambda_i\}_{i=0}^n$ and for all $\mathbf{x} \in \Omega$, there holds

$$0 \leq E_n^{\mathbb{T}(\Omega)}[\|\cdot\|^2](\mathbf{x}) \leq E_n[\|\cdot\|^2, \lambda](\mathbf{x}).$$

This optimality condition also characterizes Delaunay triangulation.

6.6 Numerical experiments in 3D

In this section we provide numerical test, which we perform in order to validate our theoretical predictions. We have considered the following function of three variables as test function

$$g(x, y, z) = \exp(ax + by + cz),$$

and the domain of integration is the pyramid Pyr given in the Cartesian coordinate system (x, y, z) by the inequalities:

$$Pyr = \{(x, y, z) \in \mathbb{R}^3 : 0.3z < x < 1 - 0.3z, 0.3z < y < 1 - 0.3z, 0 < z < 1\}. \quad (6.47)$$

The algorithm for computing the approximate values of the integral is as follows:

1. Pyramid should be decomposed into tetrahedra, see Figure 6.2, a.
2. Each of tetrahedra should be mapped onto the reference one, see Figure 6.2, b.
3. For integration of function g over the reference tetrahedron the method, $Q^{\text{TraR}}(g)$ should be applied, where $Q^{\text{TraR}}(g)$ is defined by the formula (6.43).
4. The results are the sums of approximate values of integrals over all tetrahedra in the decomposition of the pyramid.

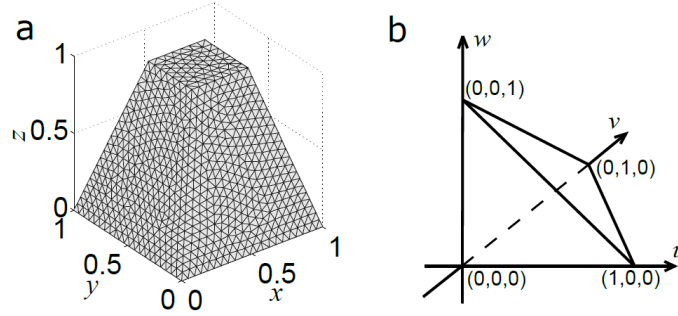


Figure 6.2 – Domain of the pyramid and its decomposition into tetrahedra generated by DistMesh (a). The characteristic linear size of tetrahedra is $1/21$. Reference tetrahedron (b)

Let us give more details about these steps.

For decomposition of the domain P_{yr} the DistMesh package was used that is a simple triangular mesh generator in MATLAB based on Delaunay triangulation. A detailed description of the program is provided in [19, 20] or <http://persson.berkeley.edu/distmesh>. Specifically, we used the code of the Problem #3 from the webpage available at the address :

https://people.sc.fsu.edu/~jburkardt/m_src/distmesh_3d/distmesh_3d.html

For computing the errors of our methods we need to compute the exact value of integral of function $g(x, y, z)$ over the pyramid P_{yr} , assuming that P_{yr} is given by its H -representation (6.47) or, alternatively, in its corresponding V -representation. We should mention that some useful methods for computing such integrals are discussed in [15, Section 2]. The exact value of this integral is

$$I_{pyr}(g) = K(Aa^3 + Ba^2b + Ca^2c + Db^2a + Eb^2c + Fb^3 + Gc^3 + Hc^2a + Ic^2b + Jabc),$$

$$\text{where } K = \frac{10}{ab(3a - 3b - 10c)(3a + 3b - 10c)(3a - 3b + 10c)(3a + 3b + 10c)},$$

$$A = 27(e^b - e^a + e^{a+b} + \alpha + \beta - \gamma - \theta - 1),$$

$$B = 27(e^b - e^a - e^{a+b} - \alpha + \beta - \gamma + \theta + 1),$$

$$C = 90(e^{a+b} - e^a - e^b - \alpha + \beta + \gamma - \theta + 1),$$

$$D = 27(e^a - e^b - e^{a+b} - \alpha - \beta + \gamma + \theta + 1),$$

$$E = 90(-e^a - e^b + e^{a+b} - \alpha + \beta + \gamma - \theta + 1),$$

$$F = 27(e^a - e^b + e^{a+b} + \alpha - \beta + \gamma - \theta - 1),$$

$$G = 1000(e^a + e^b - e^{a+b} + \alpha - \beta - \gamma + \theta - 1),$$

Generalized barycentric coordinates and sharp strongly negative definite multidimensional numerical integration

N	4	8	16	32	64	128
$E_N^{\text{Tra}}(g)$	3.441E-01	6.520E-02	1.478E-02	3.420E-03	8.312E-04	2.074E-04

Table 6.1 – Errors obtained while integrating g with $a = 1$, $b = 2$, $c = 3$ over pyramid Pyr

N	8	16	32	64
$E_N^{\text{Tra}}(g)$	2.46763	2.14976	2.13358	2.05323

Table 6.2 – Orders of convergence obtained while integrating g with $a = 1$, $b = 2$, $c = 3$ over the pyramid Pyr

$$H = 300(e^a - e^b - e^{a+b} - \alpha - \beta + \gamma + \theta + 1),$$

$$I = 300(e^b - e^a - e^{a+b} - \alpha + \beta - \gamma + \theta + 1),$$

$$J = 180(-e^a - e^b - e^{a+b} + \alpha + \beta + \gamma + \theta - 1),$$

$$\alpha = e^{0.3a+0.3b+c}, \beta = e^{0.7a+0.3b+c}, \gamma = e^{0.3a+0.7b+c}, \theta = e^{0.7a+0.7b+c}.$$

After applying the above algorithm we got the asymptotics of the relative errors of our formulas for the case of function g with $a = 1$, $b = 2$, $c = 3$. The expression of the relative error for Q_3 cubature formula is as follows:

$$E_N^{\text{Tra}}(g) = \frac{Q^{\text{tra}}(g) - I_{pyr}(g)}{I_{pyr}(g)},$$

where $Q^{\text{tra}}(g)$ is defined by the formula (6.45). In Table 6.1, the values of the relative errors of integration are given for the case of test with $a = 1$, $b = 2$, $c = 3$. Table 6.2 shows the orders of convergence obtained for the test with $a = 1$, $b = 2$, $c = 3$. The orders are close to 2.

Bibliography

- [1] H. Brass, *Quadraturverfahren*, Vandenhoeck & Ruprecht, Göttingen (1977).
- [2] L. Chen, J. Xu, *Optimal Delaunay triangulations*, *J. Comput. Math.* 22 (2004) 299-308.
- [3] P.J. Davis, P. Rabinowitz, *Methods of Numerical Integration*, 2nd edn. Academic Press, Orlando (1984).
- [4] S.S. Dragomir, C.E.M. Pearce, *Selected Topics on HermiteHadamard Inequalities and Applications*, Internet Publication, <http://rgmia.vu.edu.au>, (2000).
- [5] G. Farin, *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, San Diego, (1993).
- [6] A. Guessab, *Approximations of differentiable convex functions on arbitrary convex polytopes*, *Appl. Math. Comput.* 240 (2014) 326-338.
- [7] A. Guessab, G. Schmeisser, *Construction of positive definite cubature formulae and approximation of functions via Voronoi tessellations*, *Adv. Comput. Math.* 32 (2010) 25-41.
- [8] A. Guessab, O. Nouisser, G. Schmeisser, *A definiteness theory for cubature formulae of order two*, *Constr. Approx.* 24 (2006) 263-288.
- [9] A. Guessab, T. Azimi Roushan, *Generalized barycentric coordinates and sharp strongly negative definite multi-dimensional Numerical Integration*, Accepted for publication in a Springer volume, entitled: *Approximation theory and analytic inequalities*, (2019).
- [10] A. Guessab, G. Schmeisser, *Convexity results and sharp error estimates in approximate multivariate integration*, *Math. Comput.* 73 (247) (2004) 1365-1384.
- [11] A. Guessab, G. Schmeisser, *Negative definite cubature formulae, extremality and delaunay triangulation*, *Constr. Approx.* 31 (2010) 95-113.

Bibliography

- [12] M. Joswig, *Beneath-and-beyond revisited*, In: Joswig, M., Takayama, N. (eds.) *Algebra, Geometry and Software Systems*, 121. Springer, Berlin (2003).
- [13] T. Ju, P. Liepa, J. Warren, *A general geometric construction of coordinates simplicial polytope*. *Computer Aided Geometric Design*, 24 (3) (2007) 161-178.
- [14] P.J. Kelly, M.L. Weiss, *Geometry and Convexity*, John Wiley & Sons, New York, (1979).
- [15] J.B. Lasserre, E.S. Zeron, *Solving a class of multivariate integration problems via Laplace techniques*, *applicaciones mathematicae* 28 (2001) 391-405.
- [16] D.H. Lehmer, *Euler constants for arithmetical progressions*, *Acta Arith.* 27 (1975) 125-142.
- [17] M. Mikolás , *On certain sums generating the Dedekind sums and their reciprocity laws*, *Pacific J. Math.* 7 (1957) 1167-1178.
- [18] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*, Kluwer-Academic. (2003).
- [19] P.O. Persson, G. Strang, *A Simple Mesh Generator in MATLAB*, *SIAM Review*, 46 (2004) 328-345.
- [20] P.O. Persson, *Mesh Generation for Implicit Geometries*, Ph.D. thesis, Department of Mathematics, MIT, (2004).
- [21] Siu-Wing Cheng, Tamal K. Dey, Jonathan R. Shewchuk, *Delaunay Mesh Generation*, CRC press Taylor and Francis Group, LLC, (2013).
- [22] C. Talischi, G.H. Paulino, A. Pereira, I.F.M. Menezes, *PolyMesher: a general-purpose mesh generator for polygonal elements written in Matlab*, *Struct Multidisc Optim* 45 (2012) 309-328.
- [23] J. Warren, *Barycentric coordinates for convex polytopes*. *Advances in Computational Mathematics*, 6(2) (1996) 97-108.
- [24] E. Wolfstetter, *Topics in microeconomics: industrial organization, auctions, and incentives*, Cambridge University, (2000).
- [25] G.M. Ziegler, *Lectures on Polytopes*, Springer Verlag, New York (1995)

Conclusion and final remarks

As our main results are appeared in chapters two to five, we are going to discuss about these chapters briefly and separately to mention the main results and final remarks of each chapter as follows:

*In extending the definition of convex function to operator convex function, since we deal only with the sum of the operators with respect to continuous functional calculus it can be easily done but in defining the operator geometrically convex functions as we are dealing with the multiplication of operators and since the multiplication of operators are not commutable one must be cautious. **In chapter two**, by the definition of weighted geometric mean, using some lemmas and theorems, we defined operator geometrically convex functions, and we tried to prove the Hermite-Hadamard types inequality for these functions. At the beginning of the discussion we sought to prove the exact inequality of Hermite-Hadamard for operator geometrically convex functions, but we found that there would be a contradiction if established, because, the mid term and the last one are the log-mean and the geometric mean, respectively, and it is known that the log-mean is greater than the geometric mean. By giving some samples we finished this chapter.*

*Audenaert obtained inequality (3.17) for unitarily invariant norms that interpolates between the arithmetic-geometric mean inequality and the Cauchy-Schwarz inequality for matrices. A refined version of this inequality was given in [1] of **chapter three** for the Hilbert-Schmidt norm by applying k th antisymmetric tensor powers of the matrices and their properties. In this chapter, we refined the mentioned inequality for general case, that is for all unitarily invariant norms, not just for Hilbert-Schmidt norm and by completely different approach. In this regard, we refined two other well-known inequalities too.*

***In chapter four**, the main goal was to present numerical radius inequalities for matrices and extend some unitarily invariant norms inequality, such as Heinz and Young inequalities, to the context of numerical radius, but it is notable that due to weak unitary invariant of the numerical radius, one can not expect the exact form holds true. Actually, the obtained inequality maybe is weaker than the corresponding version. In the sequel, we obtained some Berezin number inequalities based on the definition of Berezin symbol and we established some inequalities*

Conclusion and final remarks

involving generalization of Berezin number inequalities.

In chapter five, first we discussed about approximation a class of convex functions. Namely, it was provided by Guessab a simple and elegant characterization of upper approximation operators then it was showed that it is possible to apply the given error estimate to a differentiable convex function with Hölder continuous gradient. In this chapter, we tried to extend the result to operator version, and apply the obtained result to refining the Hermite-Hadamard inequality.

In chapter six, we established a new and efficient way of approximating a given function of multiple variables by linear operators, which approximate integral of all strongly convex functions from above. We also assumed that these functions are continuously differentiable functions with Lipschitz continuous gradients. This additional information is used to characterize sharp error estimates which are always controlled by the Lipschitz constants of the gradients, the strong convexity parameter and the error associated with using the quadratic function. Then, we turned to a practical construction of snd-cubature formulas. Finally, we provided some numerical tests, which we performed in order to validate our theoretical predictions.