The University of Hong Kong

Department of Mathematics

Uncertainty principle and its variations

Mini research report for

MATH 4910

Senior mathematics seminar

Supervised by Dr. Xin Zhang

Lyu Chenxin

Abstract

This research proposal contains three major parts including the the problems I concerned, my study objectives and the methodologies that can be applied to solve those problems. Optimal control is important as a method to create the provably best solution for control problems in engineering and science. My thesis will develop an application of differential equations to solve optimal control problems. I will first investigate optimal control from a mathematical standpoint and gain an understanding of the theory. Then, I will apply these techniques to the problem to be studied.

Contents

1 Introduction

In quantum physics, the Uncertainty Principle (also called Heisenberg's Uncertainty Principle) is a kind of mathematical inequality that claims that fundamental constraints on the accuracy of the values of pairs of variables such as the position and momentum of a particle can be predicted from initial conditions. Those pairs of variables are called complementary variables or canonically conjugate variables. When translated into the language of harmonic series, uncertainty principle states that for any non-zero function, the function itself and its Fourier transform can't be simultaneously localized with great precision. Heisenberg's seminal paper of 1927 first introduced the idea of uncertainty principle from a physical perspective. The independent mathematical research by Kennard and Weyl on the corresponding content leads to a wide discussion of mathematical formulation of uncertainty principle. This topic has been intensively discussed for nearly a century, and new results are constantly published today. So while this report studied many perspectives of uncertainty principle, it is not a comprehensive survey.

Surprisingly, the Heisenberg's uncertainty principle has a practical meaning in harmonic series. We can denote $f(t)$ as the amplitude of a sound wave at a certain time t, then $f(\hat{t})$ (the Fourier transform of $f(t)$) may give us an idea about how $f(t)$ could be expressed by sine wave functions with distinct frequencies. The uncertainty principle provides a lower bound on how well a signal can be both time-limited and bandwidth-limited.

At the same time, another sort of uncertainty principle was developed which focusing not on the localization of a function and its Fourier transform but rather on its decay at infinity. Godfrey Harold Hardy is the first one to prove that it is not possible for a function and its Fourier transform to have a greater rate of decay than $e^(- x^2)$ at the same time.

Driven by applications in signal processing, Donoho and Stark began to investigate a new type of uncertainty principle that deals not only with functions defined on R, but on finite groups. Many concepts discussed in previous study are not well-defined when one is dealing with functions over finite groups. However, the size of the function support and Many concepts discussed above, such as variance and decay at infinity, have no meaning when dealing with functions over finite groups. However, the size of the function support and other non-localization measure are well defined over finite groups. Donoho–Stark uncertainty principle is developed to extend the localization problem to a finite group base.

Many different types of uncertainty principles share great similarities in intuition. For the same reason their proofs apply various techniques and a wide range of Fourier transform properties. So in this report, we will build up a universal framework to prove different kinds of uncertainty principles.

2 The primary uncertainty principle and and k-Hadamard matrices

2.1 The primary uncertainty principle

Before we proceed to the main content of the report, let's give a brief recall about the definition operator norm. Let V, U be any two real or complex vector spaces, and let $\|\cdot\|_V$ and $\|\cdot\|_U$ be the norms on V, U, respectively. We set $A : V \to U$ to be a linear map. We define the operator norm of A to be

$$\|A\|_{V \to U} = \sup_{0 \neq v \in V} \frac{\|Av\|_U}{\|v\|_V} = \sup_{\substack{v \in V \\ \|v\|_V = 1}} \|Av\|_U.$$

For $1 \leq p, q \leq \infty$, we will denote by $\|A\|_{p \to q}$ the operator norm of A when $\|\cdot\|_V$ is the L^p norm on V and $\|\cdot\|_U$ is the L^q norm on U.

Primary uncertainty principle Let V, U be real or complex vector spaces, each equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$, and let $A : V \to U$ and $B : U \to V$ be linear operators. Suppose that $\|A\|_{1\to\infty} \leq 1$ and $\|B\|_{1\to\infty} \leq 1$. Suppose too that $\|BAv\|_\infty \geq k\|v\|_\infty$ for all $v \in V$, for some parameter $k > 0$. Then for any $v \in V$,

$$\|v\|_1 \|Av\|_1 \geq k\|v\|_\infty \|Av\|_\infty.$$

Proof. Since $\|A\|_{1\to\infty} = \sup_{\substack{v \in 1 \\ \|v\|_1 = 1}} \|Av\|_\infty \leq 1$, we have that

$$\|Av\|_\infty \leq \|v\|_1.$$

Similarly, since we have $\|B\|_{1\to\infty} \leq 1$,

$$\|BAv\|_\infty \leq \|Av\|_1.$$

Multiplying these two inequalities together, we find that

$$\|v\|_1\|Av\|_1 \geq \|Av\|_\infty\|BAv\|_\infty \geq k\|v\|_\infty\|Av\|_\infty,$$

as we want.

Remark This primary uncertainty principle holds no matter what dimensions U and V have. We choose L^1 and $L^(\infty)$ because these two norms are well defined. L^1 and L^∞ norms on R^n or C^n. In the modern study of uncertainty principle, many generally would make $B = A^*$. As we have learner in linear algebra, $1 \to \infty$ norm of a matrix is equal to the maximum absolute value of the entries in the matrix. Then we can claim that the absolute values of all entries of A are bounded by 1 if and only if $\|A\|_{1\to\infty} \leq 1$. Therefore, [7]$\|B\|_{1\to\infty} = \|A\|_{1\to\infty}$ given that $B = A^*$. Also, we need to check that $\|A^*Av\|_\infty \geq k\|v\|_\infty$ for all $v \in V$ so that we can apply the primary uncertatin principle in this case. To do this, we need to introduce the k - Hadamard matrices.

2.2 k - Hadamard matrices

Definition Assume that $A \in C^{m\times n}$ and $k > 0$. Then A is k - Hadamard if all the entries of A is bounder by 1 in absolute value and $\|A^*Av\|_\infty \geq k\|v\|_\infty$ for all $v \in C^n$.

Proposition We say A is a k-Hadamard matrix if every entry of A has absolute value at most 1, A^*A is invertible, and $\left\|(A^*A)^{-1}\right\|_{\infty\to\infty} \leq 1/k$.

Rephrased primary uncertainty principle Assume that $A \in C^{m\times n}$ and k is Hadamard. Then for any $v \in C^n$, we have that

$$\|v\|_1\|Av\|_1 \geq k\|v\|_\infty\|Av\|_\infty.$$

Proof. Given that k is Hadamard, then we have $\|A^*Av\|_\infty \geq k\|v\|_\infty$ for all $v \in C^n$ and all the entries of A is bounder by 1 in absolute value. We can conclude that $\|A\|_{1\to\infty} \leq 1$ and $\|B\|_{1\to\infty} \leq 1$. Then all the conditions needed in the primary uncertainty principle are satisfied. Hence we complete the proof.

The general version of primary uncertainty principle Say that V, U are real or complex vector spaces, and $A : V \to U$ and $B : U \to V$ are linear operators. Let

$\|\cdot\|_{V^{(a)}}, \|\cdot\|_{V^{(b)}}$ be two norms on V, and let $\|\cdot\|_{U^{(a)}}, \|\cdot\|_{U^{(b)}}$ be two norms on U. Suppose that $\|A\|_{V^{(a)}\to U^{(b)}} \leq 1$ and $\|B\|_{U^{(a)}\to V^{(b)}} \leq 1$, and suppose that $\|BAv\|_{V^{(b)}} \geq k\|v\|_{V^{(b)}}$ for all $v \in V$ and some $k > 0$. Then for any $v \in V$,

$$\|v\|_{V^{(a)}}\|Av\|_{U^{(a)}} \geq k\|v\|_{V^{(b)}}\|Av\|_{U^{(b)}}.$$

Proof. Since $\|A\|_{V^{(a)}\to U^{(b)}} \leq 1$, we have that

$$\|Av\|_{U^{(b)}} \leq \|v\|_{V^{(a)}}.$$

Similarly, since we have $\|B\|_{U^{(a)}\to V^{(b)}} \leq 1$,

$$\|BAv\|_{U^{(b)}} \leq \|Av\|_{V^{(a)}}.$$

Multiplying these two inequalities together, we find that

$$\|v\|_{V^{(a)}}\|Av\|_{U^{(a)}} \geq \|Av\|_{U^{(b)}}\|BAv\|_{V^{(b)}} \geq k\|v\|_{V^{(b)}}\|Av\|_{U^{(b)}},$$

as expected.

3 Uncertainty principles in finite dimensions

3.1 The Donoho–Stark support-size uncertainty principle.

Before we proceed to the first subsection, let's recall the definition of support. For a vector $v \in C^n$, we can denote $\mathrm{supp}(v)$ as its support which refers to the set of coordinates i where $v_i \neq 0$. Let $f : G \to C$ be a function over finite group, we write supp (f) to denote the set of $x \in G$ for which $f(x) \neq 0$.

Support-size uncertainty principle Let $A \in C^{m\times n}$ be a k - Hadamard matrix. Then for any nonzero $v \in C^n$,

$$|\operatorname{supp}(v)||\operatorname{supp}(Av)| \geq k.$$

Proof. By the rephrased primary uncertainty principle we have already proved, that for any nonzero vector v, we have $\|v\|_1\|Av\|_1 \geq k\|v\|_\infty\|Av\|_\infty$. Then it suffices for us to find a lower bound to the product of the support size of a function. Let v be any vector, we have

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| = \sum_{i\in\operatorname{supp}(v)} |v_i| \leq |\operatorname{supp}(v)|\|v\|_\infty.$$

Then we can prove that

$$|\operatorname{supp}(v)||\operatorname{supp}(Av)|\|v\|_\infty\|Av\|_\infty \geq \|v\|_1\|Av\|_1 \geq k\|v\|_\infty\|Av\|_\infty$$

Hence we complete the proof.

3.2 Support-size uncertainty principles over finite groups

In this subsection, I want to extend the Donoho–Stark support-size uncertainty principle to the Fourier transform over any finite groups and show that our framework mentioned above is also applicable in this situation. Before we do that, we have to define the Fourier transform in general finite groups.

Assume that G is an arbitrary finite group with order n, and we denote its group algebra as $C[G]$. Recall that an irreducible representation of a group is a group representation that has no nontrivial invariant subspaces. Let $\rho_1, \rho_2, \ldots, \rho_t$ be the irreducible representations of G over C, where W_i refers to a vector space over C with dimension d_i. Let's further assume that $\rho_i(x)$ is a unitary transformation on W_i for all $x \in G$.

The Fourier transform in general finite groups Given a function $f : G \to C$, its Fourier transform is defined by $\hat{f}(\rho_i) = \sum_{x \in G} f(x)\rho_i(x)$, so that $\hat{f}(\rho_i)$ is a linear transformation $W_i \to W_i$.

Then we try to build up an orthonormal basis for each vector space W_i and we denote this orthonormal basis by E_i. With the representation of the basis E_i, now we can consider $\rho_i(x)$ and $\hat{f}(\rho_i)$ as $d_i \times d_i$ matrices. To complete the definition of Fourier transform matrix, we can denote its columns as matrix entry vectors. We can define the matrix entry vector $c(i; j, k) \in C^n$ with indices $i \in [t]$ and $j, k \in [d_i]$. To be specific, the x coordinate of a matrix entry vector is the (j, k) entry of the matrix $\rho_i(x)$.

Orthogonality of matrix entries We have

$$\langle c(i; j, k), c(i'; j', k')\rangle = \begin{cases} n/d_i & \text{if } i = i', j = j', k = k' \\ 0 & \text{otherwise} \end{cases}$$

With these properties and indexes, we are able to give a formal definition to the Fourier transform matrix.

Fourier transform matrix Suppose F is a $n \times n$ matrix with rows indexed by the linear map G and its columns are indexed by tuples $(i; j, k)$ in alphabet order. The $(i; j, k)^{th}$ column of F is the vector $\sqrt{d_i}c(i; j, k)$. Then the vector F is called a Fourier transform matrix.

Suppose we have an element $f = \sum_{x \in G} f(x)x$, T_f refers to the left-multiplication of f in $C[G]$. T_f is also said to be a linear map $C[G] \rightarrow C[G]$. Since $C[G]$ have a standard basis, T_f could be represented as an $n \times n$ matrix. Both addition and multiplication of matrices could be understood as addition and multiplication in $C[G]$, then $C[G]$ can be seen as the subspace of $C^{n \times n}$ consisting of all matrices T_f

Remark By the orthogonality of matrix entries, we can conclude that $F^*F = FF^* = nI$.

If G is an abelian group, we can diagonalize T_f by conjugating with the Fourier transfrom matrix. And the diagonal of the entries of conjugated T_f would be $\hat{f}$. But if G is not an abelian group, we can conjugate T_f by the Fourier transform which turns T_f into a block-diagonal matrix, whose blocks are just d_i copies of the matrices $n\hat{f}\left(\rho_i\right)$. We define the Fourier transform of T_f to be the block-diagonal matrix $\widehat{T_f} := FT_fF^*$. Therefore, the subspace of $C^{n \times n}$ consisting of all block-diagonal matrices with d_i identical blocks of size $d_i \times d_i$ could be understood as the Fourier subspace.

Since $\widehat{T_f}$ has the block $\hat{f}\left(\rho_i\right)$ appeared d_i times at its diagonal, we can claim that

$$|\operatorname{supp}(\hat{f})| = \sum_{i=1}^{t} d_i \left|\operatorname{supp}\left(\hat{f}\left(\rho_i\right)\right)\right|$$

where $\operatorname{supp}\left(\hat{f}\left(\rho_i\right)\right)$ refers to the set of nonzero entries of $\hat{f}\left(\rho_i\right)$. Remember that the matrix $\hat{f}\left(\rho_i\right)$ changes with the choice of the basis E_i, so we need to define a new term called minimum support-size.

Minimum support-size The minimum support-size of $\hat{f}$ is

$$|\min - \operatorname{supp}(\hat{f})| = \min_{E_1, \ldots, E_t} |\operatorname{supp}(\hat{f})|$$

where the minimum is over all the orthonormal bases $E_1, \ldots, E_t$.

One may consider the minimum support-size of $\hat{f}$ to be its support-size in its most efficient representation. A terminology called rank-support was later proposed by Meshulam to be used as an alternative description for the support-size of $\hat{f}$

Rank-support For any function $f : G \to C$, the rank-support of $\hat{f}$ is denoted as $\mathrm{rk} - \mathrm{supp}(\hat{f}$ and $\mathrm{rk} - \mathrm{supp}(\hat{f}) = \mathrm{rank}\, T_f$.

As one matrix shares the same rank with its similar matrices, we see that rank $T_f = \mathrm{rank}\, \widehat{T_f}$. Since $\widehat{T_f}$ is a block-diagonal matrix, we see that

$$\mathrm{rank}\, \widehat{T_f} = \sum_{i=1}^{t} d_i \,\mathrm{rank}\left(\hat{f}\left(\rho_i\right)\right).$$

Again, recall that if $f(x) = \overline{f\left(x^{-1}\right)}$ for all $x \in G$, then f is called as a Hermitian function.

The relationship between rk-supp $(\hat{f})$ **and** $|\min - \mathrm{supp}(\hat{f})|$ Given a function $f : G \to C$, we can claim that rk-supp $(\hat{f}) \leq |\min - \mathrm{supp}(\hat{f})|$. Additionally, the equality holds if f is a Hermitian function.

Proof. Suppose there are k nonzero entries in $\widehat{T_f}$, then there would be at most s nonzero columns in $\widehat{T_f}$. Hence we are able to conclude that $\mathrm{rank}\, \widehat{T_f} \leq |\mathrm{supp}(\hat{f})|$ over all the bases $E_1, \ldots, E_t$. This implies that rk-supp $(\hat{f}) \leq \min_{E_1,\ldots,E_t} |\mathrm{supp}(\hat{f})| = |\min - \mathrm{supp}(\hat{f})|$. If we further assume that f is Hermitian, then it implies that

$$\left(\hat{f}\left(\rho_i\right)\right)^* = \sum_{x \in G} \overline{f(x)}\left(\rho_i(x)\right)^* = \sum_{x \in G} \overline{f(x)} \rho_i\left(x^{-1}\right) = \sum_{x \in G} f\left(x^{-1}\right) \rho_i\left(x^{-1}\right) = \hat{f}\left(\rho_i\right).$$

Hence each $\hat{f}\left(\rho_i\right)$ is Hermitian. For this reason we are able to build an orthonormal basis E_i for W_i where all the $\hat{f}\left(\rho_i\right)$ are diagonal matrices. With kind of orthonormal basis, we notice that $\widehat{T_f}$ is also diagonal. Moreover, the rank of $\widehat{T_f}$ exactly equals to the number of nonzero diagonal entries of $\widehat{T_f}$. This implies that $(\hat{f}) = |\min - \mathrm{supp}(\hat{f})|$.

The following theorem was proven by Meshulam in his paper at 1992 and we will state it without proof here.

Uncertainty principles for the support and the rank-support Given a finite group G and a function $f : G \to C$, $|\mathrm{supp}(f)|\, \mathrm{rk} - \mathrm{supp}(\hat{f}) \geq |G|$.

Together with the relationship between rk-supp $(\hat{f})$ and $|\min - \operatorname{supp}(\hat{f})|$, we can imply the following theorem:

Proposition For any finite group G and any $f : G \to C$,

$$|\operatorname{supp}(f)||\min - \operatorname{supp}(\hat{f})| \geq |G|.$$

Remark One may also notice that if the function f is Hermitian, the two inequalities above are equivalent.

Modified uncertainty principles for the support and the rank-support Given a function $f : G \to C$,

$$|\operatorname{supp}(f)| \operatorname{rk} - \operatorname{supp}(\hat{f}) \geq \frac{|G|}{4}.$$

Proof. To see this relationship, we construct a function $g : G \to C$ with $g(x) = f(x) + \overline{f(x^{-1})}$. One can easily check that g is Hermitian. By the relationship between rk-supp $(\hat{f})$ and $|\min - \operatorname{supp}(\hat{f})|$, we have that $\operatorname{rk} - \operatorname{supp}(\hat{g}) = |\min - \operatorname{sup}(\hat{g})|$. With the definition of support and rank-support, we have that

$$|\operatorname{supp}(g)| \leq 2|\operatorname{supp}(f)| \quad \text{and} \quad \text{rk-supp}(\hat{g}) \leq 2 \operatorname{rk} - \operatorname{supp}(\hat{f})$$

Then we add these two equations together, we get

$$|\operatorname{supp}(f)| \operatorname{rk} - \operatorname{supp}(\hat{f}) \geq \frac{1}{4}|\operatorname{supp}(g)| \operatorname{rk} - \operatorname{supp}(\hat{g}) = \frac{1}{4}|\operatorname{supp}(g)||\min - \operatorname{supp}(\hat{g})| \geq \frac{|G|}{4}$$

Hence, we finish the proof.

4 Acknowledgement

This report would not have been possible without the guidance of Dr. Tak Kwong WONG and Dr. Xin Zhang. I am especially indebted to the Department of Mathematics, in the University of Hong Kong. Each of the members of the department has provided me with extensive professional guidance and taught me a great deal about both academic knowledge and scientific research in general. Nobody has been more important to me in the pursuit of my academic objectives than the members of my family. I would like to thank my parents, whose love is with me in whatever I pursue.

References

[1] Anderson, J., Moradi, S., Rafiq, T. (2018). *Non-Linear Langevin and Fractional Fokker–Planck Equations for Anomalous Diffusion by Lévy Stable Processes.* Entropy (Basel, Switzerland), 20(10), 760.

[2] Oksendal, B. (2013). *Stochastic differential equations: an introduction with applications.* Springer Science Business Media.

[3] Evans, L. C. (2012). *An introduction to stochastic differential equations (Vol. 82).* American Mathematical Soc..

[4] Ren, J., Wu, J. (2019). *Probabilistic approach for nonlinear partial differential equations and stochastic partial differential equations with Neumann boundary conditions.* Journal of Mathematical Analysis and Applications, 477(1), 1-40.

[5] Marin, M., Öchsner, A. (2018). *Essentials of Partial Differential Equations.* Cham: Springer International Publishing AG.

[6] Evans, L. C. (1983). *An introduction to mathematical optimal control theory version 0.2.* Lecture notes available at http://math. berkeley. edu/ evans/control. course. pdf.

[7] Hoff, P., Thallmair, S., Kowalewski, M., Siemering, R., Vivie-Riedle, R. (2012). *Optimal control theory - closing the gap between theory and experiment.* Physical Chemistry Chemical Physics : PCCP, 14(42), 1446-14485.

[8] Hu, M., Ji, S., Peng, S., Song, Y. (2014). *Backward stochastic differential equations driven by G-Brownian motion.* Stochastic Processes and Their Applications, 124(1), 759-784.

[9] Carmona, R., Delarue, F. (2015). *Forward-Backward stochastic differential equations and controlled Mckean–Vlasov dynamics.* The Annals of Probability, 43(5), 2647-2700.

[10] Misztela, A. (2019). *Representation of Hamilton–Jacobi Equation in Optimal Control Theory with Compact Control Set.* SIAM Journal on Control and Optimization, 57(1), 53-77.

[11] Mortezazadeh, M., Wang, L. (2017). *A high-order backward forward sweep interpolating algorithm for semi-Lagrangian method.* International Journal for Numerical Methods in Fluids, 84(10), 584-597.

[12] Liu, Y., Du, Y., Li, H., He, S., Gao, W. (2015). *Finite difference/finite element method for a nonlinear time-fractional fourth-order reaction–diffusion problem.* Computers Mathematics with Applications, 70(4), 573-591.

[13] Crandall, M. G., Lions, P. L. (1983). *Viscosity solutions of Hamilton-Jacobi equations.* Transactions of the American mathematical society, 277(1), 1-42.

[14] Calvez, V., Lam, K. (2020). *Uniqueness of the viscosity solution of a constrained Hamilton–Jacobi equation.* Calculus of Variations and Partial Differential Equations, 59(5), Calculus of variations and partial differential equations, 2020-09-11, Vol.59 (5).

[15] Aseev, S. M., Kryazhimskii, A. V. (2007). *The Pontryagin maximum principle and optimal economic growth problems.* Proceedings of the Steklov institute of mathematics, 257(1), 1-255.

[16] Itô, K. (1951). *On stochastic differential equations (No. 4).* American Mathematical Soc..

[17] Zwart, H., Le Gorrec, Y., Maschke, B., Villegas, J. (2010). *Well-posedness and regularity of hyperbolic boundary control systems on a one-dimensional spatial domain.* ESAIM: control, optimisation and calculus of variations, 16(4), 1077-1093.

[18] Risken, H. (1996). *Fokker-planck equation.* In The Fokker-Planck Equation (pp. 63-95). Springer, Berlin, Heidelberg.

[19] De Palma, B., Erba, M., Mantovani, L., Mosco, N. (2019). *A Python program for the implementation of the [GAMMA]-method for Monte Carlo simulations.* Computer Physics Communications, 234, 294.

[20] Beard, R. W., Saridis, G. N., Wen, J. T. (1997). *Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation.* Automatica, 33(12), 2159-2177.

[21] Provencher, S. W. (1976). *Provencher, S. W. (1976). An eigenfunction expansion method for the analysis of exponential decay curves. The Journal of Chemical Physics, 64(7), 2772-2777.* The Journal of Chemical Physics, 64(7), 2772-2777.

[22] Straughan, B. (2013). *The energy method, stability, and nonlinear convection (Vol. 91).* Springer Science Business Media.

[23] Mc Reynolds, S. R., Bryson Jr, A. E. (1965). *A successive sweep method for solving optimal programming problems.* HARVARD UNIV CAMBRIDGE MA CRUFT LAB.