

(YITP-W-10-14)

基研研究会 量子科学における双対性とスケール (Duality and Scales in Quantum-Theoretical Sciences)

日程：2010年11月4日(木) - 6日(土)

場所：京都大学 基礎物理学研究所 湯川記念館 Panasonic 国際交流ホール

ホームページ：http://www.th.phys.titech.ac.jp/shikano/yitp10/

はじめに

本研究会では、下記のプログラムを見ても分かるように量子科学という題材を中心に物理、数学、情報、論理などの様々な領域横断的な研究者が集まった。更には、量子科学だけでなく生物学の分野からも植物ネットワークに関する招待講演があり、問題の多様性を様々な視点から論じる場として機能した。その上で、量子力学の基礎である量子測定理論の基礎の構築や近年ではハイゼンベルクの不確定性関係を取り入れた量子測定の文脈に即した新しい不確定性関係に関して導出をされ、量子論理の範疇で測定を特徴づけようとなされている名古屋大学の小澤正直先生の還暦祝いを懇親会の場で行った。最後に、本研究会の研究会報告が出版が当初の予定より大幅に遅れてしまったことを世話人一同お詫び致します。

概要

「量子的ミクロが本物で、マクロ古典は粗視化による虚像」との標準的見方では、ミクロからの演繹のみが唯一の理論的に「正しい」方法論だが、演繹展開の出発点で仮定されたミクロ量子系の理論的特徴づけの成否は如何に検証されるのか？実験観測データとの比較以外の方途はないが、どんなデータも誤差等「信頼不能なマクロ性」を免れない以上、「虚像」による「本物」の品質保証という本末転倒の恐れはないか？この Duheme-Quine paradox を記述領域・側面・精度の制約と整合する形で解決するのが、「ミクロ・マクロ双対性」とそれに基づく帰納・演繹往復の方法論である：適切な縮尺で描いた局所地図が対象領域を「正確に再現」するように、記述・分類・解釈さるべき対象・現象とそのための語彙・参照系・理論枠とは、適切な限定条件下、相互にフーリエ双対 or 圏論的随伴の関係で結ばれ、対象系と記述系との「マッチング」=圏論的普遍性により帰納・演繹の自由な往復が実現する。内部対称性と超選択則の非可換力学系的表現論による定式化と破れた対称性へのその拡張、相対論的量子場の局所非平衡状態の定式化、測定+増幅過程の物理的数学的一般化等々、具体例は数多く、外観の異なる多分野の問題が共通の表現論的・幾何学的・論理的照明の下で比較可能となる。その際、「窮極理論」とは対照的に、限定条件毎に描かれた局所地図=部分理論がネットワークを形作り、「隣り合う」局所地図相互の関係(=近傍貼合せ)が非自明で重要な問題となる。それを扱う方法の核心がスケール概念で、くりこみ群、べき乗則と非加法的統計、階層移行等々、今日的諸課題が目白押しに絡んでくる。双対性とスケール、が担う役割の本質がここにある。この方法論的視点での具体例は理論的内実の豊富化に直結し、間口の拡大と内容深化は互いに双対関係にあって矛盾しない。よって、本研究会を開催する。

関連テーマ

1. 量子場理論：量子場の内部自由度と時空構造，ゲージ構造，非相対論的量子電気力学とその数学的構成，作用素論・作用素環論
2. 統計熱力学：平衡・非平衡統計力学と熱力学，非加法的統計
3. 量子論の数学的基礎と工学的応用：量子情報理論 = 量子通信 + 量子計算，量子論理・「量子集合論」，量子測定論
4. 確率と確率過程：古典・量子確率論，確率過程論，White-Noise Analysis，グラフ・ネットワーク解析
5. ミクロ・マクロ双対性とその方法論：推定理論，制御理論，学習理論

世話人

堀田 昌寛（東北大学），谷村 省吾（京都大学），筒井 泉（高エネルギー加速器研究機構），佐々木 隆（京都大学 基礎物理学研究所），細谷 暁夫（東京工業大学），早川 尚男（京都大学 基礎物理学研究所），森川 雅博（お茶の水女子大学），鹿野 豊（東京工業大学, MIT），小嶋 泉（京都大学 数理解析研究所）

プログラム

敬称略，†=原稿未掲載

11月4日(木)

9:45-9:50 開会と連絡事項

9:50-12:30 座長：西郷 甲矢人（京都大学）

9:50-10:30 岡村 和弥（京都大学） 共同研究者：小嶋 泉

Inverse Problem and Emergence in Large Deviation Strategy

10:30-11:00 松岡 隆志（諏訪東京理科大学）

共同研究者：Dariusz Chruscinski, Andrzej Kossakowski, 大矢 雅則

Quantum Correlation and Mutual Information via Entanglement

11:00-11:20 Coffee Break

11:20-11:50 田中 篤史（首都大学東京） 共同研究者：全 卓樹, Sang Wook Kim

固有値・固有空間のアンホロノミーの階層的構成

11:50-12:20 Andrzej Jamiolkowski (Nicolaus Copernicus University)

Effective methods in investigations of positive maps

12:20-12:30 Poster Session 1 分紹介

12:30-14:00 Lunch

14:00-15:30 Poster Session

15:30-18:00 座長：細谷 暁夫 (東京工業大学)

15:30-16:20 小谷 眞一 (関西学院大学)

佐藤の可積分系理論の1次元ランダムシュレーディンガー作用素

16:20-16:50 Coffee Break

16:50-17:20 Mohammad Ali (近畿大学) 共同研究者：田中 宗, 近藤 康, 中原 幹夫

Hamiltonian Identification for a 3-Spin Chain

17:20-18:00 全 卓樹 (高知工科大学) Egoistic-altruistic duality in quantum game theory

Poster Session

井桁 和浩 (NTT 物性基礎科学研究所) 共同研究者：井元 信之, 小芦 雅斗

† 量子ビットのコヒーレント制御におけるミクロ・マクロ双対性

入山 聖史 (東京理科大学) 共同研究者：大矢 雅則

量子計算の EXPTIME 問題に対する応用と計算量について

作道 直幸 (京都大学) 共同研究者：川上 則雄, 上田 正仁

† キュムラント展開を用いた量子気体の大分配関数の解析

鹿野 豊 (東京工業大学/MIT) 共同研究者：桂 法称

Discrete Time Quantum Walk with Incommensurate Coin

市東 拓郎 (日本大学)

伏見関数と弱い測定

鈴木 理 (日本大学) 共同研究者：堀口 俊 Algebraic description of ternary structure

高江洲 俊光 (九州大学)

準相対論的な粒子と量子場が相互作用する系の enhanced binding について

田中 宗 (近畿大学) 共同研究者：鹿野 豊 weak measurement によるスピン相互作用の推定

富田 博之 (近畿大学) 確率過程における Weak Value とその応用

11月5日(金)

9:45-12:40 座長：鹿野 豊 (東京工業大学/MIT)

9:45-10:20 鈴木 増雄 (東京理科大学)

Duality of Linearity and Nonlinearity to Reveal Irreversibility in Transport Phenomena

10:20-10:50 Yidun Wan (近畿大学)

† Braided Matter of Quantum Geometry:

A meeting point with Quantum Computing/Information

10:50-11:10 Coffee Break

11:10-11:40 佐久間 弘文 ((独) 海洋研究開発機構横浜研究所) 共同研究者：小嶋 泉

共形不変な電磁場放射が満たす“双対”の幾何力学構造について

11:40-12:10 Dariusz Chruscinski (Nicolaus Copernicus University)

Non-Markovian dynamics of quantum systems

12:10-12:40 西郷 甲矢人 (京都大学) 共同研究者：小嶋 泉 Who has seen a free photon?

12:40-14:10 Lunch

14:10-16:20 座長：森川 雅博 (お茶の水女子大学)

14:10-14:50 細谷 暁夫 (東京工業大学) 共同研究者：鹿野 豊
弱値とヒルベルトシュミット束

15:00-15:40 中ノ 勇人 (NTT 物性基礎科学研究所)
†SQUID をプローブとする超伝導量子ビットの状態測定を時間発展過程として理解する

15:40-16:20 大槻 東巳 (上智大学)
2次元系におけるアンダーソン転移と普遍クラスの話

16:20-16:50 Coffee Break

16:50-18:30 座長：小嶋 泉 (京都大学)

16:50-17:40 朽津 和幸 (東京理科大学) 植物の生き方と情報ネットワーク

17:40-18:30 小澤 正直 (名古屋大学) 量子力学における測定・相関・文脈性

19:15- 懇親会 (小澤正直先生還暦記念) at 和洋割烹 Bowsprit/バウスプリット

11月6日(土)

9:50-12:20 座長：堀田 昌寛 (東北大学)

9:50-10:30 沙川 貴大 (東京大学)
共同研究者：鳥谷部 祥一, 宗行 英朗, 佐野 雅己, 上田 正仁
†Principles and Applications of Information Thermodynamics

10:30-11:00 北川 勝浩 (大阪大学) †Heisenberg 限界位相差検出への2つの道

11:00-11:20 Coffee Break

11:20-11:50 飛田 武幸 (名古屋大学) 共同研究者：Si Si 空間・時間・ノイズ

11:50-12:20 市川 翼 (近畿大学) 共同研究者：佐々木 寿彦, 筒井 泉
同種粒子系の量子もつれについて

12:20-14:00 Lunch

14:00-15:40 座長：全 卓樹 (高知工科大学)

14:00-14:30 廣島 文生 (九州大学)
Feynman-Kac type formula with Cadlag path and generalized Schroedinger operator with Spin

14:30-15:00 坂東 将光 (近畿大学) 共同研究者：市川 翼, 近藤 康, 中原 幹夫
On Robust Quantum Control

15:00-15:40 高橋 公也 (九州工業大学) 共同研究者：池田 研介
非インスタントトンネル効果：半古典論的および量子論的解釈

15:40-16:00 Coffee Break

16:50-18:30 座長：谷村 省吾 (京都大学)

16:00-16:40 Ondrej Turek (高知工科大学) 共同研究者：Pavel Exner, 全 卓樹
Vertex couplings in quantum graphs

16:40-17:20 渡辺 優 (東京大学) 共同研究者：沙川 貴大, 上田 正仁
量子推定における不確定性関係

大偏差戦略における逆問題と創発

京都大学数理解析研究所 小嶋 泉* , 岡村 和弥†

1 導入

実験技術の発展と共に量子系制御の問題への関心が高まり、理論・実験双方に互る新たな進展が期待されている。量子系の制御は、系の状態、動力学および外力・外場への応答に関わる知識を不可欠に要求すると同時に、理論的記述に現れる諸概念の操作的意味と理論の適用範囲に対する根本的理解が問われる問題でもある。例えば、望む状態をどのように準備するか？という状態準備の問題も、“状態”概念の理解一つで結果に大きな差が生じる。なぜなら、状態概念は、測定および制御の過程を左右し、系を特徴づける量によって定まる一方で、系の在り方全てを規定づけるわけではないから。この故に、状態概念には常に有限試行・有限精度の制約がつきまとい、その制約下で「最尤状態」(most likely state)の現実的指定を可能にするような方法論が要求される。この要請に応えるには、その前に立ちはだかる障害 = No-Go 定理である次の逆理の解消がまず課題となる。

Duheme-Quine の逆理 [20]: 測定量の試行回数と実験精度の有限性のため、現象データを生み出すような理論を一意に決定することはできない。

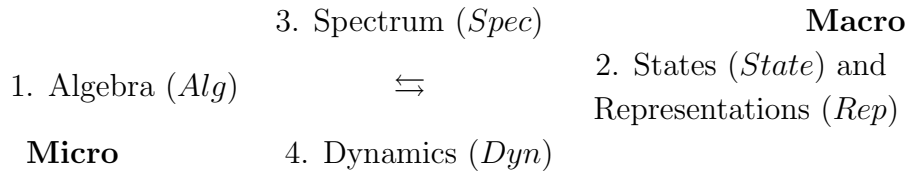
この逆理の意味は、試行回数・精度と無関係に理論の一意性はなく、その制約ゆえ対象とする系に対して、その存在、その在り様について、絶対的なものの言い方をしても意味がない、ということである。これを当たり前の言明として聞き流す方もあれば、拒否する方もあるかもしれないが、我々の解釈を明らかにしておこう。有限試行・有限精度という現実的制約から「試行回数・精度と無関係な」状況はあり得ず、したがって、理論の諸概念に有限試行・有限精度の下での位置づけを適切に与えさえすれば Duheme-Quine の逆理と対峙する場面は回避され、こうした意味で逆理の解消は達成される。それには具体的方法が必要となり、そこで活躍するのが統計学的手法なのである。

その説明のためまず言及しなければならないのが、「マイクロ・マクロ双対性 (Micro-Macro duality)」[19] と「4 項図式 (quadrality scheme)」[20] の概念である。マイクロ・マクロ双対性とは、マイクロとマクロの双方向的・整合的記述に不可欠な双対性と圏論におけるその一般化である随伴を意味すると同時に、それを基礎に展開される方法論を意味する。物理量代数とその上の状態の間の双対性に代表されるように、現象記述のいたるところで双対性は顔を出し、通常 “unknown, invisible” であるマイクロと “known, visible” と見做されるマクロの間との双方向的結びつきとして捉えるべき双対性が様々にある。本稿で中心的役割を果たす物理量代数と状態の間の双対性では、物理量をマイクロ側、状態をマクロ側と捉える (より正確には状態概念は Micro-Macro interface と見るのが最適だが)。物理的には、物理量は対象系に内属するものであるのに対し、状態は実験設定等の外的な影響を被った状況で (不可視のマイクロ量子系をマクロレベルへ可視化する際に) 我々に知りうる範囲を提示するものとして了解される。数学的には、 $(*)$ 代数とその上の規格化された正值線型汎関数という概念間の対形成から生じる往復関係が双対性に他ならない。可換 C^* -代数での

*ojima@kurims.kyoto-u.ac.jp

†kazuqi@kurims.kyoto-u.ac.jp

Gel'fand 表現定理はその代表例である（より詳しくは [19] を参照されたい）。次に，4 項図式の概念の説明に移ろう。以下の図がその 4 項図式である：



基本的発想はミクロ・マクロ双対性に基づく：動力学 (Dyn) (=動き・変化) が自然の根底にあって，条件的に安定な対象 (=モノ) を作り出し，後者が物理量代数 (Alg) で記述される。そして，系の外部との接点を記述するのが 状態 (State) および表現 (Rep) である。状態・表現にはテンソル積と直和による合成演算を定義することができ，それに基づく合成・分解での最小基本単位 (e.g., 既約表現, 因子表現等々) を選び出して，それらの単位を寄せ集める形で分類空間 (Spec) が創発する。ただし，“創発”という言葉の使用には多少注意が必要である。というのは，系の動力的記述を通じて「新たに」創発した自由度が，実は「既知の」時空などと同定すべき場合がしばしば起きるからで，その場合重要なことは，創発した時空自由度に対して物理系がどういう物理的つながりを持つかを明らかにする作業を通じて，理論の出発点では単なる数学的概念としてのみ持ち込まれた「既知の」ad hoc な幾何学的時空概念がその物理的本性を顕わす，という認識である。マクロ側に立つ状態・表現 (State/Rep) と分類空間 (Spec) には，対象系と外界との接点としての機能が備わり，それなしに制御ということもあり得ないわけだが，分類空間のパラメータには，対象系の動力学 (Dyn) と couple することで，本来不可視な動力学過程を可視化し記述可能にする，という重要な働きがある。対象系に元々内属していなかった 1-parameter 時間 $t \in \mathbb{R}$ を用いて時間発展を記述する: $t \mapsto U_t$, ということが現実的物理的意味を獲得するためには，parameter 時間 t の「出自」を物理系内部に求めなければならないはずであろう。

4 項図式の各項につけた番号は，我々にアクセス可能な実験的状況と整合し，量子論の定式化にも則した順序を示すものである。この順序が，上の説明における $Dyn \rightarrow Alg \rightarrow State/Rep \rightarrow Spec$ とずれているのは，(互いに双対な) 認識と存在それぞれの出発点が食い違っていることの理論的表明である。

最後に，動力学について言及したい。数学的抽象的表示を別にすれば，動力学それ自体の現実的記述は，4 項図式の他の 3 項を介して間接的に了解できるかたちでしか実現し得ない。というのは，動力学とは変化そのもの，変化の真っ只中にある概念だから，安定化した存在の形態を指示することによって実現される言語表現には元々馴染まないものである。我々が知る動力学の実際的記述法は，現実に動的に生起した動力学過程の「結果」を初期状態と比較して「後知恵」的になぞるものでしかなく，個々の文脈を選んで限定された側面ごとに記述するしかないのである。物理量による検証を通すため，候補となる動力学モデルは利用した幾つかの物理量のデータを統合した知見により「再」構成されるもので，有限試行・有限精度の制約の下で候補の中から状況に見合うモデルが選択されることとなる。その際，系に内在する動力学そのものとはしばしば同定されがちな，測定データに現れる各物理量の (時) 系列的・動的振る舞いは，各物理量に対応する測定器との相互作用によって擾乱を受けた動力学として了解すべきものである。とすれば，ラグランジアンお

よびハミルトニアンによる記述も動力学のとり得る形態を不変性原理の支配下に制限し，事前 (ad hoc) に限定しているからその範囲でうまく行くのであって，広いように見える適用範囲も，実は，数多くの諸側面を取り落としている（例えば，破壊現象，相転移の渦中，のように）。

統計学は用いる手法の制約が常に目に見える理論だが，(人工とは限らない)「設計」された系の制御を考察する際系の構造を抽出する必要性から，普遍的に適用可能な方法を準備し，構造を特定すべき状況下では非常に有効な概念・結論を提示する。本稿における統計学は，通常の枠組とは異なる視角に立ち，4項図式の各項それぞれを定量的に同定する作業を実行しようとするものである。統計学の中でも大量にデータが得られる場合を扱う大標本理論は，“真”とみなすべき基準への収束レートを問題にする。大偏差原理はこの収束レートを評価する確率論の中心的話題で，本稿のテーマである大偏差戦略は4項図式の番号の順に大偏差原理に基づく統計的推測を行うものである。大偏差戦略については3章において詳しく議論する。本論考の内容に関するより詳細な議論は [22] を参照していただきたい。

2 セクターと量子相対エントロピー

2.1 状態の重心分解，中心分解とセクター

この小節ではセクターおよび状態の重心分解，中心分解を定義し，これらの物理的な意味を論ずる。大偏差戦略第2段階で必要なのはしかしながら，状態概念とミクロ・マクロ双対性の本質に関わる点で本論考において非常に重要な位置を占める。

状態の積分分解を通じて現れるいくつかの作用素環の概念を準備しよう。まず，セクターの概念から。 C^* -代数 \mathfrak{A} に対し， \mathfrak{A} 上の規格化された正值線型汎関数として定義される \mathfrak{A} 上の状態の全体を $E_{\mathfrak{A}}$ で表す。状態 $\omega \in E_{\mathfrak{A}}$ は GNS 表現 $\{\mathfrak{H}_{\omega}, \pi_{\omega}\}$ に対応する von Neumann 代数 $\pi_{\omega}(\mathfrak{A})''$ の中心 $\mathfrak{Z}_{\omega}(\mathfrak{A}) := \pi_{\omega}(\mathfrak{A})'' \cap \pi_{\omega}(\mathfrak{A})' = \mathbb{C}1_{\mathfrak{H}_{\omega}}$ が自明であるとき，ファクター (因子) 状態と呼び，その全体を $F_{\mathfrak{A}}$ と書く。 π を \mathfrak{A} の表現とする： $\pi : \mathfrak{A} \rightarrow B(\mathfrak{H})$ 。 \mathfrak{A} の状態 ω は $\pi(\mathfrak{A})''$ の正規状態 ρ が存在して， $A \in \mathfrak{A}$ に対し

$$\omega(A) = \rho(\pi(A)) \quad (1)$$

と表されるとき π -正規であると呼ばれる。 C^* -代数 \mathfrak{A} の2つの表現 π_1 と π_2 は任意の π_1 -正規状態が π_2 -正規であり，その逆も成立するとき，準同値と呼ばれ， $\pi_1 \approx \pi_2$ と表される。この準同値は多重度を無視したユニタリ同値と等価であることが示されている。以上の準備の下，セクターの概念を定義する。

定義 2.1 ([18]). C^* -代数 \mathfrak{A} のファクター状態の準同値類を \mathfrak{A} のセクターと呼ぶ。

異なるセクター π_1, π_2 間にはゼロでない繋絡作用素 (intertwiner) T が存在しない，すなわち， $T\pi_1(A) = \pi_2(A)T \Rightarrow T = 0$ ($\forall A \in \mathfrak{A}$) であるから，この2つの表現を繋ぐ非対角項が存在せず互いに移りあうことはない。セクターは熱力学的な相概念の一般化になっており，マクロに異なる構造の分類をミクロから生成したものとして扱うことを可能にする。

任意の状態をセクターに分解することが可能で、かつその分解は一意であることが示される。そのための準備を行う。

定義 2.2. $\omega_1, \omega_2 \in \mathfrak{A}_+^*$ が次の等価な 3 つの条件を満たすとき、 ω_1 と ω_2 は直交すると言われ、 $\omega_1 \perp \omega_2$ と表す:

1. $\omega' \in \mathfrak{A}_+^*$ に対し $\omega' \leq \omega_1$ かつ $\omega' \leq \omega_2$ ならば $\omega' = 0$ である;
2. $\omega_1(A) = \langle P\Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle$ かつ $\omega_2(A) = \langle (1 - P)\Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle$ となる射影作用素 $P \in \pi_\omega(\mathfrak{A})'$ が存在する;
3. $\omega = \omega_1 + \omega_2$ に伴う表現は ω_1, ω_2 それぞれに伴う表現の直和となる,

$$\mathfrak{H}_\omega = \mathfrak{H}_{\omega_1} \oplus \mathfrak{H}_{\omega_2}, \quad \pi_\omega = \pi_{\omega_1} \oplus \pi_{\omega_2}, \quad \Omega_\omega = \Omega_{\omega_1} \oplus \Omega_{\omega_2}.$$

定義 2.3. $E_{\mathfrak{A}}$ 上の正則 Borel 測度 μ は任意の Borel 集合 $S \subset E_{\mathfrak{A}}$ に対し、

$$\left(\int_S \rho \, d\mu(\rho) \right) \perp \left(\int_{E_{\mathfrak{A}} \setminus S} \rho \, d\mu(\rho) \right), \quad (2)$$

を満たすとき、 $E_{\mathfrak{A}}$ 上の直交測度と呼ぶ。

次の定理が状態の分解の要となる。

定理 2.1 (冨田分解定理 ([6] 参照)). \mathfrak{A} を C^* -代数とし、 ω を \mathfrak{A} 上の状態とする。次の 3 つの集合に 1 : 1 対応が存在する:

- (1) 直交測度 $\mu \in M_\omega(E_{\mathfrak{A}})$; (2) 可換 von Neumann 代数 $\mathfrak{B} \subseteq \pi_\omega(\mathfrak{A})'$;
- (3) \mathfrak{H}_ω 上の射影作用素 P で、 $P\Omega_\omega = \Omega_\omega$, $P\pi_\omega(\mathfrak{A})P \subseteq \{P\pi_\omega(\mathfrak{A})P\}'$ を満たす。

μ, \mathfrak{B}, P は上の対応があるとき、次を満たす:

- (1) $\mathfrak{B} = \{\pi_\omega(\mathfrak{A}) \cup P\}'$; (2) $P = [\mathfrak{B}\Omega_\omega]$;
- (3) $\mu(\widehat{A}_1 \widehat{A}_2 \cdots \widehat{A}_n) = \langle \Omega_\omega, \pi_\omega(A_1)P\pi_\omega(A_2)P \cdots P\pi_\omega(A_n)\Omega_\omega \rangle$;
- (4) \mathfrak{B} は $\langle \Omega_\omega, \kappa_\mu(f)\pi_\omega(A)\Omega_\omega \rangle = \int d\mu(\rho) f(\rho) \widehat{A}(\rho)$ で定義される写像 $L^\infty(\mu) := L^\infty(E_{\mathfrak{A}}, \mu) \ni f \mapsto \kappa_\mu(f) \in \pi_\omega(\mathfrak{A})'$ の像に $*$ -同型であり、 $A, B \in \mathfrak{A}$ に対し

$$\kappa_\mu(\widehat{A})\pi_\omega(B)\Omega_\omega = \pi_\omega(B)P\pi_\omega(A)\Omega_\omega$$

が成立する。

上の測度 μ は状態 ω の重心測度と呼ばれ、一方、状態 ω を μ の重心 $\omega = b(\mu) := \int_{E_{\mathfrak{A}}} \rho \, d\mu(\rho)$ と呼ぶ。そして、 ω を重心とする $E_{\mathfrak{A}}$ 上の直交確率測度 μ の集合を $\mathcal{O}_\omega(E_{\mathfrak{A}})$ と表す。また、上の定理において、可換 von Neumann 代数 \mathfrak{B} に対応する測度 μ を $\mu_{\mathfrak{B}}$ と表す。

定義 2.4. μ に対応する可換 von Neumann 代数 \mathfrak{B} が ω の GNS 表現 π_ω の中心 $\mathfrak{Z}_\omega(\mathfrak{A})$ の部分代数であるとき、直交測度 $\mu = \mu_{\mathfrak{B}} \in \mathcal{O}_\omega(E_{\mathfrak{A}})$ を ω の準中心測度と呼ぶ。任意の Borel 集合 $\Delta \subset E_{\mathfrak{A}}$ に対し、 π_ω の部分表現の対 $\int_\Delta^\oplus \pi_\rho \, d\mu(\rho)$ と $\int_{E_{\mathfrak{A}} \setminus \Delta}^\oplus \pi_\rho \, d\mu(\rho)$ との間に、ゼロでない繋絡作用素がないという意味で、互いに素である、という条件を準中心測度 μ は満たす。 $\mathfrak{B} = \mathfrak{Z}_\omega(\mathfrak{A})$ のとき、対応する測度を ω の中心測度と呼び、 $\mu_\omega := \mu_{\mathfrak{Z}_\omega(\mathfrak{A})} \in \mathcal{O}_\omega(E_{\mathfrak{A}})$ と表す。

中心測度はファクター状態 $F_{\mathfrak{A}}$ に準台をもつ測度であり，状態をセクターに過不足なく分解する唯一の測度である。

κ_{μ_ω} は $*$ -代数的埋め込み $L^\infty(\mu_\omega)$ であるから，状態空間 $E_{\mathfrak{A}}$ の Borel 部分集合 $\Delta \in \mathfrak{B}(\text{supp } \mu_\omega)$ 上の射影値測度 (PVM) $E_\omega : (\mathfrak{B}(\text{supp } \mu_\omega) \ni \Delta \mapsto E_\omega(\Delta) \in \text{Proj}(\mathfrak{Z}_\omega(\mathfrak{A})))$ を $E_\omega(\Delta) := \kappa_{\mu_\omega}(\chi_\Delta) \in \text{Proj}(\mathfrak{Z}_\omega(\mathfrak{A}))$ によって定義できる。これは，次を満たす：

$$\langle \Omega_\omega, E_\omega(\Delta)\Omega_\omega \rangle = \mu_\omega(\Delta). \quad (3)$$

ここで $E_{\mathfrak{A}}$ の部分集合 Δ に対する定義関数 χ_Δ は通常通り定義される：

$$\chi_\Delta(\rho) = \begin{cases} 1 & (\rho \in \Delta), \\ 0 & (\rho \notin \Delta). \end{cases}$$

以上から， $\text{supp } \mu_\omega$ に属する状態族 $\rho \in \text{supp } \mu_\omega$ を実現値とするような確率変数に対応する測定過程 E_ω が存在することがわかった。こうして，セクターの測定は Born の統計公式から正当化される。また，各元 $\kappa_{\mu_\omega}(f) \in \kappa_{\mu_\omega}(L^\infty(\mu_\omega)) = \mathfrak{B} = \mathfrak{Z}_\omega(\mathfrak{A})$ は

$$\kappa_{\mu_\omega}(f) = \int f(\rho) dE_\omega(\rho). \quad (4)$$

と表示される。それ故， \mathfrak{A} の中心 $\mathfrak{Z}_\omega(\mathfrak{A})$ は状態の非線型関数として与えられた物理量のなす代数として理解できる。

本章の方法を実際的情況に適用する際には，中心測度 μ_ω の台の状態が秩序変数によってパラメータづけられている場合の考察で十分と言ってよい：

$$\omega = \int \rho d\mu_\omega(\rho) = \int_{\Xi} \rho_\xi d\tilde{\mu}(\xi). \quad (5)$$

ここで， $\{\rho_\xi | \xi \in \Xi : \text{秩序変数}\} \subset F_{\mathfrak{A}}$ である。尚，測定過程 [21, 10] を考慮することにより，ここでの方法論の適用領域をセクター内部にまで拡大することが可能である。

2.2 量子相対エントロピーの操作的意味

本節では量子相対エントロピーの性質と操作的意味を議論する。量子相対エントロピーの定義は [5, 13] 等を参照されたい。量子相対エントロピーの記法は論文によって違いがあるが，[13] の記法に従った。

\mathfrak{A} を可分な C^* -代数とし， ψ を \mathfrak{A} 上の状態とする。このとき， $E_{\mathfrak{A}}$ は距離

$$d(\omega_1, \omega_2) = \sum_{j=1}^{\infty} \frac{1}{2^j} \frac{|\omega_1(A_j) - \omega_2(A_j)|}{\|A_j\|}, \quad (6)$$

によって距離付けされる。ただし，集合 $\{A_j \in \mathfrak{A} | A_j \neq 0, j = 1, 2, \dots\}$ は \mathfrak{A} の稠密な部分集合である。この距離の生成する位相は弱 $*$ -位相と一致し， $E_{\mathfrak{A}}$ はコンパクト距離空間となる。したがって $\psi \in E_{\mathfrak{A}}$ の中心測度 μ_ψ の台 $\text{supp } \mu_\psi$ はコンパクトであり，チコノフの定理から $(\text{supp } \mu_\psi)^{\mathbb{N}}$ もまたコンパクトである。 $\tilde{\rho} = (\rho_1, \rho_2, \dots) \in (\text{supp } \mu_\psi)^{\mathbb{N}}$ に対し， $Y_j(\tilde{\rho}) = \rho_j$ と定める。 μ_ψ はファクター状態の集合 $F_{\mathfrak{A}}$ の閉部分集合を台にもつので，各 ρ_j はファクター状態である。このとき， $\{Y_j\}_{j=1}^{\infty}$ は $(\text{supp } \mu_\psi)$ -値確率変数の集合で，次の条件を満たす：

マッチング条件 1. $\{Y_j\}$ は独立同分布 (“i.i.d.”) な確率変数である。

この条件は構成の仕方から数学的には自明だが、物理的には一つの要請である。セクターを独立同分布な確率変数として測定する状況を構成するため、状態推定を大偏差戦略第2段階で定式化し、上記の $\{Y_j\}$ の数学的構成を与えれば、それによって問題を適切に扱う舞台装置が整うことになる。

$M_1(\Sigma)$ により、Polish 空間 Σ 上の Borel 確率測度の空間を表し、 $\mathcal{B}(\Sigma)$ で Σ 上の有界な Borel 可測関数のなすベクトル空間を表す。 $\phi \in \mathcal{B}(\Sigma)$ に対し、 $\tau_\phi : M_1(\Sigma) \rightarrow \mathbb{R}$ を $\tau_\phi(\nu) = \langle \phi, \nu \rangle = \int_\Sigma \phi d\nu$ で定義する。そして、 $\mathcal{B}_{cy}(M_1(\Sigma))$ で $M_1(\Sigma)$ 上の筒状集合のなす σ -加法族、すなわち、全ての $\{\tau_\phi\}$ を可測にする最小の σ -加法族を表す ([8] 参照)。

任意の測度 m に対し、 $P_m := m^{\mathbb{N}}$ で m の可算個の積測度を表す。任意の $\tilde{\rho} \in (\text{supp } \mu_\psi)^{\mathbb{N}}$ 、 $A \in \mathcal{B}(\text{supp } \mu_\psi)$ と $\Gamma \in \mathcal{B}_{cy}(M_1(E_{\mathfrak{A}}))$ に対し、経験測度

$$L_n(\tilde{\rho}, A) = \frac{1}{n} \sum_{j=1}^n \delta_{Y_j(\tilde{\rho})}(A), \quad (7)$$

および

$$Q_n^{(2)}(\Gamma) = P_{\mu_\psi}(L_n \in \Gamma). \quad (8)$$

を定義する。経験測度は測定から得られたデータとしてのセクターから構成した $E_{\mathfrak{A}}$ 上の確率測度で、 $Q_n^{(2)}(\Gamma)$ はデータ数が n のときの経験測度が $E_{\mathfrak{A}}$ 上の確率測度の集合 Γ に入る確率を表す。

相対エントロピー $D(\cdot \|\cdot)$ を次で定義する:

$$D(\nu \|\mu) = \begin{cases} \int d\nu(\rho) \log \frac{d\nu}{d\mu}(\rho) & (\nu \ll \mu) \\ +\infty & (\text{otherwise}). \end{cases} \quad (9)$$

次の定理 [13] は大偏差原理 (LDP) を示す重要な鍵となる:

定理 2.2 (HOT83). μ, ν を $\omega, \psi \in E_{\mathfrak{A}}$ を重心とする $E_{\mathfrak{A}}$ 上の Borel 確率測度とする。 $\mu, \nu \ll m$ となる $E_{\mathfrak{A}}$ 上の準中心測度 m が存在するならば、 $S(\psi \|\omega) = S(b(\nu) \|\mu) = D(\nu \|\mu)$ 。

この定理により、量子相対エントロピー $S(\psi \|\omega)$ を、比較する2つの状態 ω, ψ の重心測度間の測度論的相対エントロピー $D(\nu \|\mu)$ で評価することが可能となる。

定理 2.3. \mathfrak{A} を可分な C^* -代数とし、 ψ を \mathfrak{A} 上の状態とする。このとき、 $Q_n^{(2)}$ は $S(b(\cdot) \|\psi)$ をレート関数とする LDP を満たす。すなわち、任意の $\Gamma \in \mathcal{B}_{cy}(M_1(E_{\mathfrak{A}}))$ に対して、

$$\begin{aligned} - \inf_{\nu \in \Gamma^o, \nu \ll \mu_\omega} S(b(\nu) \|\psi) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log Q_n^{(2)}(\Gamma) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log Q_n^{(2)}(\Gamma) \leq - \inf_{\nu \in \bar{\Gamma}, \nu \ll \mu_\omega} S(b(\nu) \|\psi) \end{aligned} \quad (10)$$

が成立する。ただし、集合 $\{\nu \in \Gamma^o \mid \nu \ll \mu_\omega\}$ もしくは $\{\nu \in \bar{\Gamma} \mid \nu \ll \mu_\omega\}$ が空集合である場合、 $\inf_{\nu \in \Gamma^o, \nu \ll \mu_\omega} S(b(\nu) \|\psi)$ もしくは $\inf_{\nu \in \bar{\Gamma}, \nu \ll \mu_\omega} S(b(\nu) \|\psi)$ の値を ∞ と定める。

この定理は「Sanov の定理」と呼ばれる定理の量子版にあたり、非常に重要な観点を提供する。セクターを測定し、そのデータから状態空間上の測度 (経験測度) をつくり、データ数が無限に増える漸近的状况を考えると、次第に“真”と見做す状態の中心測度へと経験測度は収束する。極限での測度におけるレート関数の値が 0 に近いほど、極限での測度が“真”と見做す状態の中心測度に近いものであることを意味し、このことが収束レートの評価に当たる。その際、“真”と見做す状態の中心測度への収束レートが相対エントロピーで与えられるが、その中心測度から構成される状態および“真”と見做す状態との間の量子相対エントロピーは、レート関数としての中心測度間の相対エントロピーと一致することがわかるのである。こうして、データを用いて状態を構成する、という目的が理想的な状況下で達成されることがわかった。また、データが非常に多いときにしか有効に機能しない経験測度ではなく、統計的推定により得られる状態を用いることにし、その状態と“真”と見做す状態の“近さ”を量子相対エントロピーで評価することが射程に入る。定理 2.3 の設定を振り返ると、今まで“真”と思っていた状態は実際はあくまで仮説的に定めていたに過ぎず、定理 2.3 では本当に“真”と見做す状態を探すことを行っていたのであった。数学的に“真”と見做す状態を現実的にも“真”のものとして扱うのが適切であるという意味で、このような定式化になるのである。統計的推定では立場を変えて、“真”と見做す状態が定まっているものとして、データに基づいてモデルを構成し、その中で最も“真”と見做す状態に近いものを選択する方法を探すことに目的が変わる。

3 大偏差戦略：第 1 段階と第 2 段階を中心に

3.1 大偏差原理から大偏差戦略へ

物理学において、大偏差原理は統計力学がこれまで主たる適用先であった。系の自由度を N としたとき、自由度 N が増加するときの系の状態の収束先の発見およびその収束レートの検証が専らの目的である。古典的な本では [9] が詳しい。これに対し、統計学では N はデータ数 (標本数とも) を意味する。どちらも漠然とは大数の法則が関与する範疇で対象の構造を抽出することに差はなく、その精密化としての大偏差原理はレート関数という概念によってその妥当性の判定をも可能にする。[22, 26] などを眺めると、ゆらぎの中から法則とすべき構造を見出すという類似点からその解析法まで統計力学と統計学は似通ってしまうのは不思議なことではあるが、その恩恵を我々は享受しているのである。この類似点は認識論と数学的構造の観点から非常に興味深い研究対象である。

大偏差戦略は 4 項図式と結びつけて大偏差原理の評価の下で統計的推測を行う枠組である。第 1 段階で物理量代数 (Alg) の各元 (物理量)、第 2 段階で状態 ($State$) および表現 (Rep)、第 3 段階で分類空間 ($Spec$) と代数 (Alg)、そして、第 4 段階では動力学 (Dyn) の推定を行う。本稿では第 1 段階と第 2 段階に焦点を当てる。大偏差原理の評価からはじめて統計的推測による構造決定を各段階では実行する。次節および次々節を眺めることで具体的な内容を把握できるであろう。

3.2 第1段階

第1段階では物理量に焦点を当て、統計的推測を行う。 \mathfrak{A} を C^* -代数、 ψ を \mathfrak{A} 上の状態、そして A を \mathfrak{A} の元である測定対象の物理量とする。 \mathcal{A} で A から生成される \mathfrak{A} の可換部分代数を表す。物理量 A の測定を考察するので、 \mathfrak{A} 上の状態である ψ は \mathcal{A} に自然に制限される。そのため、 $\psi|_{\mathcal{A}}$ を推定する問題へと移行する。ここで、 $\psi|_{\mathcal{A}}$ は ψ の \mathcal{A} へ制限した状態を表す。問題の焦点を変え、特定の可換部分代数 \mathcal{A} に興味がある場合には、 \mathcal{A} の元である A の候補は次の定理から提供される。

定理 3.1 ([6]). 可分 Hilbert 空間 \mathfrak{H} 上の可換 von Neumann 代数 \mathfrak{M} は \mathfrak{M} に属する 1 つの元 X により生成される。

X が自己共役であるとき、 $A = X$ とおく。可換 von Neumann 代数 \mathcal{A} と \mathcal{A} 上の正規状態 ψ に対し、次の関係が成立する：

$$\begin{aligned} \langle \Omega_\psi, \pi_\psi(A) \Omega_\psi \rangle &= \psi(A) = \int \hat{A}(k) d\nu_\psi(k), \\ \pi_\psi(\mathcal{A}) \ni \pi_\psi(A) &\longleftrightarrow \hat{A} \in L^\infty(K, \nu_\psi), \\ \mathfrak{H}_\psi &\cong L^2(K, \nu_\psi) \\ (\mathfrak{H}_\psi \ni \Omega_\psi &\longleftrightarrow 1 \in L^2(K, \nu_\psi)), \\ \mathcal{A}_* &\cong L^1(K, \nu_\psi). \end{aligned}$$

ここで、 K はコンパクト Hausdorff 空間であり、 ν_ψ は K 上の Borel 測度である。上の関係から、 $\pi_\psi(\mathcal{A})$ のどの自己共役元 $\pi_\psi(A)$ も測度論的実確率変数 \hat{A} として扱える。それ故、可換部分代数上では物理量のスペクトルを扱うことが可能となる。可換部分に制限する理由を物理的に正確に意味づけることは難しいが、数学的にはこれらの構造を用いて測度論的確率論の構造を利用することが可能となる。

$\bar{k} = (k_1, k_2, \dots) \in K^{\mathbb{N}}$ と $A = A^* \in \mathcal{A}$ に対し、 $X_j(\bar{k}) = k_j$ および $\hat{A}_j(\bar{k}) := \hat{A}(X_j(\bar{k}))$ と定めると、次の条件の妥当性がわかる：

マッチング条件 2. $\{\hat{A}_j\}$ は独立同分布 (“i.i.d.”) な確率変数である。

この条件が物理的に満たされれば、大偏差原理の評価およびデータを用いた確率分布の推定を行うことが可能となる。ここからが大偏差戦略第1段階の議論となる。まずは大偏差原理の評価で、次の定理が成立する：

定理 3.2 (Cramér の定理 [8]). $M_n(\bar{k}) := \frac{\hat{A}_1(\bar{k}) + \dots + \hat{A}_n(\bar{k})}{n}$, $Q_n^{(1)}(\Gamma) := P_{\nu_\psi}(M_n \in \Gamma)$ とする。 $Q_n^{(1)}$ は $I_\psi(a) = \sup_{t \in \mathbb{R}} \{at - c_\psi(t)\}$ をレート関数とする LDP を満たす

$$\left(c_\psi(t) = \log \int_{\mathbb{R}} e^{tx} \nu_\psi(\hat{A} \in dx) \right):$$

$$- \inf_{a \in \Gamma^\circ} I_\psi(a) \leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log Q_n^{(1)}(\Gamma) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log Q_n^{(1)}(\Gamma) \leq - \inf_{a \in \bar{\Gamma}} I_\psi(a) \quad (11)$$

この定理から，物理量の相加平均がその“真”の平均値に収束するときの収束レートが議論可能となる。なお，平均値の推定の議論は次の確率分布の推定の議論の帰結の系として扱えるため，ここでは省略する。では，確率分布の推定の議論に移る。大偏差原理の評価は Sanov の定理 [8] で与えられ，相対エントロピーに操作的意味が与えられる。次に，モデルの定義を行おう。

定義 3.1. \mathbb{R}^d 上の確率分布の族 $\{p(x|w)|w \in W\}$ (W は \mathbb{R}^l のコンパクト部分集合) は， $\{x \in \mathbb{R}^d | p(x|w) > 0\}$ が $w \in W$ に依らないとき (統計的) モデルと呼ばれる。

モデルを構成する際，最低限要請されるべきことは，物理的に想定される性質を適切に取り込んだ構成を見つけることである。ただし，モデルは物理的に1つだけしか作れない理由などないのでいくつでも作ってよい。制御理論などでも扱われる，データに用いる物理量とモデルで利用する物理量が異なる場合が考察できる。この場合に代表されるように，測定する範囲を超えたモデルが必要とされる状況が現実であり，この要請が満たされていけば，あとは用いる数学の手法にあわせた数学的要請をモデルが満たしていれば良い。これらは第2段階で用いる (非可換) モデルについても当てはまる事項であることを先走って述べておく。 $\pi(w)$ を W 上の確率密度関数 (事前分布) とし， $0 < \beta < \infty$ とする。

定義 3.2. 以下で定義される確率分布 $p_{\pi,\beta}(x|x^n)$ を Bayes エスコート予測分布と呼ぶ：

$$p_{\pi,\beta}(x|x^n) = \langle p(x|w) \rangle_{\pi,\beta}^{x^n} = \frac{\int p(x|w) \prod_{j=1}^n p(x_j|w)^\beta \pi(w) dw}{\int \prod_{j=1}^n p(x_j|w)^\beta \pi(w) dw}. \quad (12)$$

この Bayes エスコート予測分布は予測分布の一種で，事前分布 $\pi(w)$ と β の選択に依存する。しかしながら，特定の性質が満たされていけば，データ数が増えていく漸近的な状況下では事前分布の違いによる影響は小さくなることが知られている。Bayes エスコート予測分布は次のリスク関数を最小にする確率分布であることがわかる。

定理 3.3. Bayes エスコート予測分布は次のリスク関数 $\mathcal{R}^n(p||\cdot) : r(\cdot|x^n) \rightarrow \mathcal{R}^n(p||r)$ を最小にする：

$$\mathcal{R}^n(p||r) = \iint D(p(\cdot|w)||r(\cdot|x^n)) \prod_{j=1}^n p(x_j|w)^\beta dx_j \pi(w) dw. \quad (13)$$

構成したモデルのうち，どれが最も適切であるか？という議論などは本質的には第2段階での手法と同じである (手法の開発は，統計学および学習理論，すなわち第1段階の方が先)。それ故，ここでは議論せず，第2段階に譲る。

3.3 第2段階

状態 (State) ・ 表現 (Rep) の推定を行うのが第2段階である。前章のセクター・中心測度を用いると，状態の推定が可能となる。このとき，大偏差原理には定理 2.3 が自然に採用される。以上の準備の下，モデルを定義し，状態の推定を行う。

定義 3.3. \mathbb{R}^d のコンパクト部分集合 Θ でパラメトライズされる状態の族 $\{\omega_\theta | \theta \in \Theta\}$ は次の β 条件を満たすとき (統計的) モデルと呼ばれる :

- (i) $\theta \in \Theta$ に対し, $\mu_{\omega_\theta} \ll m$ となる $E_{\mathfrak{A}}$ 上の準中心測度 m が存在する。
- (ii) 集合 $\left\{ \rho \in E_{\mathfrak{A}} \mid p_\theta := \frac{d\mu_{\omega_\theta}}{dm}(\rho) > 0 \right\}$ は $\theta \in \Theta$ に依らない。
- (iii) ω_θ は Bochner 可積分である。

状態のレベルでモデルを構築する理由を述べよう。状態の指定により様々な物理量のスペクトルと確率分布が考察可能になることがある一方, 状態の指定なしに特定の確率分布の指定だけでは系を記述し切れない。けれども, 一切の仮定なく状態を指定することは不可能なので, 望まれる性質を満たすパラメータ付きの状態族を想定し, パラメータの推定・調整を通じて状態を指定するという過程を経ることが必要になる。

$\rho^n = \{\rho_1, \dots, \rho_n\}$ をデータ, $\pi(\theta)$ を Θ 上の確率密度関数, そして, $0 < \beta < \infty$ とする。

定義 3.4. 与えられたモデル $\{\omega_\theta\}_{\theta \in \Theta}$ に対し,

$$\omega_{\pi, \beta}^n := \frac{\int \omega_\theta \prod_{j=1}^n p_\theta(\rho_j)^\beta \pi(\theta) d\theta}{\int \prod_{j=1}^n p_\theta(\rho_j)^\beta \pi(\theta) d\theta} \quad (14)$$

で定義される状態 $\omega_{\pi, \beta}^n$ を Bayes エスコート予測状態と呼ぶ。

この Bayes エスコート予測状態が Bayes エスコート予測分布の一般化であることは容易にわかる。

Θ 上の与えられた可測関数 $G(\theta)$ に対し, $G(\theta)$ の事後平均を次で定める。

$$\langle G(\theta) \rangle_{\pi, \beta}^{\rho^n} = \frac{\int G(\theta) \prod_{j=1}^n p(\rho_j | \theta)^\beta \pi(\theta) d\theta}{\int \prod_{j=1}^n p(\rho_j | \theta)^\beta \pi(\theta) d\theta}. \quad (15)$$

ただし, $0 < \beta < \infty$ である。このとき, 次の等式が成り立つ :

$$\omega_{\pi, \beta}^n = \int \rho \langle p(\rho | \theta) \rangle_{\pi, \beta}^{\rho^n} dm(\rho). \quad (16)$$

ここで2つの期待値 E_ρ および E_{ρ^n} を定義する : $E_{\mathfrak{A}}$ 上の可測関数 $F(\rho)$ に対し,

$$E_\rho[F(\rho)] = \int H(\rho) q(\rho) dm(\rho). \quad (17)$$

$(E_{\mathfrak{A}})^n$ 上の可測関数 $G(\rho_1, \dots, \rho_n)$ に対し,

$$E_{\rho^n}[G(\rho_1, \dots, \rho_n)] = \int G(\rho_1, \dots, \rho_n) \prod_{j=1}^n q(\rho_j) dm(\rho_j). \quad (18)$$

統計学及び学習理論の主要な計算及び推定の対象である概念を次に定義する。

定義 3.5.

(1) Bayes 汎化誤差 \mathcal{E}_{bg} および Bayes 汎化損失 \mathcal{L}_{bg} を以下で定める：

$$\mathcal{E}_{bg} = E_{\rho} \left[\log \frac{q(\rho)}{\langle p(\rho|\theta) \rangle_{\pi,\beta}^{\rho^n}} \right], \quad \mathcal{L}_{bg} = E_{\rho} \left[-\log \langle p(\rho|\theta) \rangle_{\pi,\beta}^{\rho^n} \right]. \quad (19)$$

(2) Bayes 訓練誤差 \mathcal{E}_{bt} および Bayes 訓練損失 \mathcal{L}_{bt} を以下で定める：

$$\mathcal{E}_{bt} = \frac{1}{n} \sum_{j=1}^n \left[\log \frac{q(\rho_j)}{\langle p(\rho_j|\theta) \rangle_{\pi,\beta}^{\rho^n}} \right], \quad \mathcal{L}_{bt} = \frac{1}{n} \sum_{j=1}^n \left[-\log \langle p(\rho_j|\theta) \rangle_{\pi,\beta}^{\rho^n} \right]. \quad (20)$$

(3) 汎関数分散 \mathcal{V} を以下で定める：

$$\mathcal{V} = \sum_{j=1}^n \left\{ \langle (\log p(\rho_j|\theta))^2 \rangle_{\pi,\beta}^{\rho^n} - (\langle \log p(\rho_j|\theta) \rangle_{\pi,\beta}^{\rho^n})^2 \right\}. \quad (21)$$

次の関係は容易にわかる：

$$\begin{aligned} \mathcal{E}_{bg} &= D(q \| \langle p(\cdot|\theta) \rangle_{\pi,\beta}^{\rho^n}) = S(\psi \| \omega_{\pi,\beta}^n) \\ &= \mathcal{L}_{bg} + E_{\rho} [\log q(\rho_j)], \\ \mathcal{E}_{bt} &= \mathcal{L}_{bt} + \frac{1}{n} \sum_{j=1}^n \log q(\rho_j). \end{aligned} \quad (22)$$

ここで、共通の測度に対する密度関数 $\langle p(\rho|\theta) \rangle_{\pi,\beta}^{\rho^n}$ と $q(\rho)$ との間の相対エントロピーを $D(q \| \langle p(\cdot|\theta) \rangle_{\pi,\beta}^{\rho^n})$ で表した。Bayes 汎化誤差は“真”の状態 ψ と Bayes エスコート予測状態 $\omega_{\pi,\beta}^n$ の間の量子相対エントロピーであり、Bayes エスコート予測状態 $\omega_{\pi,\beta}^n$ の性能を相対エントロピーにより評価することを目的として構成された量である。しかしながら、実際には“真”の状態 ψ を我々は知りようがなく、Bayes 汎化誤差あるいは Bayes 汎化損失は計算できない。故に、“真”の状態 ψ に依らない量を用いて Bayes 汎化誤差および Bayes 汎化損失を推定もしくは近似する必要が生じる。このとき用いるのが“真”の状態 ψ に依らず、モデルとデータ ρ^n に依存した Bayes 訓練損失であって、それに合わせ Bayes 汎化損失を Bayes 訓練損失で推定することを考える。次の定理が最も望む結果を与える定理である。

定理 3.4.

$$E_{\rho^n}[\mathcal{L}_{bg}] = E_{\rho^n}[\text{WAIC}] + o\left(\frac{1}{n}\right), \quad (23)$$

$$\text{WAIC} = \mathcal{L}_{bt} + \frac{\beta}{n} \mathcal{V}. \quad (24)$$

この定理の証明には解析性などのいくつかの重要な仮定が本来必要である。この定理の主張は、Bayes 汎化損失のデータが増えていく状況下の平均 ((23) の左辺) と Bayes 訓練損

先に補正項を加えたものの平均 ((23) の右辺および (24)) が $o\left(\frac{1}{n}\right)$ のオーダーのずれの範囲内で一致する, というものである。補正項は汎関数分散 ν の $\frac{\beta}{n}$ 倍で与えられる。(23) の右辺で平均を取っている項 (24) には WAIC, widely applicable information criteria の頭文字, と名づけられており, 情報量規準の一種である。この WAIC を最小化するモデルが最も “真” の状態に近いモデルであって, β もこの WAIC の最小化を通して定められる。また, “真” の状態の明確な実体もこの段階に至ってようやく実感されるものである。第 2 段階で議論した内容をまとめると,

$$\begin{aligned} \text{量子相対エントロピー } S(\psi||\omega) &\Rightarrow \text{Bayes エスコート予測状態 } \omega_{\pi,\beta}^n \\ &\Rightarrow \text{WAIC } (\Rightarrow \text{“真” の状態}) \end{aligned}$$

という構図が見えてくる。この構図は

$$\text{レート関数} \Rightarrow \text{モデル} \Rightarrow \text{評価基準} (\Rightarrow \text{“真” の構造})$$

という構図の大偏差戦略第 2 段階という一形態であるとみなすのが自然であり, 第 1 段階も第 2 段階と同様の構図に従う。大偏差戦略の第 3 段階・第 4 段階でも同様の構図が発見されると期待される。

4 展望

統計学の知見および手法を深める作業は今後も絶えず必要であり, 逃れられない有限の試行回数と実験精度の制約を引き受ける役割を負うことのできる理論に統計学がなるか否かは今後の課題である。本稿において, 有限試行・有限精度の制約下で量子論の諸概念の意味を深められたことは, この課題達成を目指す上できわめて大きな成果であった。特に, 大偏差戦略の第 1 段階および第 2 段階では統計学の手法の効用が遺憾なく発揮された。今後は, これまでの統計学に囚われない手法を導入することにより, 分類空間の同定を行う第 3 段階および動力学の推定を行う第 4 段階の定式化が望まれる。また, 量子推定理論 [12, 15] としてこれまで研究されてきた量子系における統計的推測の手法を今一度見直し, 普遍的方法論と系に固有の特徴を利用した方法論とに分類・整理し, 双方を発展させる時期が間違いなく到来している。この観点から我々の提起する量子仮説検定論 [22] と [14, 16, 11] で提示された量子仮説検定論とを比較することは興味深い。量子推定理論の今後の発展に期待したい。

参考文献

- [1] J. Aitchison, Goodness of prediction fit, *Biometrika* **62**, 547-554 (1975).
- [2] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: B.N. Petrov and F. Csaki, eds., 2nd international symposium on information theory (Akademiai Kiado, Budapest, 1973).

- [3] H. Akaike, IEEE Trans. Automatic Control **19**, 716-723 (1974).
- [4] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of mathematical monographs; v. 191, Amer. Math. Soc. & Oxford Univ. Press (2000).
- [5] H. Araki, Relative entropy for states of von Neumann algebras II, Publ. Res. Inst. Math. Sci. **13**, 173-192 (1977).
- [6] O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (vol. 1), Springer-Verlag (1979).
- [7] D. Buchholz, I. Ojima and H. Roos, Ann. Phys. (N.Y.) **297**, 219-242 (2002).
- [8] A. Dembo and O. Zeitouni, *Large deviations techniques and applications* (2nd ed.), (Springer, 2002).
- [9] R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, (Springer,1985).
- [10] R. Harada and I. Ojima, Open Sys. Inform. Dyn. **16**, 55-74 (2009).
- [11] *Asymptotic Theory of Quantum Statistical Inference*, edited by M. Hayashi (World Scientific, Singapore, 2005).
- [12] C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).
- [13] F. Hiai, M. Ohya and M. Tsukada, Pacific J. Math. **107**, 117-140 (1983).
- [14] F. Hiai and D. Petz, Commun. Math. Phys. **143**, 99-114 (1991).
- [15] S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
- [16] T. Ogawa and H. Nagaoka, IEEE Trans. Inform. Theory **46**, 2428-2433 (2000).
- [17] I. Ojima, RIMS Kokyuroku **1066** (1998), 121-132.
- [18] I. Ojima, Open Sys. Inform. Dyn. **10**, 235-279 (2003).
- [19] I. Ojima, "Micro-Macro Duality in Quantum Physics", pp.143-161 in Proc. Intern. Conf. on Stochastic Analysis, Classical and Quantum (World Scientific, 2005), arXiv:math-ph/0502038.
- [20] I. Ojima, J. Phys.: Conf. Ser. **201**, 012017 (2010).
- [21] I. Ojima, Space(-Time) Emergence as Symmetry Breaking Effect, Invited talk at International Conference in QIBC (= Quantum Bio-Informatics Center, Tokyo University of Sciences) 2010; Dilation and Emergence in Physical Sciences, Invited talk at International Conference, "Advances in Quantum Theory" at Linnaeus University, June 2010.
- [22] I. Ojima and K. Okamura, arXiv:1101.3690.
- [23] M. Takesaki, *Theory of Operator Algebras II*, (Springer, 2002).
- [24] F. Tanaka and F. Komaki, Phys. Rev. A **71**, 052323 (2005).
- [25] K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto and K. Nakamura, Phys. Rev. A **68**, 022314 (2003).
- [26] S. Watanabe, *Algebraic geometry and statistical learning theory*, (Cambridge University Press, 2009).
- [27] S. Watanabe, J. Phys.: Conf. Ser. **233**, 012014 (2010).
- [28] S. Watanabe, J. Mach. Learn. Res. **11**, 3571-3594 (2010).

**QUANTUM CORREALTION AND MUTUAL ENTROPY
VIA
ENATANGLEMENT**

TAKASHI MATSUOKA

*Faculty of Management of Administration and Information
Tokyo University of Science, Suwa
Chino City, Nagano 391-0292, Japan
E-mail: matsuoka@rs.swu.tus.ac.jp*

**Joint work with D. Chruściński, Y. Hirota, A. Kossakowski,
and M. Ohya**

We study a quantum correaltions from the information theoretical points of view. A quantum mutual entropy is discueeed in relation to a transmitted infromation. This report is the review of our recent works on the quantum infromation theory with entanglement.

1. Introduction

In classical probability theory the classical correlation of a physical compound system can be represented by a joint probability measure with marginal probability measures which are corresponding to each subsystem. In quantum case a quantum state of a composite system, i.e. a compound density operator describes a correlation between its marginal states. However, in quantum system, it is known that the joint probability and the conditional probability do not generally exist, which is an essential difference from classical system^{32,37}. The typical example of such difference is the existence of entangled states in quantum system.

An entangled state of two quantum system is defined by a state not written as a form

$$\sum_k \lambda_k \rho_k \otimes \sigma_k, \quad (1)$$

with any marginal states ρ_k and σ_k on subsystems. A state written as above is called a separable state, so that an entangled state is a state not belonged to the set of all separable states in the naive definition above.

However the mathematical characterization of entanglement is not yet fully understood except for some simple cases, for example pure states or very low dimensional mixed states. Further it is obvious that there exist several types of correlated states, written as separable forms. Such correlated states have been considered in several contexts in the fields of quantum information and quantum probability such as quantum measurement and filtering^{9,10}, quantum compound states^{30,31} and lifting^{1,4}. So that one have to deal with not only the correlation of entangled state but also that of separable state carefully.

In the middle of 1990s Peres and Horodecki family gave a classification of compound states by introducing the criteria^{34,22} which enable one to detect the entangled state. The famous Peres-Horodecki criterion is based on the partial transposition. States which are positive under partial transposition are called PPT sates. Clearly each separable state is necessarily PPT but the converse is not true. On the other hand Belavkin and Ohya gave a rigorous construction of quantum compound states by means of a complete co-positive map (CCP map for short) called entanglement map^{11,12}. We can construct a compound state by applying the entanglement map to Choi-Jamiołkowski isomorphism^{15,23,25,39}. Kossakowski et al. also introduced another method to construct a compound state via a notion of quantum conditional probability operator, QCPO for short, by means of a complete positive map (CP map for short)⁵. If a marginal states is given, then a compound state is constructed by operating the marginal state by QCPO. The PPT sate can be characterized by the entanglement map, also by the QCPO^{24,26,5}.

This paper is organized as follows: In the section 2 the definitions of above criteria, PPT condition, entanglement map and QCPO, are reviewed and we show some relations among such three types of criterion^{5,18,24,26}.

The section 3 we review the two types of quantum entropy, so called the quantum mutual entropy and the quantum conditional entropy^{11,12,14,19,21}, and such entropies are applied to some examples of compound states. In addition we introduce the order of quantum correlation by using the symmetrized conditional entropy called the degree of entanglement (DEN for short)^{18,28}. Interestingly it turned out that there exists a class of separable states which have stronger correlation (with respect to DEN) than entangled states¹⁶. Finally we discuss the difference between classical and quantum correlations in the section 4.

The above argument can be formulated at general C*-algebraic setting^{11,12,24,26,27}. However, for simplicity, only the finite-dimensional case

will be considered here.

2. PPT state, entanglement map and quantum conditional operator

We start with notions of positivity: the so-called complete positivity (CP for short) and complete co-positivity (CCP for short). Let \mathcal{A} and \mathcal{B} be C^* -algebras (with a unit 1). A linear map $\chi : \mathcal{A} \mapsto \mathcal{B}$ is CP if and only if

$$\chi_n : M_n(\mathcal{A}) \rightarrow M_n(\mathcal{B}); [a_{ij}] \mapsto [\chi(a_{ij})] \quad (2)$$

is positive for all n . Here $M_n(\mathcal{A})$ stands for $n \times n$ matrices with entries from \mathcal{A} . A linear map $\chi : \mathcal{A} \mapsto \mathcal{B}$ is CCP if and only if

$$\chi_n^t : M_n(\mathcal{A}) \rightarrow M_n(\mathcal{B}); [a_{ij}] \mapsto [\chi(a_{ji})] \quad (3)$$

is positive for all n . Every n -co-positive map is positive but not necessarily CP even if it is CCP, unless \mathcal{B} is Abelian, in which case a positive map is both CP and CCP.

Let us consider two quantum systems $(\mathcal{H}_k, \mathcal{A}_k, \mathcal{S}(\mathcal{A}_k))_{k=1,2}$ where \mathcal{H}_k is a Hilbert space given by $\mathcal{H}_k = \mathbb{C}^n$, $\mathcal{A}_k = \mathcal{B}(\mathcal{H}_k)$ stands for the algebra of complex $k \times k$ matrices and $\mathcal{S}(\mathcal{A}_k)$ stands for the set of all states on \mathcal{A}_k . Then the composite system $(\mathcal{H}, \mathcal{A}, \mathcal{S})$ is given by

$$\mathcal{H} \equiv \mathcal{H}_1 \otimes \mathcal{H}_2, \mathcal{A} \equiv \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) = \mathcal{A}_1 \otimes \mathcal{A}_2, \mathcal{S} \equiv \mathcal{S}(\mathcal{A}) = \mathcal{S}(\mathcal{A}_1 \otimes \mathcal{A}_2).$$

Notice that

$$\mathcal{S} \supset \text{conv}(\mathcal{S}(\mathcal{A}_1) \otimes \mathcal{S}(\mathcal{A}_2)).$$

It was mentioned that $\text{conv}(\mathcal{S}(\mathcal{A}_1) \otimes \mathcal{S}(\mathcal{A}_2))$ are called separable states and the subset of states $\mathcal{S} \setminus \text{conv}(\mathcal{S}(\mathcal{A}_1) \otimes \mathcal{S}(\mathcal{A}_2))$ is called the set of entangled states, i.e. $\mathcal{S} = \mathcal{S}_{\text{ent}} \cup \mathcal{S}_{\text{sep}}$. We are interested in the special subset of \mathcal{S} ,

$$\mathcal{S}_{PPT} \equiv \{\omega \in \mathcal{S}; \omega \circ (1 \otimes \tau) \in \mathcal{S}\}, \quad (4)$$

where τ is stands for transposition. Such states are called partial positive transposition (PPT) states^{22,34}. Clearly

$$\mathcal{S} \supset \mathcal{S}_{PPT} \supset \mathcal{S}_{\text{sep}}.$$

In general the PPT condition is not a sufficient one for the separability of state. However, in the case of \mathcal{H} identified to $\mathbb{C}^2 \otimes \mathbb{C}^2$ or $\mathbb{C}^2 \otimes \mathbb{C}^3$ the PPT condition becomes to the sufficient one²².

Theorem 2.1. *In the case of $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ or $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^3$ a state ω is separable iff ω is belonged to \mathcal{S}_{PPT} .*

We will review another criteria called an entanglement map^{11,12} and a quantum conditional probability operator (QCPO for short)⁵.

For a normal state $\omega \in \mathcal{S}$ there exist a density matrix θ such that

$$\omega(a \otimes b) = \text{Tr}(a \otimes b)\theta, a \in \mathcal{A}_1, b \in \mathcal{A}_2. \quad (5)$$

The state ω can be written as

$$\omega(a \otimes b) = \text{Tr}_1 a \phi(b) = \text{Tr}_2 \phi^*(a) b \quad (6)$$

where ϕ is a linear map from \mathcal{A}_2 to \mathcal{A}_1 given by $\phi(b) \equiv \text{Tr}_2(1 \otimes b)\theta \in \mathcal{A}_1$, and its dual map ϕ^* from \mathcal{A}_1 to \mathcal{A}_2 given by $\phi^*(a) \equiv \text{Tr}_1(a \otimes 1)\theta \in \mathcal{A}_2$. In (6) the partial trace with respect to \mathcal{H}_k is denoted by $\text{Tr}_k(\cdot)$. It is clear that the marginal densities of θ are given by

$$\rho \equiv \text{Tr}_2 \theta = \phi(1), \sigma \equiv \text{Tr}_1 \theta = \phi^*(1). \quad (7)$$

Belavkin and Ohya revealed that such maps can be reconstructed by the Hilbert-Schmidt operator, which is called the entangling operator, and they showed that both ϕ and ϕ^* are CCP, but not always CP on general C*-algebraic setting^{11,12}. For example, if ω is a pure entangled state, then its entanglement map is not CP.

On the base of above observation they introduced a general notion of quantum compound state by using such CCP maps.

Definition 2.1. A CCP map $\phi: \mathcal{A}_2 \rightarrow \mathcal{A}_1$, normalized as $\text{Tr}_1 \phi(1) = 1$, is called the (generalized) entanglement (map) from the state $\phi^*(1) \in \mathcal{S}(\mathcal{A}_2)$ to the state $\phi(1) \in \mathcal{S}(\mathcal{A}_1)$.

The density operator θ_ϕ corresponding to the entanglement ϕ with its marginal $\rho = \phi(1)$ and $\sigma = \phi^*(1)$ can be constructed by

$$\theta_\phi = \sum_{i,j} \phi(|e_j\rangle\langle e_i|) \otimes |e_i\rangle\langle e_j| \quad (8)$$

where $\{e_i\}$ is a standard base in \mathcal{H}_2 ^{11,12}.

Notice that the construction of θ_ϕ in (8) is given via the Choi-Jamiołkowski isomorphism^{15,23,25,39}. Due to this fact we can show the following^{24,26}.

Theorem 2.2. A state θ_ϕ (i.e., a density operator) is a PPT state iff its entanglement map ϕ is a CP map.

In the following discussion we use the symbol \mathcal{S} also as a set of density operators.

On the other hand Kossakowski et al. gave another construction of compound density operators⁵. For a given $\theta \in \mathcal{S}$ with $\rho = \text{Tr}_2 \theta$ one can define the operator

$$\pi_\theta \equiv \left(\rho^{-\frac{1}{2}} \otimes 1 \right) \theta \left(\rho^{-\frac{1}{2}} \otimes 1 \right), \quad (9)$$

which has the following properties

$$\pi_\theta \geq 0, \quad (10)$$

$$\text{Tr}_2 \pi_\theta = 1 \in \mathcal{A}_1, \quad (11)$$

and we assume that $\rho > 0$, that is, ρ represents a faithful state. It follows from (10) and (11) that the operator π_θ is the quantum analogue of a classical conditional probability distribution. Indeed, in the case \mathcal{A} is an Abelian algebra, π_θ coincides with a classical conditional probability.

Definition 2.2. An operator $\pi \in \mathcal{A}$ is called a quantum conditional probability operator (QCPO) if π satisfies condition (10) and (11).

One easily finds the following formula

$$\pi_\varphi = \sum_{k,l=1}^n \varphi(|e_k\rangle\langle e_l|) \otimes |e_k\rangle\langle e_l| \quad (12)$$

where φ is a CP unital map from \mathcal{A}_2 to \mathcal{A}_1 , i.e., $\varphi(1) = 1$. It is easy to check that π_φ satisfies conditions (10) and (11). For a given π_φ and any marginal state $\rho \in \mathcal{S}(\mathcal{A}_1)$ one can construct a compound state as

$$\theta_\varphi = \sum_{k,l=1}^n \rho^{\frac{1}{2}} \varphi(|e_k\rangle\langle e_l|) \rho^{\frac{1}{2}} \otimes |e_k\rangle\langle e_l|. \quad (13)$$

Applying the CP map $\rho^{\frac{1}{2}} \varphi(\cdot) \rho^{\frac{1}{2}}$ to the Choi-Jamiolkowski isomorphism, then we have

$$(1 \otimes \tau) \theta_\varphi = \sum_{k,l=1}^n \rho^{\frac{1}{2}} \varphi(|e_l\rangle\langle e_k|) \rho^{\frac{1}{2}} \otimes |e_k\rangle\langle e_l|. \quad (14)$$

This means the following Corollary⁵.

Corollary 2.1. θ_φ is a PPT state iff the map φ is a CCP map.

There exists a relation between the density operator θ_ϕ in (8) and the QCPO π_φ in (12) via the following decomposition of the entanglement map ϕ .

Lemma 2.1.¹³ Every entanglement map ϕ with $\phi(1) = \rho$ has a decomposition

$$\phi(\cdot) = \rho^{\frac{1}{2}} \varphi \circ \tau(\cdot) \rho^{\frac{1}{2}}, \quad (15)$$

where φ is a CP unital map to be found as a unique solution to

$$\varphi(\cdot) = \rho^{-\frac{1}{2}} \phi \circ \tau(\cdot) \rho^{-\frac{1}{2}}. \quad (16)$$

Theorem 2.3.¹⁸ If a compound state θ_ϕ given by (8) has a faithful marginal state $\rho = \phi(1)$, then θ_ϕ is represented by

$$\theta_\phi = \left(\rho^{\frac{1}{2}} \otimes 1 \right) \pi_\phi \left(\rho^{\frac{1}{2}} \otimes 1 \right) \quad (17)$$

where $\pi_\phi = \sum_{k,l=1}^n \rho^{-\frac{1}{2}} \phi(|e_l\rangle\langle e_k|) \rho^{-\frac{1}{2}} \otimes |e_k\rangle\langle e_l|$.

The relation (17) can be regarded as a quantum analogue of the relation between a joint probability and a conditional probability in classical probability theory:

$$r_{ij} = p_i p(j|i), \quad (18)$$

where r_{ij} is a joint probability with its marginal probability p_i and $p(j|i)$ is a conditional probability i.e., the probability of an event "j" under the observation of an event "i".

3. Mutual entropy and conditional entropy vs. quantum correlation

In classical information theory the mutual entropy is given as the relative entropy of the correlated joint probability and the non-correlated one. We extend the classical mutual entropy to a quantum system by using the Umegaki relative entropy³⁶ in the terminology of entanglement map ϕ .

Definition 3.1. Let ϕ be an entanglement map with $\rho = \phi(1)$ and $\sigma = \phi^*(1)$. One defines^{11,12,14,19}

$$\begin{aligned} I_\phi(\rho, \sigma) &\equiv S(\theta_\phi, \rho \otimes \sigma) \\ &= \text{Tr} \theta_\phi (\log \theta_\phi - \log \rho \otimes \sigma) \end{aligned} \quad (19)$$

where $S(\cdot, \cdot)$ is the Umegaki relative entropy. One calls $I_\phi(\rho, \sigma)$ the quantum mutual entropy^{14,19}.

By simple computation the quantum mutual entropy can be decomposed as follows:

$$I_\phi(\rho, \sigma) = S(\rho) + S(\sigma) - S(\theta_\phi). \quad (20)$$

The above relation gives the same decomposition form of the classical mutual entropy. Using this decomposition (20) one also defines the quantum conditional entropies^{11,12,14,21}:

$$S_\phi(\sigma|\rho) \equiv S(\sigma) - I_\phi(\rho, \sigma) = S(\theta_\phi) - S(\rho), \quad (21)$$

$$S_\phi(\rho|\sigma) \equiv S(\rho) - I_\phi(\rho, \sigma) = S(\theta_\phi) - S(\sigma). \quad (22)$$

Remark 3.1. The above entropic criteria are discussed also by Cerf and Adami¹⁴, Horodeckis²¹, Henderson and Vedral²⁰, Groisman et al.¹⁹.

Example 3.1. -Product state- For entanglement maps $\phi(b) = \rho \text{Tr}_2 \sigma b$ and $\phi^*(a) = \sigma \text{Tr}_1 \rho a$, we have $\theta_\phi = \rho \otimes \sigma$. Then

$$I_\phi(\rho, \sigma) = 0, \quad S_\phi(\sigma|\rho) = S(\sigma), \quad S_\phi(\rho|\sigma) = S(\rho). \quad (23)$$

The above results clearly means that the product state has no correlation between its marginal states. Here we introduce the following notions³²:

Definition 3.2. (1) A state θ_ϕ is called a product state if it can be represented in the form

$$\theta_\phi = \rho \otimes \sigma \quad (24)$$

where $\rho = \phi(1)$ and $\sigma = \phi^*(1)$.

(2) A state is called a correlated state if it is not a product state.

The quantum mutual entropy $I_\phi(\rho, \sigma)$ measures the correlation as the difference between the density operator itself and the product state given by its marginal densities in the sense of relative entropy. In quantum system we have two types of correlated states. The first one is a separable correlated state, and second one is an entangled correlated state.

Example 3.2. -Separable correlated state- For an entanglement map $\phi(b) = \sum p_i \rho_i \text{Tr} \sigma_i b$, $\phi^*(a) = \sum p_i \sigma_i \text{Tr} \rho_i a$, where $p_i \geq 0$, $\sum p_i = 1$, we have

$$\theta_\phi = \sum p_i \rho_i \otimes \sigma_i. \quad (25)$$

with $\rho = \phi(1) = \sum p_i \rho_i$, $\phi^*(1) = \sum p_i \sigma_i$. Then we have the following inequalities^{3,11,12}

$$0 \leq I_\phi(\rho, \sigma) \leq \min\{S(\rho), S(\sigma)\}, \quad (26)$$

$$S_\phi(\sigma|\rho) > 0, S_\phi(\rho|\sigma) > 0. \quad (27)$$

Example 3.3. -Pure entangled correlated state- For entanglement maps $\phi(b) = \sum \lambda_i \bar{\lambda}_j |e_i\rangle \langle e_j| \langle f_j, b f_i\rangle$, $\phi^*(a) = \sum \lambda_i \bar{\lambda}_j |f_i\rangle \langle f_j| \langle e_j, a e_i\rangle$, where $\lambda_i \in \mathbb{C}$, $\sum |\lambda_i|^2 = 1$, we have

$$\theta_\phi = \sum \lambda_i \bar{\lambda}_j |e_i\rangle \langle e_j| \otimes |f_i\rangle \langle f_j| = |\Psi\rangle \langle \Psi|,$$

where $|\Psi\rangle = \sum \lambda_i e_i \otimes f_i$. and its marginal states are given by $\rho = \phi(1) = \sum |\lambda_i|^2 |e_i\rangle \langle e_i|$, $\sigma = \phi^*(1) = \sum |\lambda_i|^2 |f_i\rangle \langle f_i|$. Then

$$\begin{aligned} I_\phi(\rho, \sigma) &= S(\rho) + S(\sigma) - S(\theta_\phi) \\ &= 2S(\rho) > \min\{S(\rho), S(\sigma)\}, \end{aligned} \quad (28)$$

$$S_\phi(\sigma|\rho) = S_\phi(\rho|\sigma) = -S(\rho) < 0, \quad (29)$$

where $S(\rho) = S(\sigma) = -\sum |\lambda_i|^2 \log |\lambda_i|^2$.

The classical mutual entropy is always smaller than its marginal entropies, and the conditional entropy is always positive. From these points of view, on the base of above observation, we know that the separable state has the same property of correlation with classical probability, however the entangled state does not so. Above entropies of entangled states (i.e., the correlation of entangled states) show the non-classical property.

We introduce another criteria to measure the correlation of compound states^{11,12,18,28}.

Definition 3.3. For an entanglement map ϕ with $\rho = \phi(1)$, $\sigma = \phi^*(1)$, we defines

$$D_\phi^E(\rho, \sigma) \equiv \frac{1}{2} (S_\phi(\sigma|\rho) + S_\phi(\rho|\sigma)) \quad (30)$$

$$= S(\theta_\phi) - \frac{1}{2} (S(\rho) + S(\sigma)), \quad (31)$$

$$D_\phi(\rho, \sigma) \equiv -D_\phi^E(\rho, \sigma) \quad (32)$$

We call $D_\phi^E(\rho, \sigma)$ the *degree of entanglement*^{11,12,28} and call $D_\phi(\rho, \sigma)$ the *degree of quantum correlation*¹⁸.

Proposition 3.1. ^{2,28}If θ_ϕ is a pure state, then the following statements hold:

- (1) θ_ϕ is a pure entangled state iff $D_\phi(\rho, \sigma) > 0$.
- (2) θ_ϕ is a pure separable state iff $D_\phi(\rho, \sigma) = 0$.

It is well known³⁸ that if θ is a PPT state, then

$$S(\theta) - S(\rho) \geq 0, S(\theta) - S(\sigma) \geq 0. \quad (33)$$

where ρ and σ are the marginal states of θ . Due to this fact one shows

Proposition 3.2. If θ_ϕ is a mixed PPT state, then

$$D_\phi(\rho, \sigma) \leq 0. \quad (34)$$

Suppose now that we have two entanglement maps ϕ_1 and ϕ_2 such that $\rho = \phi_k(1)$ and $\sigma = \phi_k^*(1)$ (for $k = 1, 2$).

Definition 3.4. One says that ϕ_1 has stronger correlation than ϕ_2 if

$$D_{\phi_1}(\rho, \sigma) > D_{\phi_2}(\rho, \sigma). \quad (35)$$

Remark 3.2. Interestingly, there exist quantum entangled states with weaker correlations than some separable states in the sense of the order (35). In this remark we show one of such examples¹⁶.

We analyze Bell diagonal states (see Refs. [6-8, 17]). Consider $\mathcal{H} = \mathbb{C}^3 \otimes \mathbb{C}^3$ with the standard basis $\{e_0, e_1, e_2\}$. We set

$$\Omega_{0,0} = \frac{1}{\sqrt{3}} \sum_{i=0}^2 |e_i\rangle \langle e_i| \quad (36)$$

and define $\Omega_{k,l}$ for any k, l ($0 \leq k, l \leq 2$) as

$$\Omega_{k,l} = (W_{k,l} \otimes 1) \Omega_{0,0}, \quad (37)$$

where $W_{k,l}$ means the circle action given by

$$e_i \mapsto \exp\left(\frac{2\pi\sqrt{-1}k(i-l)}{3}\right) e_{i-l}, \quad (i = 0, 1, 2) \quad (38)$$

Finally, we define

$$\theta(\alpha, \beta) = \frac{1-\alpha-\beta}{9} 1 \otimes 1 + \alpha |\Omega_{0,0}\rangle \langle \Omega_{0,0}| + \frac{\beta}{2} (|\Omega_{1,0}\rangle \langle \Omega_{1,0}| + |\Omega_{2,0}\rangle \langle \Omega_{2,0}|). \quad (39)$$

Due to the computation for the eigenvalues of $\theta(\alpha, \beta)$ one knows that $\theta(\alpha, \beta)$ defines a state iff the parameters α, β satisfy

$$\alpha + \beta \leq 1, \quad 2\alpha - 7\beta \leq 2, \quad -8\alpha + \beta \leq 1. \quad (40)$$

Let us consider for example the line $\beta = \frac{3}{5}$. It was shown⁷ that $\theta(\alpha, \frac{3}{5})$ is entangled iff

$$\alpha \in [0, \frac{1}{4}) \cup (\frac{13}{40}, \frac{2}{5}], \quad (41)$$

and it is separable iff

$$\alpha \in [\frac{1}{4}, \frac{13}{40}]. \quad (42)$$

Now we can compute its degree of quantum correlation DQC denoted by $D(\theta(\alpha, \frac{3}{5}))$ as

$$\begin{aligned} D\left(\theta\left(\alpha, \frac{3}{5}\right)\right) &= \log 3 + \frac{31 - 10\alpha}{45} \log\left(\frac{31 - 10\alpha}{90}\right) \\ &\quad - \frac{10\alpha - 4}{15} \log\left(\frac{2 - 5\alpha}{45}\right) - \frac{40\alpha + 2}{45} \log\frac{40\alpha + 2}{45}. \end{aligned} \quad (43)$$

By a simple calculation, it is easily be verified that $D(\theta(\alpha, \frac{3}{5}))$ is monotonically increasing if $\alpha \in [\frac{1}{20}, \frac{2}{5})$. Hence the values of DQC for $\alpha \in [\frac{1}{20}, \frac{1}{4})$ (i.e. $\theta(\alpha, \frac{3}{5})$ is entangled.) is smaller than the ones for $\alpha \in [\frac{1}{4}, \frac{13}{40}]$ (i.e. $\theta(\alpha, \frac{3}{5})$ is separable.).

4. Summary and Discussion

In the section 3 we drew a quantum analogue of the classical probability relation (18) as the relation (17) between a compound density θ_ϕ and a quantum conditional probability operator π_ϕ via an entanglement map ϕ . However we have not yet discussed the roll of the relation (17) in the context of quantum information theory. Before the discussion we summarize the results of the section 3. The quantum mutual entropy $I_\phi(\rho, \sigma)$ and conditional entropies $S_\phi(\sigma|\rho)$ and $S_\phi(\rho|\sigma)$ have a non-classical property (see the example 3.3). How can we understand such non-classical property? Due to the non-classical property we can classify a subset of \mathcal{S}_{ent} (the set of entangled states) by using the DEN or DQC (see the proposition 3.1. and 3.2.). This is, in some sense, one of the answers above question. The non-classical property of such entropies are found on the subset of entangled states, however such property does not give the necessary condition of entangled states. Anyway it is natural to think that the mutual entropy $I_\phi(\rho, \sigma)$ of entangled

state represents the correlation of entanglement. From this points of view during the last one decade there have been a lot of studies of quantum correlation by means of the mutual entropy^{2,3,11,12,14,19,20,21,29,33,38}. On the base of above observation we introduced the order of quantum correlation in the sense of (35). Interestingly there exist quantum entangled states with weaker correlations than some separable states (see Remark 3.2.). How can we understand such correlation between separable and entangled states?

In the classical information theory the mutual entropy can be recognized as the criterion to measure the transmitted information through a channel. We can also represent the quantum mutual entropy $I_\phi(\rho, \sigma)$ by channel representation using the relation (17). If a QCPO π is given, then we can define a channel Λ^* for any input state ρ_{in} by

$$\rho_{\text{out}} = \Lambda^*(\rho_{\text{in}}) \equiv \text{Tr}_1 \left(\rho_{\text{in}}^{\frac{1}{2}} \otimes 1 \right) \pi \left(\rho_{\text{in}}^{\frac{1}{2}} \otimes 1 \right). \quad (44)$$

In this scheme we can call π a channel density. Applying this scheme to the density operator θ_ϕ , then we have

$$\begin{aligned} I_\phi(\rho, \sigma) &= \text{Tr} \theta_\phi (\log \theta_\phi - \log \rho \otimes \sigma) \\ &= \text{Tr} \left(\rho^{\frac{1}{2}} \otimes 1 \right) \pi_\phi \left(\rho^{\frac{1}{2}} \otimes 1 \right) \\ &\quad \left(\log \left(\rho^{\frac{1}{2}} \otimes 1 \right) \pi_\phi \left(\rho^{\frac{1}{2}} \otimes 1 \right) - \log \rho \otimes \Lambda_\phi^*(\rho) \right) \\ &\equiv I_\phi(\rho; \Lambda_\phi^*), \end{aligned} \quad (45)$$

where $\Lambda_\phi^*(\rho) \equiv \text{Tr}_1 \left(\rho^{\frac{1}{2}} \otimes 1 \right) \pi_\phi \left(\rho^{\frac{1}{2}} \otimes 1 \right)$. However this mutual entropy $I_\phi(\rho; \Lambda_\phi^*)$ does not satisfy the following Shannon's fundamental inequality on the subset of entangled states (see the example 3.3.):

$$0 \leq I_\phi(\rho; \Lambda_\phi^*) \leq \min \{ S(\rho), S(\Lambda_\phi^*(\rho)) \}. \quad (46)$$

From this point of view the quantum mutual entropy $I_\phi(\rho; \Lambda_\phi^*)$ is not a proper measure of the transmitted information through the channel.

If one restricts the above argument to the set of separable states, then one knows that $I_\phi(\rho, \sigma)$ satisfies the quantum Shannon's inequality (see the example 3.2.). In addition we can define the quantum conditional entropy as an extension of the classical definition of conditional entropy:

$$\begin{aligned} S_{\phi_{\text{sep}}}(\sigma|\rho) &\equiv \text{Tr} \left(\rho^{\frac{1}{2}} \otimes 1 \right) \pi_{\phi_{\text{sep}}} \left(\rho^{\frac{1}{2}} \otimes 1 \right) \log \pi_{\phi_{\text{sep}}} \\ &= \text{Tr} \theta_{\phi_{\text{sep}}} \log \pi_{\phi_{\text{sep}}} \end{aligned} \quad (47)$$

Due to the separability of $\theta_{\phi_{\text{sep}}}$ we can draw the same decomposition of $S_{\phi_{\text{sep}}}(\sigma|\rho)$ given as the definition (21) from the definition (47):

$$S_{\phi_{\text{sep}}}(\sigma|\rho) = S(\theta_{\phi}) - S(\rho) = S(\sigma) - I_{\phi}(\rho, \sigma). \quad (48)$$

One may observe that on the set of separable state the scheme of classical information theory may be sifted to that of quantum information theory via the relation (17), however it can not be done on the set of entangled states.

Now we have another question from the information theoretical point of view: how can we connect $I_{\phi}(\rho, \sigma)$ to the transmitted information through the entanglement channel Λ_{ϕ}^* on the base of the relation (17)? We will continue to study the above problems in the forthcoming paper.

References

1. L. Accardi, D. Chruściński, A. Kossakowski, T. Matsuoka and M. Ohya, "On classical and quantum liftings", *Open. Syst. Info. Dyn.* **17**, 361-386 (2010).
2. L. Accardi, T. Matsuoka, M. Ohya, "Entangled Markov chains are indeed entangled", *Infin. Dim. Anal. Quantum Probab. Top.* **9**, 379-390 (2006).
3. L. Accardi, T. Matsuoka, M. Ohya, "Entangled Markov chain satisfying entanglement condition", *RIMS* **1658**, 84-94 (2009).
4. L. Accardi, M. Ohya, "Compound channels, transition expectation and liftings", *J. Appl. Math. Optim.*, **39**, 33-59 (1999).
5. M. Asorey, A. Kossakowski, G. Marmo, E.C.G. Sudarshan, "Relation between quantum maps and quantum state", *Open. Syst. Info. Dyn.* **12**, 319-329 (2006).
6. B. Baumgartner, B. Hiesmayer and H. Narnhofer, "State space for two qutrits has a phase space structure in its cone", *Phys. Rev. A* **74**, 032327 (2006).
7. B. Baumgartner, B. Hiesmayer and H. Narnhofer, "A special simplex in the state space for entangled qutrits", *J. Phys. A: Math. Theor.*, **40**, 7919 (2007).
8. B. Baumgartner, B. Hiesmayer and H. Narnhofer, "The geometry of biparticle qutrits including bound entanglement", *Phys. Lett. A*, **372**, 2190 (2008).
9. V. P. Belavkin, *Radio Eng. Electron. Phys.*, **25**, 1445-1453 (1980).
10. V. P. Belavkin, *Found. Phys.*, **24**, 685-714 (1994).
11. V. P. Belavkin, M. Ohya, "Quantum entropy and information in discrete entangled state", *Infin. Dim. Anal. Quantum Probab. Top.* **4**, 33-59 (2001).
12. V. P. Belavkin, M. Ohya, "Entanglement, quantum entropy and mutual information", *Proc. R. Soc. London A* **458**, 209-231 (2002).
13. V. P. Belavkin, X. Dai, "An operational algebraic approach to quantum channel capacity", *Int. J. Quantum Inf.* **6**, 981 (2008).
14. N. J. Cerf and C. Adami, "Negative entropy and information in quantum mechanics", *Phys. Rev. Lett.* **79**, 5194-5197 (1997).
15. M. D. Choi, "Completely positive maps on complex matrix", *Lin. Alg. Appl.*, **10**, 285 (1975).

16. D. Chruściński, Y. Hirota, T. Matsuoka and M. Ohya, "Remarks on the degree of entanglement" *QP- PQ Quantum Probab. & White Noise Anal.* **28**, 145-156 (2011).
17. D. Chruściński, A. Kossakowski, T. Matsuoka, K. Młodawski, "A class of Bell diagonal states and entanglement witnesses", *Open. Syst. Info. Dyn.* **17**, 213-231 (2010).
18. D. Chruściński, A. Kossakowski, T. Matsuoka and M. Ohya, "Entanglement mapping vs. quantum conditional probability operator", *QP- PQ Quantum Probab. & White Noise Anal.* **28**, 223-236 (2011).
19. B. Groisman, S. Popescu and A. Winter, "Quantum, classical, and total amount of correlations in a quantum state" *Phys. Rev A* **72**, 0323187 (2005).
20. L. Henderson and V. Vedral, "Classical, quantum and total correlation", *J. Phys. A* **34** 6913 (2001)
21. M. Horodecki and R. Horodecki, "Information-theoretical aspect of quantum inseparability of mixed states", *Phys. Rev. A* **54**, 1838-1843 (1996).
22. M. Horodecki, P. Horodecki and R. Horodecki, *Phy. Lett.*, **A 223**, 1 (1996).
23. A. Jamiolkowski, "Linear transformation which preserve trace and positive semidefiniteness of operators", *Rep. Math. Phys.* **3**, 275 (1972).
24. A. Jamiolkowski, T. Matsuoka and M. Ohya, "Entangling operator and PPT condition", TUS preprint (2007).
25. G. Kimura, A. Kossakowski, "A note on positive maps and classification on states", *Open Sys. & Information Dyn.* **12**, 1 (2005).
26. W. A. Majewski, T. Matsuoka and M. Ohya, "Characterization of partial positive transposition states and measures of entanglement", *J. Math. Phys.* **50**, 113509 (2009).
27. T. Matsuoka, "On generalized entanglement", *QP- PQ Quantum Probab. & White Noise Anal.* **21**, 170-180 (2007).
28. T. Matsuoka, M. Ohya, "Quantum entangled state and its characterization", *Foud. Probab. Phys.* **3 750**, 298-306 (2005).
29. K. Modi, T. Paterek, W. Son, V. Vedral and M. Williamson, "Unified view of quantum and classical correlations", *Phys. Rev. Lett.*, **104**, 080501 (2010).
30. M. Ohya, "On coomound state and mutual information in quantum information theory", *IEEE Info. Theory*, **29**, 77-774 (1983).
31. M. Ohya, "Note on quantum probability", *Nuovo Cimento*, **38**, 402-406 (1983).
32. M. Ohya, I. V. Volovich, *Mathematical Foundation of Quantum Information and Computation*, Springer, New Youk, 2011.
33. H. Ollivier, W. Z. Zurek, "Quantum discord: A measure of the quantumness of correlations", *Phys. Rev. Lett.*, **88**, 017901 (2002).
34. A. Peres, *Phys.Rev.Lett.*, **77**, 1413 (1996).
35. C. E. Shannon, "A mathematical theory of communication", Urbana, IL: Univ. Illinois (1949).
36. H. Umegaki, "Conditional expectation in an operator algebras IV", *Kodai Math. Sem. Rep.*, **14**, 59-85 (1962).
37. K. Urbanik, "Joint probability distribution of observables in quantum mechanics", *Stud. Math. T.* **21**, 317-323 (1961).

38. K.G.H. Vollbrecht, M.M. Wolf, "Conditional entropies and their relation to entanglement criteria", e-print arXiv: quant-ph/0202058v1.
39. S. L. Woronowicz, "Positive maps of low dimensional matrix algebra", Rep. Math. Phys., **10**, 165 (1976)

固有値・固有空間のアンホロノミーの階層的構成¹

首都大学東京 理工物理 田中篤司
 高知工科大学 理論物理研究室 全卓樹
 Pusan National University Sang Wook Kim

量子系の持つパラメータを充分ゆっくり変動させると、量子系の時間発展は断熱定理 [1] に従うことは良く知られています。Berry は、パラメータの軌跡を閉じた場合、断熱的な時間発展は自明では無く、位相におけるアンホロノミーが発現することを指摘しました [2]。これは Berry の位相、あるいは、微分幾何学の用語から位相のホロノミーとも呼ばれます [3]。さて、断熱的な時間発展での中心的な役者は、状態ベクトルの位相以外にも、固有値と固有空間が居ます。固有値や固有空間にはアンホロノミー的な現象は起きないのでしょうか？ この疑問は Berry の仕事以来見過されてきたようですが、20 世紀の末に、一般化された点状相互作用下の質点で具体例が見いだされました [4]。このモデルでは、パラメータ空間中の適当な閉経路に沿って固有エネルギーを追跡すると始点と終点は閉じません。つまり、断熱的な時間発展であっても、この閉経路に沿うことで系は励起します (あるいはその逆)。これは、固有値・固有空間のアンホロノミー [4, 5] (または、新奇な量子ホロノミー [6]) が発現したと呼ぶべき状況です。

近年、固有値・固有空間のアンホロノミーを持つモデルが見い出され [5, 7, 6, 8, 9] (図 1 も参照)、理論的な理解も進展してきました。例えば、藤川による Berry 位相の取り扱い [11, 12] の固有空間のアンホロノミーへの拡張 [6, 13]、複素例外点との関連づけ [14]、断熱 Floquet 理論を通じた Abelian gerbes を用いた定式化 [15] を挙げるすることができます。

一方、既存の例は全て“一体”のものに限られています。つまり、これらの例を複数の構成要素に分解する手続は不明でした。これに対し、本稿では、 N -体系での例を説明します。この新しい例は N -qubit 上の量子回路上的のもので、小さい系の例を構成要素としてより大きな例を順々に構成していきます。特に、この構成法で得られる最も簡単な例を説明します。

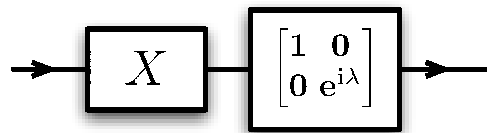


図 1: 固有値・固有空間のアンホロノミーを持つ最も簡単な量子回路。量子 NOT ゲートと位相シフトゲート (強度 λ) から成り、パラメータ λ の住む空間は周長 2π の閉経路。本文では、これとユニタリー等価な回路 (式 (1)) と、その多 qubit への拡張を調べます。

¹ 基研研究会「量子科学における双対性とスケール」(2010 年 11 月 4 日-6 日), 2010 年 11 月 4 日の口頭発表より。

1 階層的構成の種 — 1 qubit の量子回路にて

固有値・固有空間のアンホロノミーを持つ系は珍しいように思えますが、1 qubit 上の、二つの量子ゲートから成る簡素な量子回路がその一例を与えます (図 1)。ここでは、固有ベクトルの表記を簡素にするため、

$$\hat{u}(\lambda) \equiv \exp(i\lambda|y\rangle\langle y|) \hat{Z} \quad (1)$$

を考えます。ここで、 $\hat{X} \equiv |0\rangle\langle 1| + |1\rangle\langle 0|$, $\hat{Y} \equiv i(|1\rangle\langle 0| - |0\rangle\langle 1|)$, $\hat{Z} \equiv |0\rangle\langle 0| - |1\rangle\langle 1|$ 。 $|y\rangle \equiv (|0\rangle - i|1\rangle)/\sqrt{2}$ は $\hat{Y}|y\rangle = -|y\rangle$ を満たします。式 (1) 右辺第一因子は、“y 軸に関する位相ゲート” を記述します： $|y\rangle$ と平行な成分は λ だけ位相をずらし、 $|y\rangle$ と直交する成分はそのままにするものです。 $\hat{u}(\lambda)$ の固有値問題を解きます。これはユニタリーなので、固有角

$$\theta(n; \lambda) = n\pi + \frac{1}{2}\lambda \quad (2)$$

($n = 0, 1$) を用いると、固有値は $z(n; \lambda) = e^{i\theta(n; \lambda)}$ 。対応する固有ベクトルは

$$|0(\lambda)\rangle = |0\rangle \cos \frac{\lambda}{4} + |1\rangle \sin \frac{\lambda}{4}, \quad |1(\lambda)\rangle = |1\rangle \cos \frac{\lambda}{4} - |0\rangle \sin \frac{\lambda}{4}. \quad (3)$$

$\hat{u}(\lambda)$ は λ について周期 2π を持ちます。 λ の一周分増大は閉経路とみなせ、これを C と呼びます。 N 体系での導入を兼ね、 $\hat{u}(\lambda)$ の固有角のアンホロノミーを調べます。出発点の量子数 n が規定する整数 $s(n)$, $r(n)$ を導入し、 $\theta(n; \lambda)$ のパラメータ依存性を

$$\theta(n; \lambda + 2\pi) = \theta(s(n); \lambda) + 2\pi r(n) \quad (4)$$

とまとめます。つまり、 C を一周することで、量子数は n から $s(n)$ に化けます。これに対応して、経路 C が規定する行列 $S(C)$ を、その行列要素

$$\{S(C)\}_{n', n} = \delta_{n', s(n)} \quad (5)$$

で定義します。特に、 $\hat{u}(\lambda)$ では、 $s(n) = \bar{n}$, $r(n) = n$ となり、固有角のアンホロノミーが陽に現れます (ここで、 $\bar{0} = 1$, $\bar{1} = 0$ と約束します)。これに対応し、 $S(C)$ は非自明な置換行列です。

ユニタリー作用素での固有値と固有空間の対応から、固有空間のアンホロノミーも発現します。以下、これをゲージ接続を通じて調べます [6]。基底ベクトル列 (“枠”)

$$f(\lambda) \equiv [|0(\lambda)\rangle, |1(\lambda)\rangle] \quad (6)$$

の微分は非アーベル的なゲージ接続

$$A(\lambda) \equiv \{f(\lambda)\}^\dagger i \partial_\lambda f(\lambda) \quad (7)$$

を誘導します。これは 2×2 のエルミート行列です。逆に、接続を積分すると $f(\lambda)$ を得ることができます:

$$f(\lambda'') = f(\lambda') \exp \left(-i \int_{\lambda'}^{\lambda''} A(\lambda) d\lambda \right). \quad (8)$$

ここで、 \exp_{\rightarrow} は反順序的指数関数。固有値の縮退が無い場合、 $A(\lambda)$ はゲージ接続 $A^D(\lambda)$ を誘導します：

$$\{A^D(\lambda)\}_{n',n} = \delta_{n',n} \{A(\lambda)\}_{n,n}. \quad (9)$$

これは、全ての固有空間について、Mead-Truhlar-Berry のゲージ接続 [16, 2] をひとまとめにしたものです。

固有空間のアンホロノミーを特徴付けるホロノミー行列 $M(C)$ は、 $\langle n(\lambda) |$ と $|n(\lambda + 2\pi)\rangle$ の重なり積分で定義されます。ただし、これは、 $|n(\lambda)\rangle$ のゲージを Berry 位相の扱いにおける平行移動条件 [17] を満たすように選んだ場合の話です。 $M(C)$ のゲージ共変な表式として、上で導入したゲージ接続を使ったものを示します [6]：

$$M(C) = \exp_{\rightarrow} \left(-i \int_C A(\lambda) d\lambda \right) \exp_{\leftarrow} \left(i \int_C A^D(\lambda) d\lambda \right). \quad (10)$$

$\hat{u}(\lambda)$ では、

$$A(\lambda) = \frac{1}{4} Y, \quad \text{ここで} \quad \begin{bmatrix} Y_{00} & Y_{01} \\ Y_{10} & Y_{11} \end{bmatrix} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (11)$$

なので、 $A^D(\lambda) = 0$ 。 $M(C)$ を計算していくと、

$$\{M(C)\}_{n',n} = \{S(C)\}_{n',n} (-1)^n. \quad (12)$$

となります。つまり、 $M(C)$ は、その骨組み $S(C)$ と位相因子から構成されます。

2 N 体系を階層的に構成する

$N - 1$ 体系がアンホロノミーを持つとします。ここに一つ qubit を付けて、アンホロノミーを持つ N 体系を作ります。これを実現するための材料として、 \hat{U} を引数とする回路を導入します：

$$\hat{D}[\hat{U}] \equiv \hat{C}^y[\hat{U}](\hat{Z} \otimes \hat{1}), \quad (13)$$

ここで

$$\hat{C}^y[\hat{U}] \equiv (\hat{1} - |y\rangle\langle y|) \otimes \hat{1} + |y\rangle\langle y| \otimes \hat{U} \quad (14)$$

は、制御 bit が “y 軸方向” の制御ユニタリゲートです。試しに、 \hat{U} として補助 qubit の大域的な位相ゲート $e^{i\lambda\hat{1}}$ を使ってみると、 $\hat{D}[e^{i\lambda\hat{1}}] = \hat{u}(\lambda) \otimes \hat{1}$ となります。つまり、 $\hat{D}[\cdot]$ には、アンホロノミーを起こす量子回路 $\hat{u}(\lambda)$ が潜んでいます。

さて、 N -qubit 上の量子回路 $\hat{U}^{(N)}(\lambda)$ を導入します。まず、 $\hat{U}^{(1)}(\lambda) \equiv \hat{u}(\lambda)$ とします。そして、 $N > 1$ では漸化式

$$\hat{U}^{(N)}(\lambda) \equiv \hat{D}[\hat{U}^{(N-1)}(\lambda)] \quad (15)$$

を課します。これを構成するのに必要な量子ゲートの数の増え方は、 N を増大するにつれ、たかだか N の多項式程度であるという意味で、 $\hat{U}^{(N)}(\lambda)$ は効率的に構成可能です。

3 固有値問題を解く

$\hat{U}^{(N)}(\lambda)$ は漸化式 (15) を通じて定義されたので、これに関する諸量の解析において、自然に漸化式が現われます。ここでは、最も簡素な例として、固有値問題を再帰的に解いていきます。簡潔さのため結果を先取りします： $\hat{U}^{(N)}(\lambda)$ の固有状態を指定する量子数として、 $(n_N, n_{N-1}, \dots, n_1)$ を用います。ここで、 n_j は $\{0, 1\}$ のいずれか。そこで、固有角を $\theta^{(N)}(n_N, n_{N-1}, \dots, n_1; \lambda)$ 、固有ベクトルを $|n_N, n_{N-1}, \dots, n_1; \lambda\rangle$ と記します。 $N = 1$ では $\theta^{(1)}(n_1; \lambda) = \theta(n_1; \lambda)$ 、 $|n_1; \lambda\rangle = |n_1(\lambda)\rangle$ 。 $N > 1$ にて固有角と固有ベクトルの漸化式を求めると、

$$\theta^{(N)}(n_N, n_{N-1}, \dots, n_1; \lambda) = \theta(n_N; \theta^{(N-1)}(n_{N-1}, \dots, n_1; \lambda)) \quad (16)$$

および

$$|n_N, n_{N-1}, \dots, n_1; \lambda\rangle = |n_N(\theta^{(N-1)}(n_{N-1}, \dots, n_1; \lambda))\rangle \otimes |n_{N-1}, \dots, n_1; \lambda\rangle. \quad (17)$$

これを解くと、

$$\theta^{(N)}(n_N, n_{N-1}, \dots, n_1; \lambda) = \frac{2\pi}{2^N} \left\{ m_N(n_N, n_{N-1}, \dots, n_1) + \frac{\lambda}{2\pi} \right\}, \quad (18)$$

ここで、

$$m_N(n_{N-1}, \dots, n_1) = \sum_{j=1}^N 2^{j-1} n_j \quad (19)$$

は主量子数と呼ぶべきものです。固有ベクトルは、

$$|n_N, n_{N-1}, \dots, n_1; \lambda\rangle = \bigotimes_{j=1}^N |n_j(\theta^{(j-1)}(n_{j-1} \dots n_1; \lambda))\rangle \quad (20)$$

です。

4 固有角・固有空間のアンホロノミーの解析

λ を経路 C に沿って一周する (つまり 2π 増やす) 場合のアンホロノミーを調べます。最初に $\hat{U}^{(N)}(\lambda)$ の固有状態の量子数 n_N, \dots, n_1 を指定します。 λ を一周させた後では、アンホロノミーに対応して系の量子数が変わります。終点での j 番目の量子数を $s_j^{(N)}(n_N, \dots, n_1; C)$ ($1 \leq j \leq N$) と記します。これらをひとまとめたものは $2^N \times 2^N$ の置換行列

$$\{S^{(N)}(C)\}_{(n'_N, \dots, n'_1), (n_N, \dots, n_1)} \equiv \prod_{j=1}^N \delta_{n'_j, s_j(n_N, \dots, n_1)} \quad (21)$$

です。これは、“量子数のアンホロノミー” を表現したものだとも言えます。 $S^{(N)}(C)$ に位相の情報を加味したものがホロノミー行列 $M^{(N)}(C)$ です。その行列要素は

$$\{M^{(N)}(C)\}_{m'_N, m_N} = \langle m'_N; \lambda | \psi^{(C)}(m_N; \lambda) \rangle \quad (22)$$

です。ただし、 $|\psi^{(C)}(m_N; \lambda)\rangle$ は $|m_N; \lambda\rangle$ を C に沿って平行移動させたものです。また、 n_N, \dots, n_1 を m_N と略記しました。ユニタリー作用素での固有値と固有空間の対応から $M^{(N)}(C)$ は $S^{(N)}(C)$ および固有空間毎の位相因子 $\sigma^{(N)}(m_N)$ から構成されます:

$$\{M^{(N)}(C)\}_{m'_N, m_N} = \{S^{(N)}(C)\}_{m'_N, m_N} \sigma^{(N)}(m_N; C). \quad (23)$$

以下、 $S^{(N)}(C)$ と $M^{(N)}(C)$ を求めていきます。

固有角のアンホロノミーを調べる $s_j^{(N)}(m_N; C)$ を求めるため、固有角のパラメータ依存性を観察します。 $s_j^{(N)}(m_N; C)$ と整数 $r^{(N)}(m_N; C)$ を用いると、 C の終点での固有角の値は

$$\theta^{(N)}(m_N; \lambda + 2\pi) = \theta^{(N)}(s^{(N)}(m_N; C); \lambda) + 2\pi r^{(N)}(m_N; C), \quad (24)$$

と整理できます。ここで、量子数の並び $s_N^{(N)}(m_N; C), \dots, s_1^{(N)}(m_N; C)$ を $s^{(N)}(m_N; C)$ と略記しました。

我々の模型 $\hat{U}^{(N)}(\lambda)$ では、 $N = 1$ のとき $s_1^{(1)}(n_1; C) = \bar{n}_1$, $r^{(1)}(n_1; C) = n_1$ 。 $N > 1$ では、

$$s_N^{(N)}(n_N, \dots, n_1) = \begin{cases} \bar{n}_N & n_{N-1} \cdots n_1 = 1 \text{ の場合} \\ n_N & \text{他} \end{cases} \quad (25)$$

および、 $r_N^{(N)}(n_N, \dots, n_1) = n_N \cdots n_1$ 。よって、置換行列 $S^{(N)}(C)$ は長さ 2^N の巡回置換を記述します。

ホロノミー行列 $M^{(N)}(C)$ を求める $\sigma^{(N)}(m_N)$ を求めるには、 $M^{(N)}(C)$ のゲージ共変な表式 (10) を利用します。ここで用いる非 Abel なゲージ接続は、 $2^N \times 2^N$ のエルミート行列

$$\{A^{(N)}(\lambda)\}_{m'_N, m_N} \equiv \langle m_N; \lambda | (i\partial_\lambda | m_N; \lambda) \rangle. \quad (26)$$

対応する Mead-Truhlar-Berry のゲージ接続達は $\{A^{D(N)}(\lambda)\}_{m'_N, m_N} \equiv \delta_{m'_N, m_N} \{A^{(N)}(\lambda)\}_{m_N, m_N}$ 。 $M^{(N)}(C)$ の行列要素に関する漸化式を求めると、

$$\{M^{(N)}(C)\}_{m'_N, m_N} = \left\{ e^{-i\pi r^{(N-1)}(m_{N-1}; C)Y/2} \right\}_{n'_N, n_N} \{M^{(N-1)}(C)\}_{m'_{N-1}, m_{N-1}}. \quad (27)$$

この結果、 $\sigma^{(N)}$ の漸化式を得ます:

$$\sigma^{(N)}(m_N; C) = (-1)^{r^{(N)}(m_N; C)} \sigma^{(N-1)}(m_{N-1}; C). \quad (28)$$

Manini-Pistolesi のゲージ不変量 ホロノミー行列 $M^{(N)}(C)$ はゲージ共変であるものの、ゲージ不変量ではありません。しかし、ここから、Manini-Pistolesi のゲージ不変量 [18] を引き出すことができます。詳細は割愛しますが、ここでの非自明なゲージ不変量は、

$$\gamma_{\text{MP}}^{(N)}(C) \equiv \prod_{n_N=0}^1 \cdots \prod_{n_1=0}^1 \sigma^{(N)}(n_N, \dots, n_1; C) \quad (29)$$

のみです。実は、 $\gamma_{\text{MP}}^{(N)}(C)$ は、閉経路 C を 2^N 回繰り返して元に戻ったときの Berry 位相のことです。結局、我々の模型では、任意の $N > 0$ で

$$\gamma_{\text{MP}}^{(N)}(C) = -1 \quad (30)$$

となり、Longuet-Higgins 位相 [19] に類似した状況が示唆されます。

例: $N = 3$ での量子数の軌跡 経路 C を繰り返すことによる固有角の軌跡を理解するには、以上から定まる量子数の軌跡を理解すれば充分です。特に簡単なのは、主量子数 m_N の値を追跡することで、 C を一周する毎に m_N は 1 増えます。ただし、 m_N が 2^N に達すると $m_N = 0$ に戻ります。 n_N, \dots, n_1 は m_N の二進展開の係数であることに注意すると、その軌跡が容易に求まります。例えば、 $N = 3$ では

$$000 \mapsto 001 \mapsto 010 \mapsto 011 \mapsto 100 \mapsto 101 \mapsto 110 \mapsto 111 \mapsto 000. \quad (31)$$

これに付随する位相因子として興味深いのはゲージ不変な $\gamma_{\text{MP}}^{(N)}(C) = -1$ のみですが、これは一般論を用いる方がより見通し良く求まるものです。

5 まとめと展望

多体系 (多 qubit 系) において、固有値・固有空間のアンホロノミーを持つ一族を見だし、拡張された藤川の定式化 [6] を用いて固有空間のアンホロノミーを調べました。この例は階数 1 の摂動下の量子写像とみなせる [20] ので、既存の例の構成要素からの構成法を与えたともみなせます。一方、この多体系の一族の構成法の単純な拡張を通じて、既存の例には還元できない模型も構成できます。これについては、別に報告する予定です [20]。

謝辞

本研究の一部について、科研費 (22540396) および JST の助成を受けました。

参考文献

- [1] M. Born and V. Fock, Z. Phys. A **51**, 165 (1928).
- [2] M. V. Berry, Proc. R. Soc. London **A 392**, 45 (1984).
- [3] B. Simon, Phys. Rev. Lett. **51**, 2167 (1983).
- [4] T. Cheon, Phys. Lett. A **248**, 285 (1998).
- [5] A. Tanaka and M. Miyamoto, Phys. Rev. Lett. **98**, 160407 (2007).

- [6] T. Cheon and A. Tanaka, Europhys. Lett. **85**, 20001 (2009).
- [7] M. Miyamoto and A. Tanaka, Phys. Rev. A **76**, 042115 (2007).
- [8] T. Cheon, A. Tanaka, and S. W. Kim, Phys. Lett. A **374**, 144 (2009).
- [9] A. Tanaka and T. Cheon, Phys. Rev. A **82**, 022104 (2010).
- [10] F. Wilczek and A. Zee, Phys. Rev. Lett. **52**, 2111 (1984).
- [11] K. Fujikawa, Phys. Rev. D **72**, 025009 (2005).
- [12] K. Fujikawa, Ann. Phys. (N.Y.) **322**, 1500 (2007).
- [13] A. Tanaka and T. Cheon, Ann. Phys. (N.Y.) **324**, 1340 (2009).
- [14] S. W. Kim, T. Cheon, and A. Tanaka, Phys. Lett. A **374**, 1958 (2010).
- [15] D. Viennot, J. Phys. A. **42**, 395302 (2009).
- [16] C. Mead and D. G. Truhlar, J. Chem. Phys. **70**, 2284 (1979).
- [17] A. J. Stone, Proc. R. Soc. London **A 351**, 141 (1976).
- [18] N. Manini and F. Pistolesi, Phys. Rev. Lett. **85**, 3067 (2000).
- [19] H. C. Longuet-Higgins, Proc. R. Soc. London **A 344**, 147 (1975).
- [20] A. Tanaka and T. Cheon and S.W. Kim, *in preparation*.

On certain effective methods in investigation of positive maps

Andrzej Jamiolkowski
Institute of Physics, Nicolaus Copernicus University,
87–100 Toruń, Poland,
e-mail: jam@fizyka.umk.pl

May 3, 2011

Abstract

The main purpose of this paper is to discuss some effective methods of study of certain characterizations of quantum operations and quantum channels. In particular, an effective method of checking if a given quantum operation is irreducible or not is presented. Our methods are based on analyzing an explicit form of a fixed quantum operation in its Kraus representation. Some examples based on theory of open quantum systems are discussed.

1 Introduction

Non-classical correlations among subsystems of a composite quantum system, known as entanglement, can be precisely described mathematically but, at least in the case of mixed states, in a rather ineffective way. Here, by an effective way (effective procedure) we mean a method which uses only a finite number of arithmetic operations on elements of a given density matrix (representing a quantum state) and allows us to formulate an answer to the question: is a given composite quantum system in a separable or entangled state?

It appears, that the questions of the above type can be naturally connected with some problems of changes in the space of quantum states, that is with some aspects of quantum dynamics and properties of dynamical maps. Here, by dynamical maps we understand linear transformations that take one density operator to another. Moreover, this kind of connection between some linear transformations and states of composed quantum systems is a one-to-one type. At first sight this seems strange that two entirely different issues, namely, the inner structure of states and, on the other hand, their dynamics are so strongly connected. But in fact there exists an intricate and strong link between them — there exists an isomorphism between positive maps on quantum states (these maps very often are called superoperators on the space of quantum states) and some properties of inner structure

of density matrices representing composite quantum systems [1, 2]. At present, various methods of study of entangled systems based on properties of positive maps are discussed in hundreds of papers and many monographs and textbooks (cf. [3, 4, 5]).

In this paper we will discuss some effective methods of study of certain properties of quantum operations and we will use some methods and results which are typical for the research area known as *non-commutative Perron-Frobenius theory*. In particular, in our study of dynamical maps we will use the natural identifications

$$M_k(\mathbb{C}) \otimes B(\mathcal{H}) \cong M_k(B(\mathcal{H})) \cong B\left(\bigoplus_{i=1}^k \mathcal{H}\right), \quad (1.1)$$

where \mathcal{H} denotes a finite dimensional Hilbert space, $B(\mathcal{H})$ represents the set of all linear operators on \mathcal{H} and symbols $M_k(\mathcal{A})$ for $k, 2, 3, \dots$ denote $k \times k$ matrices with elements from an algebra \mathcal{A} . In our approach we will concentrate on a quantum analogy of the classical theory of positive maps also known as Perron-Frobenius theory. Perron theorem on positive matrices [6] and its generalization by Frobenius on nonnegative matrices [7] have interested mathematicians since the results appeared at the beginning of last century. Later these theorems have been generalized to operators on a partially ordered real Banach spaces [8]. This has motivated several authors to consider linear maps on a finite dimensional space which leave a fixed cone invariant (cf., e.g. [9, 10]).

Almost at the same time it appeared that some natural questions connected with fundamentals of quantum mechanics (more precisely, with theory of open quantum systems) lead to investigations of linear maps (*superoperators*) in a real Banach space of self-adjoint operators on a fixed Hilbert space ([11, 12]). This concept of Banach space with the partial order defined by a specific cone, namely, the cone of positive semidefinite operators, constitutes a basic idea in description of open quantum systems. It should be stressed that in this case we consider linear maps on $B(\mathcal{H})$ not only as maps on a vector space but also as maps on $B(\mathcal{H})$ equipped in a natural structure of an algebra. Such linear maps on the set of operators are often called superoperators and their general form is well known. Namely, for a given superoperator Φ ,

$$\Phi : B(\mathcal{H}) \rightarrow B(\mathcal{H}), \quad (1.2)$$

there always exists an operator-sum representation of maps given by

$$\Phi(X) = \sum_{i=1}^{\kappa} A_i X B_i, \quad (1.3)$$

where A_i, B_i are elements of $B(\mathcal{H})$. A particular class of such maps, the so-called completely positive maps (or in physical terminology — *quantum operations*), plays a prominent role in formulation of *evolution* of open quantum systems and in the theory of *quantum measurement*. A comprehensive description of these problems from the physical point of view one can find in books [2, 13].

It is important to observe that most of the papers devoted to the classical Perron-Frobenius theory of nonnegative linear maps are concerned mainly with the existence problems. The purpose of this paper is, on the one hand, to connect the Frobenius theory of irreducible linear operators with the so-called Kraus representation of the completely positive linear maps and, on the other hand, to show that there exist some effective procedures which allow us to verify if a given quantum operation (i.e. a given completely positive map) is irreducible or not. Here by irreducibility we mean the natural generalization of this concept, introduced by Frobenius in his famous paper from 1912 year ([7]) and based on the specific block representation of nonnegative matrices, to a geometric approach formulated in terms of the invariance of faces of a fixed cone. In this context, by an effective procedure we understand a method which uses only a finite number of arithmetic operations on matrices A_i (Kraus coefficients) which define a fixed quantum operation (completely positive map)

$$\Phi(X) = \sum_{i=1}^{\kappa} A_i^* X A_i. \quad (1.4)$$

The paper is organized as follows. In Section 2 we will review the concepts needed for our discussion based on geometric and operator theoretic formulation of the Perron-Frobenius theory and, on the other hand, we describe the structure of the cone of positive definite operators defined on a given Hilbert space. Moreover, some properties of faces of this cone are analyzed. Section 3 describes some properties of amplifications of general positive maps (superoperators). The main results of the paper are discussed in Section 4, where we formulate an effective method of checking if a given superoperator, i.e., a fixed quantum operation is irreducible or not.

2 Spectral properties of maps which have invariant cones

First, let us recall some results of the classical theory of nonnegative matrices. In this theory a prominent role is played by so-called irreducible matrices. By definition, a matrix A is irreducible if no permutation matrix P exists such that

$$P^T A P = \begin{pmatrix} X & Z \\ 0 & Y \end{pmatrix}, \quad (2.1)$$

where X and Y are square matrices and 0 denotes a block of zeros. There are three main categories of results in Perron and, respectively, Frobenius approach to linear operators which preserve the nonnegative orthant \mathbb{R}_+^n :

C I. If A is strictly positive matrix, $A > 0$, i.e., all entries of A satisfy the inequality $a_{ij} > 0$, then

- a) the spectral radius of the matrix A , $r(A)$, is a simple eigenvalue of A , greater than the magnitude of any other eigenvalue;
- b) there exists a corresponding eigenvector which is positive (componentwise);
- c) if $A \leq B$ and $A \neq B$, then $r(A) < r(B)$.

C II. If A is nonnegative matrix, $A \geq 0$, that is some entries a_{ij} can be equal to zero, then

- a) the spectral radius $r(A)$ of the matrix A is an eigenvalue of A ;
- b) there exists a corresponding eigenvector which is nonnegative;
- c) if $A \leq B$, then $r(A) \leq r(B)$.

C III. If A is irreducible and nonnegative, $A \geq 0$, then we have

- a) $r(A)$ is a simple eigenvalue;
- b) there exists a corresponding eigenvector which is positive;
- c) if $A \leq B$ and $A \neq B$, then $r(A) < r(B)$.

Simple proofs of the Perron-Frobenius results for matrices can be found, e.g., in [16]. Let us observe that the assumption of irreducibility (Frobenius) of the nonnegative matrices gives us a reproduction of results obtained by Perron for positive matrices. The generalizations mentioned in the Introduction involve two important extensions of the classical theory. First, the considered spaces are very often assumed to be infinite dimensional and second, positivity (nonnegativity) is replaced by the assumption that the considered operator leaves a fixed cone invariant. In this paper we concentrate on this second case and, in order to obtain stronger results, important for quantum information theory, we consider only finite dimensional cases.

Now, let us recall some notions and results important in analyzing partially ordered vector spaces. Let V be a normed and partially ordered real linear space with a fixed cone K , i.e., K is a nonempty closed subset of V which satisfies :

K I $K + K \subseteq K$;

K II. $\lambda K \subseteq K$ for all $\lambda \geq 0$;

K III. $K \cap (-K) = \{\mathbf{0}\}$,

where $\mathbf{0}$ denotes the zero element of V . If we have the equality $V = K - K$, then the cone K is called *generating* or *reproducing* (sometimes one also uses the term *full cone*).

If Φ is a linear transformation on V , $\Phi : V \rightarrow V$, then we denote by $r(\Phi)$ the *spectral radius* of Φ , i.e.,

$$r(\Phi) := \max\{|\lambda|; \lambda \in \sigma(\Phi)\}, \quad (2.2)$$

where $\sigma(\Phi)$ denotes the spectrum of Φ . For any cone K we let K° denote the interior of K and by ∂K we denote its boundary.

As is well known any fixed cone K in V determines a partial order in V . For this order we use the following terminology:

1. x is *nonnegative*, $x \geq 0$, iff $x \in K$;

2. x is *positive*, $x > 0$, iff $x \geq 0$ and $x \neq 0$;
3. x is *strictly positive*, $x \gg 0$, iff $x \in K^\circ$.

Now, let us define the concept of face which plays a basic role in the theory of irreducible operators. Let K be a cone in V . By a *face* F of K one understands a subset of K which is a cone and satisfies an extra condition: if $0 \leq y \leq x$ and $x \in F$, then y also belongs to F , $y \in F$.

Of course, if we fixed a basis in V , then we may regard vectors in V as column vectors in \mathbb{R}^n . In this case the positive orthant \mathbb{R}_+^n constitutes a cone in \mathbb{R}^n and exact description of $\partial\mathbb{R}_+^n$ is obvious, namely,

$$F_M := \{x \in \mathbb{R}_+^n; x_i = 0 \text{ if } i \notin M\}, \quad (2.3)$$

where $M \subseteq \{1, 2, \dots, n\}$.

If $E \subset K$, then we will denote by $\Omega(E)$ the intersection of all faces containing E . It is easily seen that $\Omega(E)$ is a face. It is called the *face generated by E* .

The set of all operators $\Phi : V \rightarrow V$ such that $\Phi(K) \subseteq K$ we will denote by $\Pi(K)$. Let $B(V)$ be the set of all linear operators on V . Then we have

$$\Pi(K) := \{\Phi; \Phi(K) \subseteq K\} \subseteq B(V) \quad (2.4)$$

and $\Pi(K)$ is a cone in $B(V)$. The elements of the set $\Pi(K)$ are said to be *K -nonnegative operators*. In particular, the operator Φ is called *K -positive* in case $\Phi(K \setminus \{0\}) \subseteq K^\circ$. The set of all K -positive maps will be denoted by $\Pi^+(K)$.

Now we introduce one of the main ideas of the Perron-Frobenius theory both in classical and quantum case. For a fixed K in V a natural generalization of the concept of an irreducible matrix is the following: Φ is K -irreducible if and only if Φ leaves invariant no face of K except $\{0\}$ and K itself. In other words, an operator in $\Pi(K)$ is K -reducible iff it leaves invariant a nontrivial face of K .

Another, strictly equivalent, definition of K -irreducibility can be given by the following theorem: An operator $\Phi \in \Pi(K)$ is K -irreducible if and only if no eigenvector of Φ lies on the boundary of K . In fact, one can say even more: An operator $\Phi \in \Pi(K)$ is K -irreducible if and only if Φ has exactly one (up to scalar multiples) eigenvector in K and this vector belongs to K° . Moreover, for any proper cone K we have

$$\Pi^+(K) \subseteq \tilde{\Pi}(K) \subseteq \Pi(K), \quad (2.5)$$

where $\tilde{\Pi}(K)$ denotes the set of all K -irreducible operators. If $K = \mathbb{R}_+^n$, then the both definitions, Frobenius one and the above, coincide. For details, see e.g. [10, 14]

Some important spectral properties of K -nonnegative operators are summarized in the following theorems, which one can consider as natural generalizations of CI, CII and CIII.

Theorem 1. Let $\Phi \in \Pi^+(K)$. Then we have

- a) the spectral radius of the operator Φ is a simple eigenvalue of Φ , greater than the magnitude of any other eigenvalue;
- b) an eigenvector of Φ corresponding to $r(\Phi)$ belongs to K° ;
- c) no other eigenvector of Φ (up to scalar multiples) belongs to K .

Theorem 2. Let $\Phi \in \Pi(K)$. Then the following hold

- a) $r(\Phi)$ is an eigenvalue of Φ ;
- b) K contains an eigenvector of Φ corresponding to $r(\Phi)$;
- c) if $\Phi \leq \Psi$, then $r(\Phi) \leq r(\Psi)$.

Theorem 3. Let $\Phi \in \tilde{\Pi}(K)$. Then the following hold

- a) $r(\Phi)$ is a simple eigenvalue of Φ ;
- b) no eigenvector of Φ lies on the boundary of K ;
- c) Φ has exactly one (up to scalar multiples) eigenvector in K and this vector belongs to K° ;
- d) $(I + \Phi)^{n-1} \in \Pi^+(K)$, where $n = \dim V$.

For proofs of the above theorems consult ([9, 10, 16]).

We will conclude this Section by some comments on the two special cases which are important from the point of view of physics, namely, $K = \mathbb{R}_+^n$ and $K = M_n^+(\mathbb{C})$. Here, the last symbol denotes the set of all semipositive operators on the space \mathbb{C}^n , and \mathbb{C}^n is regarded as a representation of an n -dimensional Hilbert space.

For the case $K = \mathbb{R}_+^n$, the whole story reduces to the classical Perron-Frobenius theory and in physics we use this theory for the so-called mesoscopic description of classical systems.

The case $K = M_n^+(\mathbb{C})$ plays a fundamental role in description of representations of states for open quantum systems. In this case, $B(\mathcal{H})$ is an n^2 - dimensional vector space. We will denote by $B_*(\mathcal{H})$ the set of all self-adjoint operators on \mathcal{H} which can be naturally considered as n^2 - dimensional real Banach space. At the same time, $B(\mathcal{H})$ can be regarded as a Hilbert space with the scalar product defined by

$$[A, B] := \text{Tr}(A^*B). \quad (2.6)$$

The vector space $B(\mathcal{H})$ with scalar product (2.6) is called the Hilbert-Schmidt space. The vector space $B_*(\mathcal{H})$ of all Hermitian (self-adjoint) operators on \mathcal{H} constitutes an n^2 - dimensional, real subspace of the Hilbert-Schmidt space. One can use the "internal structure" of vectors from $B_*(\mathcal{H})$ to define a positive cone. By definition, a *semipositive* element of $B_*(\mathcal{H})$ is an operator A on \mathcal{H} such that $\langle \psi | A | \psi \rangle$ is real and nonnegative for all vectors $|\psi\rangle$ from \mathcal{H} . Of course, one can equivalently define a positive element of $B(\mathcal{H})$ as a self-adjoint operator with nonnegative eigenvalues. The set of all semipositive operators on \mathcal{H} we will denote by $B_*^+(\mathcal{H})$ or *PDS*. In particular, if we have the inequality $\langle \psi | A | \psi \rangle > 0$ for all $|\psi\rangle$ from \mathcal{H} , then we say that A is *positive*.

Let \mathcal{P}_n denote the set of all orthogonal projections, i.e., $A \in \mathcal{P}_n$ if and only if $A \in B_*(\mathcal{H})$ and $A^2 = A$. With the natural order on projections, namely, $A \leq B$ iff $Im(A) \subseteq Im(B)$, the mapping from $A \in \mathcal{P}_n$ to $F(PDS)$, where $F(K)$ denotes the set of all faces of K , is an order preserving isomorphism. Now, we formulate an important characterization of all faces of any cone PSD . It appears, that all faces of PSD_n , where the suffix n denotes $\dim \mathcal{H}$, are isomorphic to PSD_m for $0 \leq m \leq n$.

Theorem 4. *If $B \in PSD$ is of rank r , then there exists a unitary U such that*

$$\Omega(B) = U^*(PSD_m \oplus 0_{n-m})U. \quad (2.7)$$

Conversly, if U denotes a unitary operator on \mathcal{H} , then $U^(PSD_m \oplus 0_{n-m})U$ for $m = 0, 1, \dots, n$ are faces of PSD_n .*

It is well known that if K is a polyhedral cone, then for all faces of K we have

$$\text{span } F + \text{span } F^\triangleleft = V, \quad (2.8)$$

where F^\triangleleft denotes the so-called complementary face of F defined by

$$F^\triangleleft := \{z \in K^*; \langle z, x \rangle = 0 \text{ for all } x \in F\} \quad (2.9)$$

and V denotes the ambient space of K . The *residual subspace* of F is meant to measure "to what extent F is nonpolyhedral". It is defined as

$$\text{res}(F) := (\text{span } F + \text{span } F^\triangleleft)^\perp. \quad (2.10)$$

A list of several examples of cones, along with the description of their faces and residual subspaces, is contained in [23].

3 Irreducibility of positive maps on PDS

It is well known that if a linear map $\Phi : B(\mathcal{H}) \rightarrow B(\mathcal{H})$ sends the set $B_*(\mathcal{H})$ of all hermitian elements of $B(\mathcal{H})$ into itself, then Φ can be represented in the form

$$\Phi(X) = \sum_{i=1}^{\kappa} a_i A_i^* X A_i, \quad (3.1)$$

where $A_i \in B(\mathcal{H})$, and a_i for $i = 1, 2, \dots, \kappa$ are real numbers [2, 20]. In general, all maps of the above form are hermitian-preserving. However, the representation (3.1) is not unique. In general, for a given Φ , there exist many possible representations of the form (3.1). For a given Φ the smallest κ in (3.1) is called the *minimal length* of Φ and this minimal length is always smaller or equal to $(\dim \mathcal{H})^2$. If we assume that the operators A_i for $i = 1, 2, \dots, \kappa$ are linearly independent, then κ in (3.1) must be minimal.

According to the general definition introduced in Section 2 a *positive map* Φ (*PSD-positive*) is a linear map from $B(\mathcal{H})$ into itself, which

leaves PSD invariant. Now, Φ is called *k-positive* if its *k-amplification* $\Phi_{(k)} := \mathbb{I}_k \otimes \Phi$ that is the map

$$\mathbb{I}_k \otimes \Phi : M_k(\mathbb{C}) \otimes B(\mathcal{H}) \rightarrow M_k(\mathbb{C}) \otimes B(\mathcal{H}) \quad (3.2)$$

is positive. Here $M_k(\mathbb{C})$ denotes as usual the set of all $k \times k$ complex matrices. It is not difficult to observe that we can identify the set $M_k(\mathbb{C}) \otimes \mathcal{A}$, where for simplicity we denote the algebra $B(\mathcal{H})$ by \mathcal{A} , with the set of all $k \times k$ matrices $M_k(\mathcal{A})$ with entries from \mathcal{A} . In such notation one can represent $\Phi_{(k)} : M_k(\mathcal{A}) \rightarrow M_k(\mathcal{A})$ by

$$\Phi_{(k)} \left(\begin{array}{ccc} \cdots & \vdots & \cdots \\ \cdots & X_{ij} & \cdots \\ \cdots & \vdots & \cdots \end{array} \right) := \left(\begin{array}{ccc} \cdots & \vdots & \cdots \\ \cdots & \Phi(X_{ij}) & \cdots \\ \cdots & \vdots & \cdots \end{array} \right). \quad (3.3)$$

The map Φ is called completely positive if it is *k-positive* for all $k = 1, 2, \dots$. This terminology goes back to Stinespring [19], cf. also [4]. It is well known that for n -dimensional Hilbert space \mathcal{H} , n -positive maps on $B(\mathcal{H})$ are already completely positive [20].

Let us observe that all hermitian-preserving maps which are not only positive but completely positive can be written in the form (3.1) with positive a_i , $i = 1, \dots, \kappa$, i.e. by

$$\Phi(X) = \sum_{i=1}^{\kappa} K_i^* X K_i, \quad (3.4)$$

where $K_i := \sqrt{a_i} A_i$, and $\kappa \leq n^2$. The above expression is called the Kraus representation of a completely positive map Φ and, in case of the finite-dimensional Hilbert space \mathcal{H} , can be regarded also as a definition of the completely positive map. This representation is very useful in quantum information theory. In particular, completely positive maps are used to describe quantum operations, quantum channels and to model quantum devices.

It was shown by R. Timoney [21], that a positive map Φ (positive superoperator) which is m -positive, where $m = \lfloor \sqrt{\kappa} \rfloor$ must be completely positive. Here $\lfloor \sqrt{\kappa} \rfloor$ denotes the integer part of the number $\sqrt{\kappa}$. In other words, if a positive map $\Phi : B(\mathcal{H}) \rightarrow B(\mathcal{H})$, $\dim \mathcal{H} = n$, has a minimal length κ and Φ is m -positive, for some $m < n$ such that $(m+1)^2 > \kappa$, then Φ is already completely positive (cf. also [17]).

Now, let us observe that one can apply the results stated in Theorems 1 – 3 from Section 2, to the particular cone PSD_n in $B(\mathcal{H})$, $\dim \mathcal{H} = n$. In particular, Theorem 3 describes properties of irreducible superoperators and according to this theorem we have: for irreducible Φ , the spectral radius of Φ is a simple eigenvalue of the superoperator and Φ has exactly one (up to scalar coefficient) eigenvector in PSD and this vector belongs to $PSD \setminus \partial PSD$. One can say even more [22].

Theorem 5 (Farenick). *The following statements are equivalent for a positive map on PSD .*

- 1.) *There is a nontrivial (that is different from $\{0\}$ and PSD) face of PSD that is invariant under Φ ;*
- 2.) *There is nontrivial projection $P \in \mathcal{P}_n$ and a positive real number $\lambda > 0$ such that $\Phi(P) \leq \lambda P$;*
- 3.) *There is a nontrivial projection $P \in \mathcal{P}_n$ such that subalgebra $P(\text{PSD})P$ is invariant under Φ .*

In order to produce nontrivial examples of irreducible positive maps and in certain cases to characterize all irreducible maps within the class of completely positive maps we will use the following consequences of the Kraus representations of completely positive maps.

A family of closed subspaces of a given Hilbert space is called trivial if this family contains only $\{0\}$ and \mathcal{H} . For a fixed operator $X \in B(\mathcal{H})$ we will denote by $\text{Inv}(X)$ the set of all invariant subspaces of X . Now, we can state the following theorem which is a reformulation of some results from [22].

Theorem 6. *Let Φ denote a superoperator on $B(\mathcal{H})$ which is PSD-positive. If Φ is completely positive, then there exist some operators A_1, \dots, A_κ such that $\Phi(X) = \sum_j A_j^* X A_j$. Completely positive Φ is irreducible if and only if the Kraus operators A_j do not have a nontrivial common invariant subspace in \mathcal{H} .*

To better understand the above theorem, let us observe that if Kraus operators do not have a common invariant subspace, i.e., are such that $\cap_j \text{Inv}(A_j)$ is trivial and $\Phi(P) \leq \lambda P$ for some $\lambda \geq 0$ and $P \in \mathcal{P}_n$, then we have

$$\langle \Phi(P)\psi | \psi \rangle = \sum_{j=1}^{\kappa} \langle P A_j \psi | A_j \psi \rangle \leq \lambda \langle P \psi | \psi \rangle. \quad (3.5)$$

The left-hand side of the above equality is nonnegative for all $\psi \in \mathcal{H}$. On the other hand for $\psi \in \ker P$, we have $\langle P \psi | \psi \rangle = 0$ on the right-hand side. In this way the equality (3.5) implies that $\langle P A_j \psi | A_j \psi \rangle = 0$ for each $j = 1, \dots, \kappa$. This means that $\langle P A_j \psi | P A_j \psi \rangle = 0$, if we remember that $P^2 = P$, $P^* = P$. In consequence, $\ker P \in \cap_j \text{Inv}(A_j)$, that is $\ker P$ is either $\{0\}$ or \mathcal{H} .

4 Effective investigation of superoperators

According to Theorem 6 any completely positive map Φ (superoperator Φ) on $B(\mathcal{H})$ can be represented by a set of operators A_1, \dots, A_κ and the superoperator Φ is irreducible if the operators A_j ($j = 1, \dots, \kappa$) do not have a nontrivial common invariant subspace in the Hilbert space \mathcal{H} . Let us note that if operators (matrices) A_j have a common invariant subspace then they must have also a common eigenvector. So to give an effective criterion for irreducibility of Φ it is sufficient to analyze the problem of a common eigenvector for a set of operators A_1, \dots, A_κ .

Let us begin with the problem of the existence of a common eigenvector for a pair of matrices. It is rather strange that such a natural problem was not solved until 1980s [24]. The answer to the problem can be formulated in the following way.

Theorem 7 (Shemesh). *Let A and B denote two matrices acting on $\mathcal{H} \cong \mathbb{C}^n$. A common eigenvector of A and B exists if and only if the subspace \mathcal{M} defined by*

$$\mathcal{M} := \bigcap_{k,l=1}^{n-1} \ker[A^k, B^l] \quad (4.1)$$

is of positive dimension. Here the symbol $[\cdot, \cdot]$ denotes the commutator of the matrices.

As it was stressed in [25] the genuine meaning of the subspace \mathcal{M} can be stated as follows.

Theorem 8. *A subspace \mathcal{M} is invariant with respect to both matrices A and B , and moreover, A and B commute on \mathcal{M} . Every subspace of \mathcal{H} , which is invariant under A and B and on which A and B commute is contained in \mathcal{M} .*

The condition of Theorem 7, that is $\dim \mathcal{M} > 0$, can be formulated in a constructive form. To this end let us define the matrix

$$S := \sum_{k,l=1}^{n-1} ([A^k, B^l])^* [A^k, B^l]. \quad (4.2)$$

Now we can formulate Theorem 7 as follows.

Theorem 9. *The matrices A and B do not have a common eigenvector if and only if $\det S > 0$.*

Indeed, it is enough to notice that matrix S is semipositive and matrices A and B have a common invariant subspace iff $\det S = 0$.

In some cases the calculations in (4.2) can be reduced. Let us observe that if the degree of the minimal polynomial of A is m and the degree of the minimal polynomial is p , then the subspace \mathcal{M} is given by

$$\mathcal{M} = \bigcap_{k,l} \ker[A^k, B^l], \quad (4.3)$$

where $k = 1, \dots, m-1$, and $l = 1, \dots, p-1$. This makes it possible to reduce the number of terms in (4.3) to $(m-1)(p-1)$ instead of $(n-1)^2$.

Conclusions. If in the Kraus representation of any completely positive map Φ at least two Kraus operators do not have a common eigenvector, then the map Φ is irreducible. It is not difficult to check, that if A is an irreducible operator in $\mathcal{B}(\mathcal{H})$, and if Ω is an arbitrary superoperator on $\mathcal{B}(\mathcal{H})$, then the map Φ defined by

$$\Phi(X) = AXA^* + A^*XA + \Omega(X), \quad (4.4)$$

is an irreducible positive linear map. It means we have a very simple and explicit method for constructing irreducible positive superoperators on $\mathcal{B}(\mathcal{H})$.

References

- [1] A. Jamiólkowski, Rep. Math. Phys. **3**, 275 (1972); ibidem **5**, 415 (1975).
- [2] I. Bengtsson, K. Życzkowski, *Geometry of Quantum States*, Cambridge Univ. Press, 2008.
- [3] D. Bruss, G. Leuchs (Eds.), *Lectures on Quantum Information*, Wiley-VCH, 2007.
- [4] V. Vedral, *Introduction to Quantum Information Science*, Oxford Univ. Press, 2006.
- [5] F. Benatti, *Dynamics, Information and Complexity in Quantum Systems*, Springer, 2009.
- [6] O. Perron, Math. Ann. **64**, 248 (1907).
- [7] G. Frobenius, S. B. Preuss. Akad. Wiss. (Berlin), 256 (1912).
- [8] M. G. Krein, M. A. Rutman, *Linear Operators Leaving Invariant a Cone in a Banach Space*, American Mathematical Society Translations **26**, (1950).
- [9] G. Birkhoff, Amer. Math. Monthly **74**, 274 (1967).
- [10] J. S. Vandergraft, SIAM J. App. Math. **16**, 1208 (1968).
- [11] K. Kraus, Ann. Phys. **64**, 119, (1971).
- [12] A. Kossakowski, Rep. Math. Phys. **3**, 243 (1972).
- [13] M. A. Nielsen, I. Chuang, *Quantum Computation and Quantum Information*, Cambridge Univ. Press., 2000.
- [14] G.P. Barker, Linear Alg. Appl. **39**, 263, (1981).
- [15] R.S. Varga, *Matrix Iterative Analysis*, Springer, 2000.
- [16] A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, 1979 .
- [17] A. Jamiólkowski, Open Sys. Infor. Dyn. **11** 385 (2004).
- [18] A. Jamiólkowski, pp. 185-197, in: *Quantum Bio-Informatics IV*, World Scientific (2011).
- [19] W. F Stinespring, Proc. AMS **6**, 211, (1955).
- [20] M. D. Choi, Linear Alg. Appl. **10**, 285, (1975).
- [21] R. M. Timoney, Bull. London Math. Soc. **32**, 229, (2000).
- [22] D. R. Farenick, Proc. AMS **124**, 3381, (1996).
- [23] H. Wolkowicz, R. Saigal, L. Vandenberghe, Eds., *Handbook of Semidefinite Programming*, Kluwer, 2003.
- [24] D. Shemesh, Linear Alg. Appl. **62**, 11, (1984).
- [25] Yu. A. Alpin, Kh. D. Ikramov, J. of Math. Sciences **114**, 1757, (2003).

Grassmann manifold and spectral theory of 1-D Schrödinger operators

Shinichi KOTANI
Kwansei Gakuin University

1 Introduction

The purpose of this article is to give a concise introduction to Sato's theory on Grassmann manifold in a closed form. The theory was initiated by M.Sato and developed by his colleagues to give a unified approach to many completely integrable systems including KdV equation. Originally this theory was presented in a quite algebraic form, and later reformulated analytically by Segal-Wilson[8]. However, a kind of algebraic flavor still remains, which makes it difficult for researchers in some fields to understand the theory completely. This article was written for readers who are not familiar with algebra nor geometry, and intended to discuss the theory from a view point of spectral theory of underlying 1-D Schrödinger operators. In this respect the author has introduced several notions, namely **characteristic functions**, **monodromy matrix**, **Weyl function**, corresponding to the spectral theory. On the other hand, the author published a paper for this purpose in [11] and found it contains several wrong statements. In this note he tries to clarify them.

The contents is as follows.

Contents

- Grassmann manifold on L^2 ($|z| = 1$)
 - Grassmann manifold and Schrödinger operators
 - Characteristic functions of subspaces W
- Group action on $Gr^{(2)}(H)$
 - τ -function
 - Monodromy matrix
 - Baker-Akhiezer function and Weyl function
- KdV flow
- Properties of KdV flow
 - Iso-spectral property
 - Invariant submanifolds
 - Almost periodicity

2 Grassmann manifold on $L^2(|z| = 1)$

2.1 Grassmann manifold and Schrödinger operators

In this section we explain briefly why Grassmann manifold has connection with 1-D Schrödinger operators. Segal-Wilson[8] introduced the following infinite dimensional Grassmann manifold. Let $H = L^2(|z| = 1)$ and H_{\pm} be closed subspaces of H defined by

$$H_+ = \left\{ f; f(z) = \sum_{n \geq 0} f_n z^n \right\}, \quad H_- = \left\{ f; f(z) = \sum_{n \geq 1} f_n z^{-n} \right\}.$$

Denote the orthogonal projections to H_{\pm} by $P_{H_{\pm}}$. A closed subspace W of H is an element of $Gr^{(2)}(H)$ if and only if it satisfies

- (i) $P_{H_-} : W \rightarrow H_-$ is of Hilbert-Schmidt class.
- (ii) If $f \in W$, then $z^2 f \in W$.
- (iii) $P_{H_+} : W \rightarrow H_+$ is bijective.

H_+ is the simplest example. The condition(i) can be replaced by another mild condition, for instance, compact operators. The condition(ii) is due to our being interested in 1-D Schrödinger operators.

$Gr^{(2)}(H)$ has a natural group action. Let Γ be

$$\Gamma = \left\{ g(z) = e^{h(z)}; h \text{ is holomorphic on } |z| < 1 \text{ and } h'(e^{i\theta}) \in L^2, h(0) = 0 \right\}.$$

Then Γ becomes a commutative group with natural product of functions. For $g \in \Gamma$ a new subspace $g^{-1}W$ of H arises, which satisfies the properties (i),(ii) but (iii). Set

$$e_x(z) = e^{-xz} \in \Gamma.$$

Without difficulty one can see that $e_x^{-1}W \in Gr^{(2)}(H)$ holds for every sufficiently small x . Then the property (iii) for $e_x^{-1}W$ implies the unique existence of $f_W(x, \cdot) \in W$ such that $P_{H_+}(e_x^{-1}f) = 1$. From the definition $f_W(x, z)$ has an expansion as a function of z

$$f_W(x, z) = e^{-xz} \left(1 + \frac{a_1(x)}{z} + \frac{a_2(x)}{z^2} + \dots \right),$$

and taking derivatives with respect to x yields

$$\left\{ \begin{array}{l} f'_W(x, z) = -ze^{-xz} \left(1 + \frac{a_1(x)}{z} + \frac{a_2(x)}{z^2} + \dots \right) \\ \quad + e^{-xz} \left(\frac{a'_1(x)}{z} + \frac{a'_2(x)}{z^2} + \dots \right), \\ f''_W(x, z) = z^2 e^{-xz} \left(1 + \frac{a_1(x)}{z} + \frac{a_2(x)}{z^2} + \dots \right) \\ \quad - 2e^{-xz} \left(a'_1(x) + \frac{a'_2(x)}{z} + \dots \right) + e^{-xz} \left(\frac{a''_1(x)}{z} + \frac{a''_2(x)}{z^2} + \dots \right), \end{array} \right.$$

hence

$$\begin{aligned} W \ni -f''_W(x, z) - 2a'_1(x)f_W(x, z) + z^2 f_W(x, z) \\ = e^{-xz} \sum_{k=1}^{\infty} (2a'_{k+1}(x) - 2a'_1(x)a_k(x) - a''_k(x)) z^{-k}, \end{aligned}$$

namely

$$\sum_{k=1}^{\infty} (2a'_{k+1}(x) - 2a'_1(x)a_k(x) - a''_k(x)) z^{-k} \in e_x^{-1}W$$

holds. From (iii) for $e_x^{-1}W$, for every $k \geq 1$

$$2a'_1(x)a_k(x) + a''_k(x) = 2a'_{k+1}(x), \quad -f''_W(x, z) - 2a'_1(x)f_W(x, z) + z^2 f_W(x, z) = 0$$

follows. Therefore, setting

$$q_W(x) = -2a'_1(x)$$

yields a time independent Schrödinger equation

$$-f''_W(x, z) + q_W(x)f_W(x, z) + z^2 f_W(x, z) = 0.$$

If we take $e^{-xz+4tz^3} \in \Gamma$ in place of e^{-xz} we obtain a solution to the KdV equation

$$\frac{\partial u}{\partial t} = -\frac{\partial^3 u}{\partial x^3} + 6u \frac{\partial u}{\partial x}.$$

In this way we can understand that $\{Gr^{(2)}(H), \Gamma\}$ generates the KdV hierarchy.

2.2 Characteristic functions of subspaces W

Decompose $f \in W$ as

$$f = f_+ + f_-, \quad f_{\pm} \in H_{\pm}.$$

Then, (iii) implies that f_+ determines f_- , hence there exists a linear map A from H_+ to H_- , namely

$$f = f_+ + Af_+.$$

If it is necessary, we use the notation A_W . From (i) it is easy to see that A is of Hilbert-Schmidt class. To investigate more precisely relationship between A and the multiplication z^2 define a bounded operator T on H_- by

$$Tf = P_{H_-} z^2 f.$$

If we expand $f \in H_-$ as

$$f(z) = \sum_{n \geq 1} f_n z^{-n},$$

then T is equivalently described by

$$(Tf)_n = f_{n+2} \quad \text{for every } n \geq 1.$$

Now, for $f \in H_+$

$$W \ni z^2(f + Af) = z^2 f + z^2 Af = (z^2 f + P_{H_+} z^2 Af) + P_{H_-} z^2 Af$$

is valid, hence we have an identity

$$A(z^2 f + P_{H_+} z^2 Af) = P_{H_-} z^2 Af,$$

Define another bounded operator S on H_- by

$$Sf = P_{H_-}(z^2 f) - AP_{H_+}(z^2 f).$$

Then we have

$$A(z^2 f) = SAf.$$

Introduce two functions $\{\varphi, \psi\}$ of H_- by

$$\varphi(z) = A1, \quad \psi(z) = Az.$$

Then

$$Sf = Tf - f_1\psi - f_2\varphi. \quad (1)$$

holds. For a function $f \in H$ denote

$$\begin{cases} f_e(z) = \sum_{n \in \mathbb{Z}} f_{2n} z^n = \frac{1}{2} (f(\sqrt{z}) + f(-\sqrt{z})), \\ f_o(z) = \sum_{n \in \mathbb{Z}} f_{2n+1} z^n = \frac{1}{2\sqrt{z}} (f(\sqrt{z}) - f(-\sqrt{z})). \end{cases}$$

Set

$$e_n(z) = z^n,$$

and

$$V = \overline{P_{H_-} W}.$$

Lemma 1 S, T act on V and V is generated by $\{T^m \varphi, T^n \psi\}_{m, n \geq 0}$. Moreover, the followings hold.

(i) $A(z^2 f) = SAf$, for every $f \in H_+$

(ii) $Ae_{2n} = S^n \varphi$, $Ae_{2n+1} = S^n \psi$

(iii) $Af = \sum_{n=0}^{\infty} f_{2n} S^n \varphi + \sum_{n=0}^{\infty} f_{2n+1} S^n \psi = f_e(S) \varphi + f_o(S) \psi$, and they converge in H_- .

Proof. (i) has been already proved. (ii) follows immediately from $Ae_0 = \varphi$, $Ae_1 = \psi$. From (ii), we have an identity

$$\begin{aligned} Af &= \sum_{n=1}^{\infty} f_n Ae_n = \sum_{n=0}^{\infty} f_{2n} Ae_{2n} + \sum_{n=0}^{\infty} f_{2n+1} Ae_{2n+1} \\ &= \sum_{n=0}^{\infty} f_{2n} S^n \varphi + \sum_{n=0}^{\infty} f_{2n+1} S^n \psi \quad \text{for } f \in H_+. \end{aligned}$$

The identity $AH_+ = P_{H_-} W$ shows that $SV \subset V$ and $\{S^m \varphi, S^n \psi\}_{m, n \geq 0}$ generates V . Then (1) shows that $TV \subset V$ and V is generated by $\{T^m \varphi, T^n \psi\}_{m, n \geq 0}$. ■

This lemma shows $\{\varphi, \psi\}$ determine uniquely the space W . We call $\{\varphi, \psi\}$ **characteristic functions** of W .

Remark 1 We remark here

$$\sum_{n=0}^{\infty} (\|S^n \varphi\|^2 + \|S^n \psi\|^2) = \sum_{n=0}^{\infty} \|Ae_n\|^2 = \|A\|_{H.S.}^2 < \infty.$$

Denote the resolvent operator of an operator K by $R_K(\lambda)$, namely

$$R_K(\lambda) = (\lambda - K)^{-1}.$$

Set

$$\Delta(z) = \begin{pmatrix} 1 + \varphi_e(z) & \psi_e(z) \\ \varphi_o(z) & 1 + \psi_o(z) \end{pmatrix}, \quad \delta(z) = \det \Delta(z). \quad (2)$$

δ is holomorphic on $\{|z| > 1\}$ and $\Delta(\infty) = I$, $\delta(\infty) = 1$. Since $\|T\| \leq 1$, generally

$$spT \subset \{|\lambda| \leq 1\}$$

is valid. Moreover, for $|\lambda| > 1$ note identities

$$\begin{cases} (R_T(\lambda) f, e_{-1}) = \sum_{n=0}^{\infty} \frac{(T^n f, e_{-1})}{\lambda^{n+1}} = \sum_{n=0}^{\infty} \frac{f_{2n+1}}{\lambda^{n+1}} = f_o(\lambda), \\ (R_T(\lambda) f, e_{-2}) = \sum_{n=0}^{\infty} \frac{(T^n f, e_{-2})}{\lambda^{n+1}} = \sum_{n=1}^{\infty} \frac{f_{2n}}{\lambda^n} = f_e(\lambda), \end{cases}$$

which imply

$$\Delta(\lambda) = \begin{pmatrix} 1 + (R_T(\lambda) \varphi, e_{-2}) & (R_T(\lambda) \psi, e_{-2}) \\ (R_T(\lambda) \varphi, e_{-1}) & 1 + (R_T(\lambda) \psi, e_{-1}) \end{pmatrix}.$$

Therefore, Δ, δ are extendable to $\mathbb{C} \setminus spT$ as holomorphic functions.

Lemma 2 *The resolvent operators of S, T satisfy the followings.*

- (i) *If $\lambda \in \mathbb{C} \setminus spT$, then $\delta(\lambda) \neq 0$ implies $\lambda \in \mathbb{C} \setminus spS$.*
- (ii) *The following identity holds.*

$$\Delta(\lambda)^{-1} = \begin{pmatrix} 1 - (R_S(\lambda) \varphi, e_{-2}) & -(R_S(\lambda) \psi, e_{-2}) \\ -(R_S(\lambda) \varphi, e_{-1}) & 1 - (R_S(\lambda) \psi, e_{-1}) \end{pmatrix}.$$

Proof. For $f \in H_-$ consider the equation

$$(\lambda - S)g = f, \quad g \in H_-, \quad (3)$$

which is equivalent to

$$f = \lambda g - Tg + (g, e_{-2})\varphi + (g, e_{-1})\psi,$$

due to (1). Therefore, if $\lambda \in \mathbb{C} \setminus spT$, then

$$R_T(\lambda)f = g + (g, e_{-2})R_T(\lambda)\varphi + (g, e_{-1})R_T(\lambda)\psi. \quad (4)$$

holds. Taking inner product in (4) with e_{-2}, e_{-1} , we have

$$\begin{cases} f_e(\lambda) = (1 + \varphi_e(\lambda))(g, e_{-2}) + \psi_e(\lambda)(g, e_{-1}), \\ f_o(\lambda) = \varphi_o(\lambda)(g, e_{-2}) + (1 + \psi_o(\lambda))(g, e_{-1}). \end{cases} \quad (5)$$

Therefore, the equation(3) is equivalent to (5) regarding $\{(g, e_{-1}), (g, e_{-2})\}$ as unknown variables. Clearly (5) is solvable for any $f \in H_-$ if and only if

$$\delta(\lambda) = \begin{vmatrix} 1 + \varphi_e(\lambda) & \psi_e(\lambda) \\ \varphi_o(\lambda) & 1 + \psi_o(\lambda) \end{vmatrix} \neq 0,$$

and in this case we have

$$\begin{pmatrix} (g, e_{-2}) \\ (g, e_{-1}) \end{pmatrix} = \Delta(\lambda)^{-1} \begin{pmatrix} f_e(\lambda) \\ f_o(\lambda) \end{pmatrix}.$$

Letting $f = \varphi, \psi$ yields

$$\begin{pmatrix} (R_S(\lambda)\varphi, e_{-2}) & (R_S(\lambda)\psi, e_{-2}) \\ (R_S(\lambda)\varphi, e_{-1}) & (R_S(\lambda)\psi, e_{-1}) \end{pmatrix} = \Delta(\lambda)^{-1} \begin{pmatrix} \varphi_e(\lambda) & \psi_e(\lambda) \\ \varphi_o(\lambda) & \psi_o(\lambda) \end{pmatrix} = I - \Delta(\lambda)^{-1},$$

which concludes (ii). ■

To deduce further information on $\{\varphi, \psi\}$ we investigate more closely the actions $S^n\varphi, S^n\psi$. Note

$$Tf(z) = z^2f(z) - f_1z - f_2, \quad \text{for } f(z) = \sum_{n=1}^{\infty} f_n z^{-n}.$$

Then

$$\begin{aligned} S\varphi &= T\varphi - a_2\varphi - a_1\psi = z^2\varphi - a_1z - a_2 - a_2\varphi - a_1\psi \\ &= (z^2 - a_2)(1 + \varphi) - a_1(z + \psi) - z^2, \end{aligned}$$

holds, where

$$\varphi(z) = \frac{a_1}{z} + \frac{a_2}{z^2} + \dots.$$

Now let us assume

$$S^n\varphi = p_n(z^2)(1 + \varphi) + q_n(z^2)(z + \psi) - e_{2n}. \quad (6)$$

with some polynomial $p_n(z), q_n(z)$. Then

$$\begin{aligned} S^{n+1}\varphi &= TS^n\varphi - (S^n\varphi, e_{-1})\psi - (S^n\varphi, e_{-2})\varphi \\ &= z^2S^n\varphi - (S^n\varphi, e_{-1})z - (S^n\varphi, e_{-2}) - (S^n\varphi, e_{-1})\psi - (S^n\varphi, e_{-2})\varphi \\ &= \{z^2p_n(z^2) - (S^n\varphi, e_{-2})\}(1 + \varphi) + \{z^2q_n(z^2) - (S^n\varphi, e_{-1})\}(z + \psi) - e_{2(n+1)} \end{aligned}$$

is valid. Therefore, if we set $p_0(z) = 1, q_0(z) = 0$ and

$$p_{n+1}(z) = zp_n(z) - (S^n\varphi, e_{-2}), \quad q_{n+1}(z) = zq_n(z) - (S^n\varphi, e_{-1}), \quad (7)$$

then we know that (6) holds for all $n \geq 0$. Set

$$X(\lambda) = \Delta(\lambda)^{-1} = \begin{pmatrix} s(\lambda) & t(\lambda) \\ u(\lambda) & v(\lambda) \end{pmatrix}.$$

and denote s_n, t_n, u_n, v_n be the n -th Laurent expansion of s, t, u, v respectively around $z = \infty$. From (7) it follows that

$$\begin{cases} p_n(z) = z^n \left(1 - \sum_{k=0}^{n-1} z^{-(k+1)} (S^k\varphi, e_{-2}) \right), \\ q_n(z) = -z^n \left(\sum_{k=0}^{n-1} z^{-(k+1)} (S^k\varphi, e_{-1}) \right), \end{cases}$$

hence

$$p_n(z) = z^n s_n(z), \quad q_n(z) = z^n u_n(z).$$

Similarly we see

$$S^n \psi = z^{2n} t_n(z^2) (1 + \varphi(z)) + z^{2n} v_n(z^2) (z + \psi) - e_{2n+1}.$$

Summing up the argument, we see that Az^n can be expressed by $\{\varphi, \psi\}$.

Lemma 3 *If we denote by $\{s_n, t_n, u_n, v_n\}_{n \geq 0}$ the n -th Laurent expansion of s, t, u, v around $z = \infty$ respectively, then*

$$\begin{cases} Ae_{2n} = z^{2n} s_n(z^2) (1 + \varphi(z)) + z^{2n} u_n(z^2) (z + \psi(z)) - e_{2n}, \\ Ae_{2n+1} = z^{2n} t_n(z^2) (1 + \varphi(z)) + z^{2n} v_n(z^2) (z + \psi(z)) - e_{2n+1}. \end{cases} \quad (8)$$

Now set $l_n = Ae_n$. Then (8) implies

$$\begin{cases} s_n(z^2) (1 + \varphi(z)) + u_n(z^2) (z + \psi(z)) = 1 + z^{-2n} l_{2n}(z) \\ t_n(z^2) (1 + \varphi(z)) + v_n(z^2) (z + \psi(z)) = z + z^{-2n} l_{2n+1}(z) \end{cases} \quad (9)$$

Set

$$X_n(z) = \begin{pmatrix} s_n(z) & t_n(z) \\ u_n(z) & v_n(z) \end{pmatrix}.$$

Then (9) is equivalent to

$$\Delta(z) X_n(z) - I = z^{-2n} \begin{pmatrix} l_{2n,e}(z) & l_{2n+1,e}(z) \\ l_{2n,o}(z) & l_{2n+1,o}(z) \end{pmatrix}. \quad (10)$$

Theorem 1 *For $W \in Gr^{(2)}(H)$ the corresponding Δ satisfies the following conditions:*

- (i) *For every $|z| > 1$, $\delta(z) \neq 0$ and $\delta(e^{i\theta}) \neq 0$ hold for a.e. $\theta \in [0, 2\pi)$.*
- (ii) *Set $X(z) = \Delta(z)^{-1}$ and let $X_n(z)$ be its Laurent expansion around ∞ up to n . Then*

$$\sum_{n=1}^{\infty} \int_0^{2\pi} \|\Delta(e^{i\theta}) X_n(e^{i\theta}) - I\|^2 d\theta < \infty$$

holds.

Conversely, if a 2×2 square matrix $\Delta(z)$ defined by (2) with some $\{\varphi, \psi\}$ of H_- satisfies the conditions (i), (ii), then $\{\varphi, \psi\}$ become characteristic functions of a $W \in Gr^{(2)}(H)$, A linear operator A_W from H_+ to H_- defined by (8) can be extended as an operator of Hilbert-Schmidt class from H_+ to H_- , and W is given by

$$W = \{f + A_W f; \quad f \in H_+\} \in Gr^{(2)}(H).$$

Proof. Using Hilbert-Schmidt matrix norm $\|\cdot\|$, for $|z| = 1$ we see

$$\begin{aligned} 4 \|\Delta(z^2) X_n(z^2) - I\|^2 &= |l_{2n}(z) + l_{2n}(-z)|^2 + |l_{2n}(z) - l_{2n}(-z)|^2 \\ &\quad + |l_{2n+1}(z) + l_{2n+1}(-z)|^2 + |l_{2n+1}(z) - l_{2n+1}(-z)|^2. \end{aligned}$$

Observing even functions and odd functions in $L^2(|z| = 1)$ are orthogonal, we have

$$2 \int_0^{2\pi} \|\Delta(e^{i\theta}) X_n(e^{i\theta}) - I\|^2 d\theta = \int_0^{2\pi} |l_{2n}(e^{i\theta})|^2 d\theta + \int_0^{2\pi} |l_{2n+1}(e^{i\theta})|^2 d\theta.$$

On the other hand, by the orthonormal basis $\{e_n\}_{n \geq 0}$ of H_+ we can calculate the Hilbert-Schmidt norm of A_W as

$$\|A_W\|_{H.S.}^2 = \sum_{n=0}^{\infty} \frac{1}{2\pi} \int_0^{2\pi} |l_n(e^{i\theta})|^2 d\theta.$$

This combined with the above calculation gives

$$\sum_{n=0}^{\infty} \frac{1}{\pi} \int_0^{2\pi} \|\Delta(e^{i\theta}) X_n(e^{i\theta}) - I\|^2 d\theta = \|A_W\|_{H.S.}^2 < \infty.$$

Cauchy's theorem shows that, for $|z| < 1$

$$\Delta(z^{-1}) X_n(z^{-1}) - I = \frac{1}{2\pi} \int_0^{2\pi} \frac{\Delta(e^{-i\theta}) X_n(e^{-i\theta}) - I}{e^{i\theta} - z} e^{i\theta} d\theta$$

is valid, hence $\Delta(z^{-1}) X_n(z^{-1})$ and its determinant converge to I and 1 uniformly as $n \rightarrow \infty$ on each compact set of the open unit disk. Therefore, if, for a z_0 of the disc $\det \Delta(z_0^{-1}) = 0$ holds, then it leads us to a contradiction, hence $\Delta(z)$ satisfies (1). Similarly we see that on the unit circle, for almost every θ , $\Delta(e^{-i\theta})$ is non-singular. By (10) we can identify $X_n(z)$ with the n -th Taylor expansion of $\Delta(z)^{-1}$. The latter half of the assertion of the theorem is already clear in the argument before the statement of the theorem. ■

Corollary 1 *The spectrum of S is contained in the closed unit disk.*

Proof. The proof is immediate from (i) of Lemma2 and (i) of Theorem1 ■

As we have seen in Lemma3

$$W \ni e_n + Ae_n = p_n(z^2)(1 + \varphi(z)) + q_n(z^2)(z + \psi(z))$$

holds with some polynomials p_n, q_n for any $n \geq 0$. We try to find a representation of this form for a general $\chi \in W$. Consider an equation

$$p(z^2)(1 + \varphi(z)) + q(z^2)(z + \psi(z)) = \chi(z),$$

which is equivalent to

$$\begin{cases} (1 + \varphi_e(z))p(z) + \psi_e(z)q(z) = \chi_e(z), \\ \varphi_o(z)p(z) + (1 + \psi_o(z))q(z) = \chi_o(z). \end{cases}$$

Then, due to (i) of Theorem1 $\Delta(z)$ is non-singular for a.e. $|z| = 1$, hence

$$\begin{pmatrix} p(z) \\ q(z) \end{pmatrix} = \Delta(z)^{-1} \begin{pmatrix} \chi_e(z) \\ \chi_o(z) \end{pmatrix} \quad (11)$$

is valid. However, it is not clear that the RHS of (11) are extendable as a holomorphic function into the unit disk. Therefore, instead of (11) we use (8), since identities

$$\begin{aligned} f(z) + Af(z) &= \sum_{n=1}^{\infty} f_n(z^n + Az^n) \\ &= \sum_{n=0}^{\infty} f_{2n} z^{2n} s_n(z^2)(1 + \varphi(z)) + \sum_{n=0}^{\infty} f_{2n} z^{2n} u_n(z^2)(z + \psi(z)) \\ &\quad + \sum_{n=0}^{\infty} f_{2n+1} z^{2n} t_n(z^2)(1 + \varphi(z)) + \sum_{n=0}^{\infty} f_{2n+1} z^{2n} v_n(z^2)(z + \psi(z)) \end{aligned}$$

hold at least formally. We prepare bounds for these terms to know the convergence.

Lemma 4 Define p, q by

$$\begin{pmatrix} p(z) \\ q(z) \end{pmatrix} = \begin{pmatrix} \sum_{n=0}^{\infty} f_{2n} z^{2n} s_n(z^2) + \sum_{n=0}^{\infty} f_{2n+1} z^{2n} t_n(z^2) \\ \sum_{n=0}^{\infty} f_{2n} z^{2n} u_n(z^2) + \sum_{n=0}^{\infty} f_{2n+1} z^{2n} v_n(z^2) \end{pmatrix} \quad (12)$$

for an $f \in H_+$. Then

$$\begin{pmatrix} p(z) \\ q(z) \end{pmatrix} = \begin{pmatrix} f_e(z) - \sum_{n=0}^{\infty} z^n \sum_{k=0}^{\infty} f_{2n+2+k} (e_k, A^* e_{-2}) \\ f_o(z) - \sum_{n=0}^{\infty} z^n \sum_{k=0}^{\infty} f_{2n+2+k} (e_k, A^* e_{-1}) \end{pmatrix}, \quad (13)$$

holds and they have bounds

$$\begin{cases} |p(z)| \leq |f_e(z)| + \sum_{n=0}^{\infty} |z|^n \sqrt{\sum_{k \geq 2n+2} |f_k|^2} \|A^* e_{-2}\|, \\ |q(z)| \leq |f_o(z)| + \sum_{n=0}^{\infty} |z|^n \sqrt{\sum_{k \geq 2n+2} |f_k|^2} \|A^* e_{-1}\|. \end{cases} \quad (14)$$

Proof. For $f \in H_+$ we easily see

$$\begin{aligned} & \sum_{n=0}^{\infty} f_{2n} z^n s_n(z) + \sum_{n=0}^{\infty} f_{2n+1} z^n t_n(z) \\ &= \sum_{n=0}^{\infty} f_{2n} z^n \left\{ 1 - \sum_{k=0}^{n-1} z^{-(k+1)} (S^k \varphi, e_{-2}) \right\} - \sum_{n=0}^{\infty} f_{2n+1} z^n \sum_{k=0}^{n-1} z^{-(k+1)} (S^k \psi, e_{-2}) \\ &= f_e(z) - \sum_{n=0}^{\infty} z^n \sum_{k=0}^{\infty} f_{2n+2+k} (e_k, A^* e_{-2}). \end{aligned} \quad (15)$$

Similarly we have

$$\begin{aligned} & \sum_{n=0}^{\infty} f_{2n} z^n u_n(z) + \sum_{n=0}^{\infty} f_{2n+1} z^n v_n(z) \\ &= \sum_{n=0}^{\infty} f_{2n+1} z^n \left\{ 1 - \sum_{k=0}^{n-1} z^{-(k+1)} (S^k \psi, e_{-1}) \right\} - \sum_{n=0}^{\infty} f_{2n} z^n \sum_{k=0}^{n-1} z^{-(k+1)} (S^k \varphi, e_{-1}) \\ &= f_o(z) - \sum_{n=0}^{\infty} z^n \sum_{k=0}^{\infty} f_{2n+2+k} (e_k, A^* e_{-1}), \end{aligned}$$

which yields (13). From (13) estimates

$$\begin{aligned} \left| \sum_{n=0}^{\infty} f_{2n} z^n s_n(z) + \sum_{n=0}^{\infty} f_{2n+1} z^n t_n(z) \right| &\leq |f_e(z)| + \left| \sum_{n=0}^{\infty} z^n \sum_{k=0}^{\infty} f_{2n+2+k} (e_k, A^* e_{-2}) \right| \\ &\leq |f_e(z)| + \sum_{n=0}^{\infty} |z|^n \sqrt{\sum_{k \geq 2n+2} |f_k|^2} \|A^* e_{-2}\| \end{aligned}$$

and

$$\left| \sum_{n=0}^{\infty} f_{2n} z^n u_n(z) + \sum_{n=0}^{\infty} f_{2n+1} z^n v_n(z) \right| \leq |f_o(z)| + \sum_{n=0}^{\infty} |z|^n \sqrt{\sum_{k \geq 2n+2} |f_k|^2} \|A^* e_{-1}\|$$

follow. ■

Proposition 1 For $\chi \in W$ let

$$\chi = f + Af \in H_+ \oplus H_-$$

and assume f satisfies

$$\sum_{n=1}^{\infty} n |f_n|^2 < \infty. \quad (16)$$

Then p, q defined by (12) become elements of H_+ , and they have bound

$$\begin{cases} \|p\| \leq \|f_e\| + \|A^* e_{-2}\| \sum_{n=1}^{\infty} n |f_n|^2, \\ \|q\| \leq \|f_o\| + \|A^* e_{-1}\| \sum_{n=1}^{\infty} n |f_n|^2. \end{cases}$$

They coincide with

$$\Delta(z)^{-1} \begin{pmatrix} \chi_e(z) \\ \chi_o(z) \end{pmatrix} \quad \text{a.e. for } |z| = 1. \quad (17)$$

Proof. The estimates (14) show

$$\begin{aligned} \|p\| &\leq \|f_e\| + \|A^* e_{-2}\| \sqrt{\sum_{n=0}^{\infty} \sum_{k \geq 2n+2} |f_k|^2} \\ &\leq \|f_e\| + \|A^* e_{-2}\| \sum_{n=1}^{\infty} n |f_n|^2 < \infty. \end{aligned}$$

Similarly we have

$$\|q\| \leq \|f_o\| + \|A^* e_{-1}\| \sum_{n=1}^{\infty} n |f_n|^2 < \infty.$$

Since from (8),(9) it follows that for $\chi \in W$

$$\chi(z) = (1 + \varphi(z))p(z^2) + (z + \psi(z))q(z^2), \quad \text{a.e. } |z| = 1,$$

we see the identity (17). ■

Remark 2 It might be true that p, q defined by (13) have always finite limits on $|z| = 1$ a.e. for any $f \in H_+$, and any $\chi \in W$ has an expression

$$\chi(z) = p(z^2)(1 + \varphi(z)) + q(z^2)(z + \psi(z)).$$

Actually, if the spectrum of S is contained in open unit disk $|z| < 1$, then

$$\|S^n\| \leq r^n$$

for some $r < 1$ and for all sufficiently large n . Noting $Ae_{2k} = S^k\varphi$, $Ae_{2k+1} = S^k\psi$, we see the expression (13) implies that p, q have finite limits a.e. on $|z| = 1$ for any $f \in H_+$.

3 Group action on $Gr^{(2)}(H)$

3.1 τ -function

Let Γ be

$$\Gamma = \left\{ g(z) = e^{h(z)}; h \text{ is holomorphic on } |z| < 1 \text{ and } h'(e^{i\theta}) \in L^2 \right\}.$$

Define a distance on Γ by

$$d(g_1, g_2) = \sqrt{\int_0^{2\pi} |h'_1(e^{i\theta}) - h'_2(e^{i\theta})|^2 d\theta}.$$

Γ acts on $Gr^{(2)}(H)$ by multiplication and becomes a commutative group. Therefore, for $W \in Gr^{(2)}(H)$, $g^{-1}W$ is a closed subspace of H . This space certainly satisfies the properties(i),(ii). However, the property(iii) does not hold automatically. For $g \in \Gamma$, an operator

$$R_g (= R_g^W) = gP_{H_+}g^{-1}A_W$$

defines a bounded linear operator from H_+ to H_+ . Since A_W is of Hilbert-Schmidt type, so is R_g .

Lemma 5 For $g \in \Gamma$, an operator $gP_{H_+}g^{-1}$ from H_- to H_+ is of Hilbert-Schmidt type, and for $g_1, g_2 \in \Gamma$ it holds that

$$\begin{aligned} & \|g_1P_{H_+}g_1^{-1} - g_2P_{H_+}g_2^{-1}\|_{H.S.}^2 \\ &= \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \left| \frac{g_1(e^{i\theta})/g_1(e^{i\varphi}) - g_2(e^{i\theta})/g_2(e^{i\varphi})}{e^{i\theta} - e^{i\varphi}} \right|^2 d\theta d\varphi. \end{aligned}$$

Proof. For $|z| < 1$

$$\begin{aligned} & (g_1P_{H_+}g_1^{-1} - g_2P_{H_+}g_2^{-1})e_{-n}(z) \\ &= \begin{cases} \frac{1}{2\pi i} \int_{|\zeta|=1} \frac{g_1(z)/g_1(\zeta) - g_2(z)/g_2(\zeta)}{\zeta^n(\zeta - z)} d\zeta, & \text{if } n \geq 1, \\ 0, & \text{if } n \leq 0 \end{cases} \end{aligned}$$

is valid, whence

$$\begin{aligned} & \|g_1 P_{H_+} g_1^{-1} - g_2 P_{H_+} g_2^{-1}\|_{H.S.}^2 \\ &= \sum_{n=1}^{\infty} \|(g_1 P_{H_+} g_1^{-1} - g_2 P_{H_+} g_2^{-1}) e_{-n}\|^2 \\ &= \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{2\pi} \left| \frac{g_1(e^{i\theta})/g_1(e^{i\varphi}) - g_2(e^{i\theta})/g_2(e^{i\varphi})}{e^{i\theta} - e^{i\varphi}} \right|^2 d\theta d\varphi, \end{aligned}$$

holds, which implies the assertion of the lemma. ■

This lemma shows that R_g is of trace class.

Lemma 6 *A necessary and sufficient condition for R_g to satisfy $\ker(I + R_g) = \{0\}$ is $g^{-1}W \in Gr^{(2)}(H)$. In this case, the A -operator corresponding to $g^{-1}W$ is given by*

$$A_{g^{-1}W} = P_{H_-} g^{-1} A_W (I + R_g)^{-1} g.$$

Proof. Since R_g is a compact operator, $\ker(I + R_g) = \{0\}$ implies the existence of $(I + R_g)^{-1}$ as a bounded operator on H_+ . Therefore, the operator

$$A_g = P_{H_-} g^{-1} A (I + R_g)^{-1} g$$

is of trace class from H_+ to H_- . For $f \in H_+$, an identity

$$\begin{aligned} g(f + A_g f) &= gf + g P_{H_-} g^{-1} A (I + R_g)^{-1} gf \\ &= gf + A (I + R_g)^{-1} gf - R_g (I + R_g)^{-1} gf \\ &= (I + R_g)^{-1} gf + A (I + R_g)^{-1} gf \in W \end{aligned}$$

shows

$$g^{-1}W = \{f + A_g f; f \in H_+\}.$$

Hence we know that $g^{-1}W$ satisfies the property(iii), and $g^{-1}W \in Gr^{(2)}(H)$ holds. Therefore, the A -operator for $g^{-1}W$ becomes A_g . Conversely, if $f \in \ker(I + R_g)$, then, from an identity

$$g^{-1}f + P_{H_+} (g^{-1}Af) = 0$$

it follows that

$$P_{H_-} (g^{-1}Af) = g^{-1}(f + Af) - (g^{-1}f + P_{H_+} (g^{-1}Af)) = g^{-1}(f + Af) \in g^{-1}W.$$

Therefore $g^{-1}W \in Gr^{(2)}(H)$ implies $P_{H_-} (g^{-1}Af) = 0$ and

$$g^{-1}(f + Af) = 0 \implies f + Af = 0 \implies f = 0,$$

which completes the proof. ■

Since R_g is of trace class, τ -function is defined by

$$\tau_W(g) = \det(I + R_g),$$

which was introduced by Sato to describe solutions of completely integrable systems. τ -functions satisfy the following cocycle property.

Lemma 7 τ -functions satisfy

(i) The property $\tau_W(g) \neq 0$ is equivalent to $g^{-1}W \in Gr^{(2)}(H)$.

(ii) For $g_1, g_2 \in \Gamma$, if $g_1^{-1}W \in Gr^{(2)}(H)$, then

$$\tau_W(g_1g_2) = \tau_W(g_1)\tau_{g_1^{-1}W}(g_2) \quad (\text{cocycle property}). \quad (18)$$

Proof. Since R_g is of trace class, $\tau_W(g) \neq 0$ is equivalent to $Ker(I + R_g) = \{0\}$, which shows (i). (ii) follows from

$$\begin{aligned} (g_1g_2)^{-1}A &= g_2^{-1}P_{H_+}g_1^{-1}A + g_2^{-1}P_{H_-}g_1^{-1}A \\ &= g_2^{-1}g_1^{-1}R_{g_1} + g_2^{-1}A_{g_1}g_1^{-1}(I + R_{g_1}) \quad (A_g = A_{g^{-1}W}), \end{aligned}$$

and

$$P_{H_+} \left((g_1g_2)^{-1}A \right) = g_2^{-1}g_1^{-1}R_{g_1} + P_{H_+} \left(g_2^{-1}A_{g_1}g_1^{-1}(I + R_{g_1}) \right),$$

hence

$$\begin{aligned} I + R_{g_1g_2} &= I + g_1g_2P_{H_+} \left((g_1g_2)^{-1}A \right) \\ &= I + R_{g_1} + g_1g_2P_{H_+} \left(g_2^{-1}A_{g_1}g_1^{-1}(I + R_{g_1}) \right) \\ &= I + R_{g_1} + g_1R_{g_2}^{g_1^{-1}W}g_1^{-1}(I + R_{g_1}) \\ &= g_1 \left(I + R_{g_2}^{g_1^{-1}W} \right) g_1^{-1}(I + R_{g_1}). \end{aligned}$$

Consequently, if $g_1^{-1}W \in Gr^{(2)}(H)$, then

$$\begin{aligned} \tau_W(g_1g_2) &= \det(I + R_{g_1g_2}) \\ &= \det \left(I + R_{g_2}^{g_1^{-1}W} \right) \det(I + R_{g_1}) \\ &= \tau_W(g_1)\tau_{g_1^{-1}W}(g_2) \end{aligned}$$

is valid. ■

We compute the τ -function for a special g , which is well-known..

Lemma 8 For $|\zeta| > 1$ set

$$q_\zeta(z) = 1 - \frac{z}{\zeta} \in \Gamma.$$

Then, it holds that

$$\tau_W(q_\zeta) = 1 + \varphi_W(\zeta).$$

Proof. For $f \in H_-$ and $1 \leq |z| < |\zeta|$, we have a decomposition into $H_- \oplus H_+$

$$q_\zeta^{-1}f(z) = \left(1 - \frac{z}{\zeta}\right)^{-1} f(z) = \left(1 - \frac{z}{\zeta}\right)^{-1} (f(z) - f(\zeta)) + \left(1 - \frac{z}{\zeta}\right)^{-1} f(\zeta),$$

which implies

$$\left(q_\zeta P_{H_+} q_\zeta^{-1} f \right) (z) = \left(1 - \frac{z}{\zeta}\right) \left(1 - \frac{z}{\zeta}\right)^{-1} f(\zeta) = f(\zeta),$$

hence, if $W \in Gr^{(2)}(H)$, then for $f \in H_+$

$$\left(q_\zeta P_{H_+} q_\zeta^{-1} A_W f\right)(z) = A_W f(\zeta)$$

holds, which implies that $q_\zeta P_{H_+} q_\zeta^{-1} A_W$ is a linear operator of rank 1. Thus

$$\tau_W(q_\zeta) = \det\left(I + q_\zeta P_{H_+} q_\zeta^{-1} A_W\right) = 1 + (A_W 1)(\zeta) = 1 + \varphi_W(\zeta)$$

follows. ■

3.2 Monodromy matrix

In this section we investigate a kind of monodromy matrix under the group action Γ .

Lemma 9 *Suppose $u \in H_+$ satisfies*

$$\sum_{n=1}^{\infty} n^2 |u_n|^2 < \infty.$$

Then for any $f \in H_-$, $v = P_{H_+}(uf)$ satisfies

$$\sum_{n=1}^{\infty} n |v_n|^2 \leq \frac{1}{2} \|f\|^2 \sum_{n=1}^{\infty} n^2 |u_n|^2 < \infty.$$

Moreover, if u is holomorphic on $|z| < R$ for some $R > 1$, then v is also holomorphic on $|z| < R$.

Proof. Note

$$\begin{aligned} u(z) f(z) &= \sum_{n=1}^{\infty} u_n \sum_{k=1}^n z^{n-k} f_k + \sum_{n=0}^{\infty} u_n \sum_{k=n+1}^{\infty} z^{n-k} f_k \\ &= \sum_{n=0}^{\infty} z^n \sum_{k=1}^{\infty} u_{n+k} f_k + \sum_{n=1}^{\infty} z^{-n} \sum_{k=0}^{\infty} u_k f_{n+k}. \end{aligned}$$

This gives the decomposition of uf into H_+ and H_- respectively and we have

$$\sum_{n=1}^{\infty} n \left| \sum_{k=1}^{\infty} u_{n+k} f_k \right|^2 \leq \frac{1}{2} \|f\|^2 \sum_{n=1}^{\infty} n^2 |u_n|^2.$$

If u is holomorphic on $|z| < R$, then

$$|u_n| \leq Cr^{-n}$$

with $r < R$. Therefore

$$\left| \sum_{k=1}^{\infty} u_{n+k} f_k \right| \leq \|f\| \sqrt{\sum_{k=1}^{\infty} |u_{n+k}|^2} \leq \frac{Cr^{-1}}{\sqrt{1-r^{-2}}} \|f\| r^{-n}$$

holds, and we see that v is holomorphic on $|z| < R$, since r can be taken arbitrarily close to R . ■

Set an invertible matrix

$$\tilde{\Delta}_W(z) = \begin{pmatrix} 1 + \varphi_e(z) & \psi_e(z) + a_1(1 + \varphi_e(z)) \\ \varphi_o(z) & 1 + \psi_o(z) + a_1\varphi_o(z) \end{pmatrix} = \begin{pmatrix} a(z) & b(z) + a_1a(z) \\ c(z) & d(z) + a_1c(z) \end{pmatrix}$$

and, for $g(z) = e^{h(z)} \in \Gamma$ define

$$U_W(z, g) = {}^t\tilde{\Delta}_{g^{-1}W}(z) \begin{pmatrix} g_e(z) & g_o(z) \\ zg_o(z) & g_e(z) \end{pmatrix} {}^t\tilde{\Delta}_W(z)^{-1} \quad (19)$$

on $|z| = 1$. Then the identity(19) is equivalent to an identity

$$\begin{aligned} & g(z) \begin{pmatrix} 1 + \varphi_{g^{-1}W}(z) \\ z + \psi_{g^{-1}W}(z) + a_{g^{-1}W}(1 + \varphi_{g^{-1}W}(z)) \end{pmatrix} \\ &= U_W(z^2, g) \begin{pmatrix} 1 + \varphi_W(z) \\ z + \psi_W(z) + a_W(1 + \varphi_W(z)) \end{pmatrix}, \end{aligned} \quad (20)$$

where

$$\varphi_W(z) = \frac{a_W}{z} + \frac{*}{z^2} + \cdots.$$

This $U_W(z, g)$ has analytic continuation to the inside of the unit disk, namely

Theorem 2 *Assume $g^{-1}W \in Gr^{(2)}(H)$. Then the entries of $U_W(z, g)$ are elements of H_+ , and the determinant is given by*

$$\det U_W(z, g) = g(\sqrt{z})g(-\sqrt{z}) = e^{2h_e(z)}. \quad (21)$$

Moreover $U_W(z, g)$ has a cocycle property

$$U_W(z, g_1g_2) = U_{g_1^{-1}W}(z, g_2)U_W(z, g_1), \quad (22)$$

and, if g is holomorphic on $|z| < R$ for some $R > 1$, then $U_W(z, g)$ is also holomorphic on $|z| < R$.

Proof. Since $dg(e^{i\theta})/d\theta \in L^2$, the condition of Lemma9 is satisfied. Therefore, the H_+ -component of $g(z)(1 + \varphi_{g^{-1}W}(z))$ holds the condition (16), which implies that there exist elements $\{a, b\}$ of H_+ such that

$$\begin{aligned} & W \ni g(z)(1 + \varphi_{g^{-1}W}(z)) \\ &= a(z^2)(1 + \varphi_W(z)) + b(z^2)(z + \psi_W(z) + a_W(1 + \varphi_W(z))). \end{aligned}$$

Similarly we have elements $\{c, d\}$ of H_+ such that

$$\begin{aligned} & W \ni g(z)(z + \psi_{g^{-1}W}(z) + a_{g^{-1}W}(1 + \varphi_{g^{-1}W}(z))) \\ &= c(z^2)(1 + \varphi_W(z)) + d(z^2)(z + \psi_W(z) + a_W(1 + \varphi_W(z))). \end{aligned}$$

Therefore, to obtain U_W we have only to set

$$U_W(z, g) = \begin{pmatrix} a(z) & b(z) \\ c(z) & d(z) \end{pmatrix}.$$

on $|z| \leq 1$. To show the identity(22) notice an identity

$$\begin{aligned} & \left(U_{g_1^{-1}W}(z^2, g_2)U_W(z^2, g_1) - U_W(z^2, g_1g_2) \right) \begin{pmatrix} 1 + \varphi_W(z) \\ z + \psi_W(z) + a_W(1 + \varphi_W(z)) \end{pmatrix} \\ & = 0. \end{aligned}$$

Hence, replacing z by $-z$ yields

$$\left(U_{g_1^{-1}W}(z, g_2)U_W(z, g_1) - U_W(z, g_1g_2) \right) \tilde{\Delta}_W(z) = 0.$$

Since $\tilde{\Delta}_W(z)$ is non-singular for a.e. $|z| = 1$, we have

$$U_{g_1^{-1}W}(z, g_2)U_W(z, g_1) = U_W(z, g_1g_2) \quad \text{for a.e. } |z| = 1,$$

which gives (22). To see the determinant of $U_W(z, g)$, we use (19) and obtain

$$\begin{aligned} \det U_W(z, g) &= \det \tilde{\Delta}_{g^{-1}W}(z) \det \begin{pmatrix} g_e(z) & g_o(z) \\ zg_o(z) & g_e(z) \end{pmatrix} \det \tilde{\Delta}_W(z)^{-1} \\ &= \delta_{g^{-1}W}(z) \delta_W(z)^{-1} g(\sqrt{z})g(-\sqrt{z}). \end{aligned} \quad (23)$$

For

$$g_n(z) = \exp \left(\sum_{k=0}^n h_k z^k \right)$$

we know that

$$\Theta_n(z) = g_n(\sqrt{z})^{-1} g_n(-\sqrt{z})^{-1} \det U_W(z, g_n)$$

is entire, and takes value 1 at $z = \infty$ due to the identity(23). Therefore, the entire function Θ_n should be constantly 1. This leads us to the identity(21) by letting $n \rightarrow \infty$. The analyticity of U_W beyond the boundary $|z| = 1$ under the condition on g follows from Lemma9, which completes the proof. ■

We call the matrix $U_W(z, g)$ **monodromy matrix** for $W \in Gr^{(2)}(H)$.

Corollary 2 For any $g \in \Gamma$, $\delta_{g^{-1}W}(z) = \delta_W(z)$ is valid.

Proof. The corollary follows immediately from Theorem2. ■

3.3 Baker-Akhiezer function and Weyl function

Let q be a smooth function defined in a real neighborhood U of the origin. For a positive real number R , a function $f(x, z)$ of two variables $x \in U$ and $z \in \{z \in \mathbb{C}; |z| > R\}$ is called a **Baker-Akhiezer function** for the potential q if f satisfies as a function of x

$$-f''(x, z) + q(x)f(x, z) = -z^2 f(x, z), \quad \left(' = \frac{d}{dx} \right), \quad (24)$$

and as a function of z the expansion

$$e^{xz} f(x, z) = 1 + \frac{a_1(x)}{z} + \frac{a_2(x)}{z^2} + \dots \quad (25)$$

converges on $R < |z| \leq \infty$.

Lemma 10 *Let $f(x, z)$ be a Baker-Akhiezer function for q . Then, for every $k \geq 1$, a_k satisfies identities*

$$2a'_1(x) = -q(x), \quad 2a'_{k+1}(x) - a''_k(x) = -q(x)a_k(x). \quad (26)$$

Moreover, $f(x, z)/f(0, z)$ is also a Baker-Akhiezer function for q and q determines uniquely $f(x, z)/f(0, z)$.

Proof. Taking derivative of (25) yields

$$\begin{aligned} & e^{xz} (-f''(x, z) + z^2 f(x, z)) \\ &= 2a'_1(x) + \frac{2a'_2(x) - a''_1(x)}{z} + \frac{2a'_3(x) - a''_2(x)}{z^2} + \dots \end{aligned}$$

This combined with (24) implies, for $k \geq 1$

$$2a'_1(x) = -q(x), \quad 2a'_{k+1}(x) - a''_k(x) = -q(x)a_k(x).$$

It is easy to see that $\tilde{f}(x, z) = f(x, z)/f(0, z)$ is also a Baker-Akhiezer function for q . Since $\tilde{f}(0, z) = 1$ holds, then, for every $k \geq 1$, we have $a_k(0) = 0$. Hence from (26) we see that $\{a_k(x)\}_{k \geq 1}$ can be determined from q . ■

Let $W \in Gr^{(2)}(H)$. As we have seen in Introduction, $f_W(x, \cdot)$ defined as a unique element of W such that $P_{H^+}(e_x^{-1}f) = 1$ satisfies a Schrödinger equation with potential $-2a'_1(x)$. This function clearly fulfils the condition(25), hence $f_W(x, z)$ is a Baker-Akhiezer function for the potential $-2a'_1(x)$. Since $f_W(x, \cdot)$ can be expressed by A -operator for $e_x^{-1}W$ as

$$f_W(x, \zeta) = e^{-x\zeta} \left\{ 1 + \left(A_{e_x^{-1}W} 1 \right) (\zeta) \right\},$$

replacing W in Lemma8 by $e_x^{-1}W$, we see

$$e^{x\zeta} f_W(x, \zeta) = 1 + \left(A_{e_x^{-1}W} 1 \right) (\zeta) = \tau_{e_x^{-1}W}(q\zeta).$$

On the other hand

$$a_1(x) = \lim_{\zeta \rightarrow \infty} \zeta (e^{x\zeta} f_W(x, \zeta) - 1) = \lim_{\zeta \rightarrow \infty} \zeta \left(\tau_{e_x^{-1}W}(q\zeta) - 1 \right)$$

is valid, hence, from (ii) of Lemma7

$$\tau_{e_x^{-1}W}(q\zeta) - 1 = \frac{\tau_W(e_x q\zeta)}{\tau_W(e_x)} - 1 = \frac{\tau_W(e_x q\zeta) - \tau_W(e_x)}{\tau_W(e_x)}$$

follows. Thus, identifying $q\zeta(z) \cong e^{-z/\zeta}$ for large ζ , it holds that

$$\lim_{\zeta \rightarrow \infty} \zeta (\tau_W(e_x q\zeta) - \tau_W(e_x)) = \lim_{\zeta \rightarrow \infty} \zeta \left(\tau_W\left(e_{x+\frac{1}{\zeta}}\right) - \tau_W(e_x) \right) = \frac{d}{dx} \tau_W(e_x).$$

Consequently we obtain the theorem below. For $W \in Gr^{(2)}(H)$ define

$$q_W(x) = -2 \frac{d^2}{dx^2} \log \tau_W(e_x).$$

Theorem 3 *There exists a Baker-Akhiezer function f_W for q_W . In this case f_W and its first coefficient are given by*

$$f_W(x, \zeta) = e^{-x\zeta} \tau_{e_x^{-1}W}(q_\zeta), \quad a_1(x) = \frac{d}{dx} \log \tau_W(e_x).$$

q_W is defined by $W \in Gr^{(2)}(H)$. However, q_W does not determine W , hence, the next problem is to find a quantity from which q_W is determined uniquely.

Let $f_W(x, z)$ be the Baker-Akhiezer function for $W \in Gr^{(2)}(H)$ and define

$$M_W(z) = \frac{f'_W(0, z)}{f_W(0, z)},$$

which we call **Weyl function** of W . Since f_W satisfies the identity(25), for a $k \in H_-$ we have

$$\begin{aligned} f'_W(0, z) &= -z \left(1 + \frac{a_1(0)}{z} + \frac{a_2(0)}{z^2} + \dots \right) + \left(\frac{a'_1(0)}{z} + \frac{a'_2(0)}{z^2} + \dots \right) \\ &= -z - a_1(0) + k(z). \end{aligned}$$

Since $f'_W(0, \cdot) \in W$, setting $a_W = a_1(0)$, we see

$$f'_W(0, z) = -(z + \psi_W(z)) - a_W(1 + \varphi_W(z)). \quad (27)$$

The identity $f_W(0, z) = 1 + \varphi_W(z)$ shows

$$M_W(z) = -\frac{z + \psi_W(z) + a_W(1 + \varphi_W(z))}{1 + \varphi_W(z)} = -\frac{z + \psi_W(z)}{1 + \varphi_W(z)} - a_W. \quad (28)$$

We have

Theorem 4 *For $W_1, W_2 \in Gr^{(2)}(H)$ an identity $q_{W_1} = q_{W_2}$ holds if and only if $M_{W_1} = M_{W_2}$ holds. Moreover, in this case, if $g \in \Gamma$ satisfy $g^{-1}W_1, g^{-1}W_2 \in Gr^{(2)}(H)$, then*

$$U_{W_1}(z, g) = U_{W_2}(z, g), \quad M_{g^{-1}W_1} = M_{g^{-1}W_2},$$

and hence

$$q_{g^{-1}W_1} = q_{g^{-1}W_2}$$

is valid.

Proof. Suppose $q_{W_1} = q_{W_2}$. Then Lemma10 implies that Baker-Akhiezer functions $f_1(x, z), f_2(x, z)$ of W_1, W_2 satisfy

$$\frac{f_1(x, z)}{f_1(0, z)} = \frac{f_2(x, z)}{f_2(0, z)}. \quad (29)$$

Therefore

$$\frac{f'_1(0, z)}{f_1(0, z)} = \frac{f'_2(0, z)}{f_2(0, z)} \quad (30)$$

holds and $M_{W_1} = M_{W_2}$ follows. Conversely assume $M_{W_1} = M_{W_2}$. For $W \in Gr^{(2)}(H)$ set

$$\tilde{f}_W(x, z) = \frac{f_W(x, z)}{f_W(0, z)}.$$

Then \tilde{f}_W satisfies the condition in Lemma10 by setting $q = q_W$, and the coefficients $\{\tilde{a}_k(x)\}_{k \geq 1}$ arising from the expansion

$$\tilde{f}_W(x, z) = e^{-xz} \left(1 + \frac{\tilde{a}_1(x)}{z} + \frac{\tilde{a}_2(x)}{z^2} + \dots \right)$$

fulfill identities

$$2\tilde{a}'_{k+1}(x) - \tilde{a}''_k(x) = 2\tilde{a}'_1(x)\tilde{a}_k(x) \quad (31)$$

for every $k \geq 1$. Since $\tilde{f}'_W(0, z) = f'_W(0, z)/f_W(0, z) = M_W(z)$ is valid, noting $\tilde{f}_W(0, z) = 1$, we see that $\{\tilde{a}'_k(0)\}_{k \geq 1}$ can be determined from $M_W(z)$ uniquely. Noting here $\tilde{a}_k(0) = 0$ for every $k \geq 1$, we see from (31) that $\{\tilde{a}''_k(0)\}_{k \geq 1}$ can be determined from $M_W(z)$ uniquely as well. In the identity(31), taking derivatives, we can know all $\{\tilde{a}^{(n)}_k(0)\}_{n \geq 0, k \geq 1}$. Since $\{\tilde{a}_k(x)\}_{k \geq 1}$ are holomorphic in a neighborhood of 0, we see that $\{\tilde{a}_k(x)\}_{k \geq 1}$ can be determined from $M_W(z)$, which implies that $q_W(x) (= -2\tilde{a}'_1(x))$ is determined from $M_W(z)$. Therefore $q_{W_1} = q_{W_2}$ follows from $M_{W_1} = M_{W_2}$.

In order to show the last statement we assume $q_{W_1} = q_{W_2}$. From the identity (30) we have

$$f_2(0, z) = r(z)f_1(0, z) \quad \text{and} \quad f'_2(0, z) = r(z)f'_1(0, z) \quad \text{with} \quad r(z) = \frac{f_2(0, z)}{f_1(0, z)},$$

which implies

$$\left(\begin{array}{c} 1 + \varphi_{W_2}(z) \\ z + \psi_{W_2}(z) + a_{W_2}(1 + \varphi_{W_2}(z)) \end{array} \right) = r(z) \left(\begin{array}{c} 1 + \varphi_{W_1}(z) \\ z + \psi_{W_1}(z) + a_{W_1}(1 + \varphi_{W_1}(z)) \end{array} \right). \quad (32)$$

On the other hand, we have

$$\begin{aligned} & r(z)g(z) \left(\begin{array}{c} 1 + \varphi_{g^{-1}W_1}(z) \\ (z + \psi_{g^{-1}W_1}(z)) + a_{g^{-1}W_1}(1 + \varphi_{g^{-1}W_1}(z)) \end{array} \right) \\ &= U_{W_1}(z^2, g) \left(\begin{array}{c} r(z)(1 + \varphi_{W_1}(z)) \\ r(z)(z + \psi_{W_1}(z)) + a_{W_1}r(z)(1 + \varphi_{W_1}(z)) \end{array} \right) \\ &= U_{W_1}(z^2, g) \left(\begin{array}{c} 1 + \varphi_{W_2}(z) \\ (z + \psi_{W_2}(z)) + a_{W_2}(1 + \varphi_{W_2}(z)) \end{array} \right). \end{aligned}$$

Assume g is holomorphic on $\{|z| < R\}$. Then so is $U_{W_1}(z^2, g)$, hence the entries of the RHS are elements of W_2 , which implies

$$r(z)(1 + \varphi_{g^{-1}W_1}(z)) \in g^{-1}W_2.$$

The property $r(\infty) = 1$ shows

$$r(z)(1 + \varphi_{g^{-1}W_1}(z)) = 1 + \varphi_{g^{-1}W_2}(z).$$

Similarly we have

$$r(z)(z + \psi_{g^{-1}W_1}(z)) = z + \psi_{g^{-1}W_2}(z) + r_1(1 + \varphi_{g^{-1}W_2}(z)),$$

hence

$$\begin{aligned} & r(z) \left(z + \psi_{g^{-1}W_1}(z) \right) + a_{g^{-1}W_1} r(z) \left(1 + \varphi_{g^{-1}W_1}(z) \right) \\ &= z + \psi_{g^{-1}W_2}(z) + \left(r_1 + a_{g^{-1}W_1} \right) \left(1 + \varphi_{g^{-1}W_2}(z) \right) \\ &= z + \psi_{g^{-1}W_2}(z) + a_{g^{-1}W_2} \left(1 + \varphi_{g^{-1}W_2}(z) \right). \end{aligned}$$

Thus

$$\begin{aligned} & U_{W_1}(z^2, g) \left(\begin{array}{c} 1 + \varphi_{W_2}(z) \\ (z + \psi_{W_2}(z)) + a_{W_2} (1 + \varphi_{W_2}(z)) \end{array} \right) \\ &= g(z) \left(\begin{array}{c} 1 + \varphi_{g^{-1}W_2}(z) \\ (z + \psi_{g^{-1}W_2}(z)) + a_{g^{-1}W_2} (1 + \varphi_{g^{-1}W_2}(z)) \end{array} \right) \\ &= U_{W_2}(z^2, g) \left(\begin{array}{c} 1 + \varphi_{W_2}(z) \\ z + \psi_{W_2}(z) + a_{W_2} (1 + \varphi_{W_2}(z)) \end{array} \right), \end{aligned}$$

which concludes $U_{W_1}(z, g) = U_{W_2}(z, g)$. For a general g we have only to approximate it by its Taylor expansion. Here it should be noted that $q_{g^{-1}W}$ as a function of g is continuous. ■

Remark 3 If q is

$$q(x) = \frac{m(m+1)}{(x-x_0)^2},$$

then its Baker-Akhiezer function is of a form

$$f(x, z) = e^{-x_0 z} h_m((x-x_0)z),$$

where h_m is the Hankel function of half integer defined by

$$h_m(z) = (-1)^m z^{m+1} \left(\frac{d}{zdz} \right)^m \frac{e^{-z}}{z} = e^{-z} \sum_{k=0}^m \alpha_k z^{-k}.$$

The coefficients α_k are defined by $\alpha_0 = 1$ for $k = 0$ and for $k \geq 1$

$$\alpha_k = \frac{1}{2^{3k} k!} \prod_{l=0}^{k-1} \left((2m+1)^2 - (2l+1)^2 \right).$$

In this case, setting

$$\varphi(z) = f(0, z) - 1 = \sum_{k=1}^m \alpha_k (-x_0 z)^{-k},$$

$$\psi(z) = -f'(0, z) - z + \frac{\alpha_1}{x_0} f(0, z) = -\frac{1}{x_0} \sum_{k=1}^m (\alpha_{k+1} + k\alpha_k - \alpha_1 \alpha_k) (-x_0 z)^{-k}$$

$W \in Gr^{(2)}(H)$ can be constructed with its characteristic functions $\{\varphi, \psi\}$. This W is an element of $Gr^{(2)}(H)$ which was considered in Section 4.2, namely for this W , $P_H W$ is finite dimensional, and the τ -function can be computed by determinant. Moreover, in this case, the M -function becomes

$$M(z) = \frac{f'(0, z)}{f(0, z)} = \frac{ze^{-x_0 z} h'_m(-x_0 z)}{e^{-x_0 z} h_m(-x_0 z)} = -z - \frac{\sum_{k=1}^m k \alpha_k (-x_0 z)^{-k}}{\sum_{k=0}^m \alpha_k (-x_0 z)^{-k}}.$$

On the other hand, if x_0 is a real number, $f_{\pm}^q(x, -z^2)$ can be determined as a unique solution of

$$\begin{aligned} L^q f &= -z^2 f, \quad f(0) = 1, \quad \int_0^{x_0} |f(x)|^2 dx < \infty, \\ L^q f &= -z^2 f, \quad f(0) = 1, \quad \int_{-\infty}^0 |f(x)|^2 dx < \infty. \end{aligned}$$

Set

$$\tilde{h}_m(z) = z^{m+1} \left(\frac{d}{zdz} \right)^m \frac{\text{sh}z}{z} = \frac{1}{2} h_m(-z) + \frac{1}{2} h_m(z).$$

Then

$$f_+^q(x, -z^2) = \frac{\tilde{h}_m((x-x_0)z)}{\tilde{h}_m(-x_0z)}, \quad f_-^q(x, -z^2) = \frac{f(x, z)}{f(0, z)}$$

are valid, hence the Weyl functions for L^q are

$$m_+^q(-z^2) = f_+^{q'}(0, z) = \frac{z \tilde{h}'_m(-x_0z)}{\tilde{h}_m(-x_0z)}, \quad m_-^q(-z^2) = -\frac{f'(0, z)}{f(0, z)} = -M(z).$$

Namely if q is real valued, the relationship between the Weyl function and the M -function is not trivial.

Lemma 11 Suppose $W_1, W_2 \in Gr^{(2)}(H)$ satisfy $M_{W_1} = M_{W_2}$. Then

$$r(z) = \frac{1 + \varphi_{W_2}(z)}{1 + \varphi_{W_1}(z)}$$

becomes holomorphic on $1 < |z| \leq \infty$ and satisfies

$$r(z) \neq 0 \quad \text{for any } z \text{ there and for a.e. } z \text{ such that } |z| = 1.$$

Proof. From (32)

$$\begin{cases} \tilde{\Delta}_{W_2}(z) = \begin{pmatrix} r_e(z) & z r_o(z) \\ r_o(z) & r_e(z) \end{pmatrix} \tilde{\Delta}_{W_1}(z), \\ \delta_{W_2}(z) = r(\sqrt{z}) r(-\sqrt{z}) \delta_{W_1}(z). \end{cases}$$

follows. The second identity comes from

$$\delta_W(z) = \det \Delta_W(z) = \det \tilde{\Delta}_W(z).$$

Then (i) of Theorem1 shows the non-vanishing property of r . ■

4 KdV flow

In this section we construct KdV flow on a certain class of potentials Ω . For $W \in Gr^{(2)}(H)$ set

$$q(x) = -2 \frac{d^2}{dx^2} \log \tau_W(e^{-x}).$$

For $g \in \Gamma$ we would like to define a new potential by

$$(K(g)q)(x) = -2 \frac{d^2}{dx^2} \log \tau_{g^{-1}W}(e^{-x}) = -2 \frac{d^2}{dx^2} \log \tau_W(ge^{-x}), \quad (33)$$

for $q(x) = -2 \frac{d^2}{dx^2} \log \tau_W (e^{-x})$.

Let Ξ be a subset of $Gr^{(2)}(H)$ satisfying $g^{-1}\Xi = \Xi$ for every $g \in \Gamma$. Then, denoting

$$\Omega_\Xi = \left\{ q; \quad q(x) = -2 \frac{d^2}{dx^2} \log \tau_W (e^{-x}), \quad \exists W \in \Xi \right\},$$

Theorem4 enables us to define KdV flow $\{K(g)\}_{g \in \Gamma}$ on Ω_Ξ by (33). In this case, from (ii) of Lemma7 it follows that, for $g_1, g_2 \in \Gamma$

$$\begin{aligned} (K(g_1 g_2) q)(x) &= -2 \frac{d^2}{dx^2} \log \tau_W (g_1 g_2 e^{-x}) \\ &= -2 \frac{d^2}{dx^2} \log \tau_{g_2^{-1} W} (g_1 e^{-x}) = (K(g_1) K(g_2) q)(x) \end{aligned}$$

is valid, and hence the flow property $K(g_1 g_2) = K(g_1) K(g_2)$ holds. The purpose of this section is to construct KdV flow on the space Ω of generalized reflectionless potentials by this procedure.

For a fixed $E_0 \in (0, 1)$ let Σ be a set of measures on $[-\sqrt{E_0}, \sqrt{E_0}]$ satisfying

$$\int_{-1}^1 \frac{\sigma(d\xi)}{E_0 - \xi^2} \leq 1. \quad (34)$$

By this σ define

$$M_\sigma(z) = -z - \int_{-1}^1 \frac{\sigma(d\xi)}{\xi - z},$$

and set

$$W_\sigma = \{f_1(z^2) + M_\sigma(z)f_2(z^2); \quad f_1, f_2 \in H_+\}.$$

We show $W_\sigma \in Gr^{(2)}(H)$. To apply Theorem1 we compute the inverse matrix of Δ . Since

$$\varphi(z) = 0, \quad \psi(z) = \int_{-1}^1 \frac{\sigma(d\xi)}{\xi - z}$$

are valid, we see

$$\Delta(z) = \begin{pmatrix} 1 & \int_{-1}^1 \frac{\xi \sigma(d\xi)}{\xi^2 - z} \\ 0 & 1 + \int_{-1}^1 \frac{\sigma(d\xi)}{\xi^2 - z} \end{pmatrix},$$

and

$$X(z) = \Delta(z)^{-1} = \begin{pmatrix} 1 & -\frac{\int_{-1}^1 \frac{\xi \sigma(d\xi)}{\xi^2 - z}}{1 + \int_{-1}^1 \frac{\sigma(d\xi)}{\xi^2 - z}} \\ 0 & \frac{1}{1 + \int_{-1}^1 \frac{\sigma(d\xi)}{\xi^2 - z}} \end{pmatrix}.$$

Since the denominator is a Herglotz function, so is

$$-\frac{1}{1 + \int_{-1}^1 \frac{\sigma(d\xi)}{\xi^2 - z}},$$

which is expressible as

$$-\frac{1}{1 + \int_{-1}^1 \frac{\sigma(d\xi)}{\xi^2 - z}} = -1 + \int_0^1 \frac{\tilde{\sigma}(d\xi)}{\xi - z}$$

Since the support of σ is contained in $[-\sqrt{E_0}, \sqrt{E_0}]$, which is strictly in $[-1, 1]$, the entries of $\Delta(z)$ are holomorphic on $\{|z| > E_0^{-1}\}$. We prove that

$$1 + \int_{-1}^1 \frac{z\sigma(d\xi)}{z\xi^2 - 1}$$

has no zeros in $\{|z| < E_0^{-1}\}$. Suppose for some z satisfying $|z| < E_0^{-1}$ the above function vanishes. Then, in the identity

$$1 = -\int_{-1}^1 \frac{z\sigma(d\xi)}{z\xi^2 - 1} = \int_{-1}^1 \frac{\sigma(d\xi)}{z^{-1} - \xi^2} \quad (35)$$

taking the imaginary parts, we have

$$0 = (\operatorname{Im} z^{-1}) \int_{-1}^1 \frac{\sigma(d\xi)}{|z^{-1} - \xi^2|^2},$$

which is valid only when $\operatorname{Im} z^{-1} = 0$. Hence z^{-1} becomes real. However, we have $|z^{-1}| < E_0^{-1}$ and

$$\int_{-\sqrt{E_0}}^{\sqrt{E_0}} \frac{\sigma(d\xi)}{E_0 - \xi^2} \leq 1,$$

whence it is impossible for (35) to be satisfied. Therefore $X(z^{-1})$ is holomorphic on $\{|z| < E_0^{-1}\}$, and the conditions (i),(ii) of Theorem1 hold. Consequently we have

Lemma 12 *If $\sigma \in \Sigma$, then $W_\sigma \in Gr^{(2)}(H)$ is valid.*

Define a subgroup Γ_o of Γ by

$$\Gamma_o = \{g \in \Gamma ; g(z) = g(-z)^{-1}, g \text{ is real on } \mathbb{R} \text{ and } g(0) = 1\}.$$

The next task is to show $g^{-1}W_\sigma \in Gr^{(2)}(H)$ for every $g \in \Gamma_o$. The property $\tau_{W_q}(g) > 0$ shall be verified.

Let Σ_d be the set of all $\sigma \in \Sigma$ having a support consisting of finitely many points. For $\sigma \in \Sigma_d$ we compute $\tau_{W_q}(g)$. In this case assume $\{-\kappa_j\}_{1 \leq j \leq m}$ be the support of σ . Then they are distinct real numbers and are contained in $[-\sqrt{E_0}, \sqrt{E_0}]$. $M(z)$ is meromorphic and takes a form of

$$M(z) = -z - \sum_{j=1}^m \frac{\sigma_j}{-\kappa_j - z}, \quad \sigma_j > 0.$$

In this case

$$\varphi(z) = 0, \quad \psi(z) = \sum_{j=1}^m \frac{\sigma_j}{-\kappa_j - z}$$

hold. Every element f of W_σ has poles $\{-\kappa_j\}_{1 \leq j \leq m}$ of at most 1 and holomorphic on $\{|z| < 1\} \setminus \{-\kappa_j\}_{1 \leq j \leq m}$. Therefore, in this case the dimension of $P_{H_-} W_\sigma$ is m , which is finite, and T has eigenvalues $\{\kappa_j^2\}_{1 \leq j \leq m}$ and eigenvectors $\{(z + \kappa_j)^{-1}\}_{1 \leq j \leq m}$. Moreover

$$\delta(z) = \det \Delta(z) = 1 + \psi_o(z) = 1 + \sum_{j=1}^m \frac{\sigma_j}{\kappa_j^2 - z}$$

holds. Since, if $\text{Im } z \neq 0$, then $\text{Im } \delta(z) \neq 0$ holds, $\delta(z)$ has zeros only on \mathbb{R} . Moreover, the inequality (34) implies that $\delta(z)$ has no zeros on (E_0, ∞) neither. It is clear that $\delta(z)$ has no zeros for $\lambda \leq 0$. Therefore, $\delta(z)$ has zeros only on $(0, E_0]$. It is easy to see that every zero is simple, and there exists exactly one zero between two consecutive poles. Assume

$$i \neq j \implies \kappa_i^2 \neq \kappa_j^2, \quad (36)$$

and let zeros of $\delta(z)$ be

$$0 < \mu_1 < \mu_2 < \cdots < \mu_m,$$

which are eigenvalues of S . Choose ω_i such that $\omega_i^2 = \mu_i$ and $\omega_i > 0$. Then, set

$$m_i = 2\omega_i \prod_{j=1}^m \frac{\omega_i - \kappa_j}{\omega_i + \kappa_j} \prod_{j:j \neq i}^m \frac{\omega_i + \omega_j}{\omega_i - \omega_j}.$$

Lemma 13 $m_i > 0$.

Proof. Since $\delta(z)$ is a Herglotz function, so is $-\delta(z)^{-1}$, whence

$$-\delta(z)^{-1} = -1 + \sum_{i=1}^m \frac{\chi_i}{\mu_i - z}, \quad \text{with } \chi_i > 0.$$

On the other hand

$$\delta(z)^{-1} = \prod_{j=1}^m \frac{z - \kappa_j^2}{z - \mu_j} = \prod_{j=1}^m \frac{z - \kappa_j^2}{z - \mu_j}$$

is valid, hence

$$\begin{aligned} \chi_i &= \lim_{z \rightarrow \mu_i} (z - \mu_i) \delta(z)^{-1} = (\mu_i - \kappa_i^2) \prod_{j=1, j \neq i}^m \frac{\mu_i - \kappa_j^2}{\mu_i - \mu_j} \\ &= (\omega_i - \kappa_i)(\omega_i + \kappa_i) \prod_{j=1, j \neq i}^m \frac{\omega_i - \kappa_j}{\omega_i + \kappa_j} \frac{\omega_i + \omega_j}{\omega_i - \omega_j} \\ &= m_i \frac{(\omega_i + \kappa_i)^2}{2\omega_i} \prod_{j=1, j \neq i}^m \left(\frac{\omega_i + \kappa_j}{\omega_i + \omega_j} \right)^2. \end{aligned}$$

Since $\omega_i = \sqrt{\mu_i} > 0$, we completes the proof. ■

When the condition (36) is not satisfied, we have only to approximate σ by other σ s satisfying (36). Then we cite a result from ([11]).

Lemma 14 For $\sigma \in \Sigma_d$ and $g \in \Gamma_\sigma$, with some positive m_i depending only on $\{\omega_i, \kappa_j\}$ the following holds:

$$\tau_{W_\sigma}(g) = \left(\prod_{i=1}^m \frac{g(\omega_i)}{g(\kappa_i)} \right) \frac{\det \left(\delta_{ij} + \frac{\sqrt{m_i m_j}}{\omega_i + \omega_j} g(\omega_i)^{-1} g(\omega_j)^{-1} \right)}{\det \left(\delta_{ij} + \frac{\sqrt{m_i m_j}}{\omega_i + \omega_j} \right)}.$$

This lemma implies particularly $g^{-1}W_\sigma \in Gr^{(2)}(H)$ for $\sigma \in \Sigma_d$ and $g \in \Gamma_\sigma$. In this case the potential q associated with W_σ is rapidly decreasing and known to be described by the scattering data. Namely, the reflection coefficients vanish, $\{-\omega_i^2\}_{1 \leq i \leq m}$ are negative eigenvalues of L^q , and $\{m_i\}_{1 \leq i \leq m}$ are the normalization constants for the Jost solutions corresponding to $\{-\omega_i^2\}_{1 \leq i \leq m}$. By these data

$$q(x) = -2 \frac{d^2}{dx^2} \log \det \left(\delta_{ij} + \frac{\sqrt{m_i m_j}}{\omega_i + \omega_j} e^{(\omega_i + \omega_j)x} \right) \quad (37)$$

is valid.

Here we identify the M with the traditional Weyl functions for a 1-D Schrödinger operator. For a real bounded potential q on \mathbb{R} , it is known that for every $E \in \mathbb{C} \setminus \mathbb{R}$ there exist uniquely functions $f_\pm(x, E)$ satisfying

$$L^q f_\pm = E f_\pm \quad \text{with } f_\pm \in L^2(\mathbb{R}_\pm) \text{ and } f_\pm(0) = 1,$$

The classical Weyl functions m_\pm are defined by

$$m_\pm(E) = \pm \frac{d}{dx} f_\pm(x, E) \Big|_{x=0}.$$

the last argument shows that, for q associated with W_σ for $\sigma \in \Sigma_d$, identities

$$m_\pm(-E^2) = \pm M(\pm E), \quad f_+(x, E^2) = \frac{f_{W_\sigma}(x, E)}{f_{W_\sigma}(0, E)} \quad (38)$$

hold. Now define

$$\begin{cases} \Omega_{cl} = \{q; q \text{ is associated with } W_\sigma \text{ for } \sigma \in \Sigma_d\}, \\ \Omega = \{q; q \text{ is a compact uniform limit of some } q_s \text{ from } \Omega_{cl}\}. \end{cases}$$

Then, without difficulty we see the identities (38) are valid for $q \in \Omega$.

In the context of the Grassmann manifold the results by Marchenko[10] and Lundina[13] can be stated as follows.

Theorem 5 Ω coincides with the set of all potentials associated with W_σ for $\sigma \in \Sigma$. Moreover $q \in \Omega$ is holomorphic on the strip $\{|\operatorname{Im} x| < \sqrt{E_0}^{-1}\}$ having a uniform bound

$$|q(x)| < 2E_0 \left(1 - \sqrt{E_0} |\operatorname{Im} x|\right)^{-2}.$$

From Lemma14 and Theorem5 the theorem below is expected to be held. Although the proof of this theorem was given in ([11]),the proof contains an error and the detailed proof will be given elsewhere

Theorem 6 For every $\sigma \in \Sigma$ and $g \in \Gamma_o$ we have always $\tau_{W_\sigma}(g) > 0$. Namely $g^{-1}W_\sigma \in Gr^{(2)}(H)$ is valid.

Now let

$$\Xi = \{g^{-1}W_\sigma; \quad \sigma \in \Sigma, \quad g \in \Gamma_o\}.$$

Then, this theorem implies that $g^{-1}\Xi = \Xi$ is valid for any $g \in \Gamma_o$, and KdV flow $\{K(g)\}_{g \in \Gamma_o}$ can be constructed on Ω .

Remark 4 It is not clear for which W the identities (38) are valid. Properties of the Baker-Akhizer functions with respect to the variable x are hard to know, whereas properties with respect to the variable z are rather easier to see. We have to know asymptotic properties with respect to the variable x , since $f_\mp^\sigma(x, z)$ is defined as a function belonging to $L^2(\mathbb{R}_+)$. Therefore, in Theorem 26 of [11] the author made a wrong statement and it is not clear if the theorem holds or not under the condition.

5 Properties of KdV flow

5.1 Iso-spectral property

Several facts known in KdV integrable system are valid in this framework as well. For instance the following iso-spectral property holds if we find a suitable Lax-pair.

Theorem 7 For $q \in \Omega$, $g \in \Gamma_o$ there exists a unitary operator U on $L^2(\mathbb{R})$ such that

$$L_{K(g)q} = U^{-1}L_qU.$$

5.2 Invariant submanifolds

$\{K(g)\}_{g \in \Gamma_o}$ has many invariant submanifolds in Ω . The simplest one is Ω_{cl} . To introduce another type of invariant manifolds. In quantum physics potential q is real valued. Therefore we define a notion of "real" in $Gr^{(2)}(H)$.

(iv) $W \in Gr^{(2)}(H)$ is called **real** if the Weyl function M satisfies $M = \overline{M}$.

It is easy to see that the associated potential takes real values on \mathbb{R} if and only if $W \in Gr^{(2)}(H)$ is real. We introduce the following notion of reflectionlessness. We call a potential q is **reflectionless** on $F (\in \mathcal{B}(\mathbb{R}))$ if its Weyl functions satisfy

$$m_+(E + i0) = -\overline{m_-(E + i0)} \quad \text{for a.e. } E \in F,$$

or equivalently

$$M(\sqrt{-\xi} - i0) = \overline{M(-\sqrt{-\xi} + i0)} \quad \text{a.e. } \xi \in F.$$

and set

$$\mathcal{R}(F) = \{q \in \mathcal{Q}; \quad q \text{ is reflectionless on } F\}.$$

The theorem below assures that Γ_o acts on $\mathcal{R}(F)$.

Theorem 8 $\mathcal{R}(F \cup [0, \infty))$ is invariant under $\{K(g)\}_{g \in \Gamma_o}$.

Proof. Assume for some $F \subset [-1, 0]$

$$M_q \left(\sqrt{-\xi} - i0 \right) = \overline{M_q \left(-\sqrt{-\xi} + i0 \right)} \quad \text{a.e. } \xi \in F. \quad (39)$$

Since

$$M_{K(g)q}(z) = M_{g^{-1}W_q}(z),$$

from Theorem2 it follows that

$$\begin{aligned} & M_{K(g)q} \left(\sqrt{-\xi} - i0 \right) - \overline{M_{K(g)q} \left(-\sqrt{-\xi} + i0 \right)} \\ &= \frac{-c(\xi) M_q \left(\sqrt{-\xi} - i0 \right) + d(\xi)}{a(\xi) M_q \left(\sqrt{-\xi} - i0 \right) - b(\xi)} - \frac{-c(\xi) M_q \left(-\sqrt{-\xi} + i0 \right) + b(\xi)}{a(\xi) M_q \left(-\sqrt{-\xi} + i0 \right) - d(\xi)} = 0, \end{aligned}$$

where we have used the fact that $U(\xi)$ is real for $\xi \in \mathbb{R}$. ■

5.3 Almost periodicity

M.Sodin - P.Yuditskii[12] proved the almost periodicity of potentials if the associated Schrödinger operators has a homogeneous closed set as their spectrum and are assumed to be reflectionless on the spectrum from the point of view of harmonic analysis. We would like to apply this KdV flow to give a general condition on the spectrum under which potentials are almost periodic.

An approach to this problem is as follows: Set

$$\Omega_q = \{K(g)q; g \in \Gamma_o\}, \quad \Gamma_q = \{g \in \Gamma_o; K(g)q = q\}.$$

Then, denoting the natural map from Γ_o to Γ_o/Γ_q by

$$\Gamma_o \ni g \longrightarrow \tilde{g} \in \Gamma_o/\Gamma_q,$$

we define

$$\Gamma_o/\Gamma_q \ni \tilde{g} \longrightarrow K(g)q \in \Omega_q,$$

which gives a bijection

$$\Gamma_o/\Gamma_q \cong \Omega_q. \quad (40)$$

We introduce a condition on q :

Condition 1 Γ_o/Γ_q is compact.

Define a continuous map from Γ_o/Γ_q to \mathbb{R} by

$$\Theta(\tilde{g}) = (K(g)q)(0),$$

and a one-parameter group on Γ_o by

$$e_x(z) = e^{-xz}.$$

Then

$$q(x) = \Theta(\tilde{e}_x)$$

holds, and for any sequence $\{x_n\}_{n \geq 1}$ of \mathbb{R} we have

$$q(x + x_n) = \Theta(\tilde{e}_{x+x_n}) = \Theta(\tilde{e}_x \omega_n) \quad \text{with } \omega_n = \tilde{e}_{x_n} \in \Gamma_o/\Gamma_q.$$

Assuming the condition1 on q , we can find a subsequence $\{\omega_{n_k}\}$ of $\{\omega_n\}_{n \geq 1}$ and an element ω_∞ of Γ_o/Γ_q such that $\omega_{n_k} \rightarrow \omega_\infty$. Since Θ is uniformly continuous on Γ_o/Γ_q , we easily see that $q(x + x_{n_k})$ converges to $\Theta(\tilde{e}_x \omega_\infty)$ uniformly on \mathbb{R} , which implies the uniform almost periodicity of q .

Condition 2 Let F be a closed set of $[-1, 0]$. For any Herglotz function m satisfying

$$\operatorname{Re} m(\xi + i0) = 0 \quad \text{a.e. } \xi \in F,$$

the representing measure σ is absolutely continuous on F .

For such an F set

$$\Omega_F = \{q \in \mathcal{R}(F \cup [0, \infty)); \quad \operatorname{sp} L^q \subset F \cup [0, \infty)\}.$$

Then Ω_F is closed in Ω , and Theorems 7, 8 imply that Ω_F is invariant under $\{K(g)\}_{g \in \Gamma_o}$. Let

$$[-1, 0] \setminus F = \bigcup_{n=1}^N (\lambda_n, \mu_n) \quad \text{disjoint union.}$$

Then $N \leq \infty$ is valid. The Green function for L^q satisfies

$$\begin{aligned} g_E(0, 0, q) &= - \left(M(\sqrt{-E}) - M(-\sqrt{-E}) \right)^{-1}, \\ \lim_{x \downarrow 0} \frac{\partial^2}{\partial x \partial y} g_E(x, y, q) \Big|_{x > y > 0} &= \left(M(\sqrt{-E})^{-1} - M(-\sqrt{-E})^{-1} \right)^{-1}. \end{aligned}$$

Therefore, from $\operatorname{sp} L^q \subset F \cup [0, \infty)$ it follows that the representing measures of the two Herglotz functions vanish on each interval (λ_n, μ_n) . On the other hand,

$$\frac{M(z) - M(-z)}{-2z} = 1 + \int_{-1}^1 \frac{1}{\zeta^2 - z^2} \sigma(d\zeta) = \delta(z^2)$$

is valid, hence

$$- \left(M(\sqrt{-E}) - M(-\sqrt{-E}) \right)^{-1} = \frac{1}{2\sqrt{-E}\delta(-E)}$$

holds, which implies that $\delta(-E)^{-1}$ is increasing on each (λ_n, μ_n) and has at most one zero in (λ_n, μ_n) . Now set

$$\begin{cases} \xi_n = \lambda_n & \text{if } \delta(-E) > 0 \text{ for } E \in (\lambda_n, \mu_n), \\ \xi_n = \mu_n & \text{if } \delta(-E) < 0 \text{ for } E \in (\lambda_n, \mu_n), \\ \xi_n \text{ is a zero of } \delta(-E)^{-1} & \text{otherwise.} \end{cases}$$

Since $\delta(-E)$ has poles at $\sqrt{-\xi_n}$ or $-\sqrt{-\xi_n}$ of order 1, σ has a mass σ_n there. Then the representing measure of $\left(M(\sqrt{-E})^{-1} - M(-\sqrt{-E})^{-1} \right)^{-1}$ vanishes on (λ_n, μ_n) as well, hence σ can not have a mass simultaneously both at $\pm\sqrt{-\xi_n}$. Now set

$$\begin{aligned} \epsilon_n &= 1 \quad \text{if } \sigma\left(\left\{\sqrt{-\xi_n}\right\}\right) > 0, \\ \epsilon_n &= -1 \quad \text{if } \sigma\left(\left\{-\sqrt{-\xi_n}\right\}\right) > 0. \end{aligned}$$

If $\xi_n = \lambda_n$ or μ_n , we define ϵ_n to be ± 1 . We are identifying here with $(\lambda_n, 1)$ and $(\lambda_n, -1)$, and $(\mu_n, 1)$ and $(\mu_n, -1)$.

Proposition 2 Under the condition 2, Ω_F corresponds bijectively to

$$\left\{ (\xi_n, \epsilon_n)_{n \geq 1}; \xi_n \in [\lambda_n, \mu_n], \epsilon_n = \pm 1 \right\} \cong \prod_{n=1}^N \left\{ z \in \mathbb{C}; \left| z - \frac{\lambda_n + \mu_n}{2} \right| = \frac{\mu_n - \lambda_n}{2} \right\}.$$

Now, if $q \in \Omega_F$, then Theorem 8 implies $\Omega_q \subset \Omega_F$. We propose a conjecture:

Conjecture 1 Under the condition 2 $\Omega_q = \Omega_F$ and Γ_o/Γ_q is compact.

This conjecture implies the uniform almost periodicity of q . We show how this approach works in case $N < \infty$. In this case Ω_F is homeomorphic to an N dimensional torus \mathbb{T}^N which makes it possible to introduce a differential structure into Ω_F . For each $g = e^h \in \Gamma_o$ a 1-parameter group e^{th} acts smoothly on Ω_F through the KdV flow $K(g)$. Let X_k be the vector field associated with the one-parameter group $e^{-tz^{2k+1}}$ on Ω_F . The vector fields $\{X_k\}_{1 \leq k \leq N}$ are commute. The classical theorem by Liouville shows:

Lemma 15 Suppose $\{X_k\}_{1 \leq k \leq N}$ are independent at each point of Ω_F . Then, for each $q \in \Omega_F$ the map

$$\Theta : \Gamma_o/\Gamma_q \rightarrow \Omega_F$$

induces a homeomorphism. This, especially implies that $\Omega_q = \Omega_F$ and Γ_o/Γ_q is compact.

Consequently, a potential $q \in \Omega_F$, which is given by

$$q(x) = \Theta(\tilde{e}_x)$$

becomes quasi periodic. Therefore all we have to do is to prove the linear independence of $\{X_k\}_{1 \leq k \leq N}$. To describe $\{X_k\}_{1 \leq k \leq N}$ we use the parameter $\{\xi_n\}_{1 \leq n \leq N} \in \mathbb{T}^N$. Let $\{\xi_n(t)\}_{1 \leq n \leq N}$ be the parameter representing the potential $K(e^{-tz^{2k+1}})q$. Then Dubrovin[15] showed $\{\xi_n(t)\}_{1 \leq n \leq N}$ satisfies an ODE on \mathbb{T}^N . The vector field X_k precisely corresponds to the ODE and we can know the linear independence of $\{X_k\}_{1 \leq k \leq N}$.

References

- [1] K.Ishii: Localization of eigenstates and transport phenomena in one-dimensional disordered systems, Supp. Prog. Theor. Phys., 53(1973),77-138
- [2] I.Goldseid - S.Molchanov - L.Pastur: A pure point spectrum of the stochastic and one-dimensional equation, Funct. Anal. Appl. 11(1977), 1-10
- [3] S.Molchanov: The structure of eigen-functions of one-dimensional un-ordered structures, Math. USSR Izv. 12(1978), 69-101
- [4] S.Kotani: Ljapounov indices determine absolutely continuous spectra of stationary random Schrödinger operators, Stochastic Analysis (Katata/Kyoto, 1982), 225-247, North-Holland Math. Library, 32, North-Holland, Amsterdam, 1984

- [5] S.Jitomirskaya: Ergodic Schrödinger operators (on one foot), Proceedings of Symposia in Pure Mathematics Vol. 76.2(2007), 613-647
- [6] D.Damanik: Lyapunov Exponents and Spectral Analysis of Ergodic Schrödinger operators: A Survey of Kotani Theory and Its Applications, Proceedings of Symposia in Pure Mathematics Vol. 76.2(2007), 539-563
- [7] D.B.Pearson: Value distribution and spectral analysis of differential operators, J. Phys. A: Math. Gen. 26(1993), 4067-4080
- [8] G.Segal - G.Wilson: Loop groups and equations of KdV type, Publ. IHES, 61(1985), 5-65
- [9] C.Remling: The absolutely continuous spectrum of Jacobi matrices, to appear in Annals of Math.
- [10] V.A.Marchenko: The Cauchy problem for the KdV equation with non-decreasing initial data, Springer Series in Nonlinear Dynamics, What is Integrability? ed. by V.E. Zakharov (1990), 273-318
- [11] S.Kotani: KdV flow on Generalized Reflectionless Potentials, Journal of Mathematical Physics, Analysis, Geometry, vol.4, No.4(2008), 490-528
- [12] M.Sodin - P.Yuditskii: Almost periodic Sturm-Liouville operators with Cantor homogeneous spectrum, Comment Math. Helvetici, 70(1995), 639-658.
- [13] D.S.Lundina: Compactness of the set of reflectionless potentials, Teor. Funktsii Funktsional Anal. i Prilozhen, 44(1985) 55-66
- [14] R.A.Johnson: On the Sato-Segal-Wilson solutions of the K-dV equation, Pacific Jour. of Math., 132(1988) 343-355
- [15] B.A.Dubrovin: Periodic problem for the Korteweg-de Vries equation in the class of finite band potentials, Funct. Anal. Appl. 9(1975), 215-223

Hamiltonian Identification for a 3-Spin Chain ¹

Mohammad Ali Fasihi^A, 田中 宗^B, 中原 幹夫^{A,B}, 近藤 康^{A,B}

^A 近畿大学 総合理工学研究科 量子コンピュータ研究センター

^B 近畿大学 理工学部 理学科 物理学コース

我々は、端のスピンのみ操作・観測が可能であるという制限のもと、横磁場イジングモデルのスピン鎖において、相互作用や横磁場の値を推定する手法を考案した [1]。本研究で提案した推定手法は4つの仮定を前提としている。この仮定はいずれも、液体 NMR 量子コンピュータなどの多くの実験系で満たすことができるものである。

1 導入

量子情報処理に関する研究は、物理学や情報科学、あるいは数学の垣根を越えて精力的に進められてきている [2-4]。中でも、実用に耐えうる量子コンピュータの物理的実現に向けた研究が盛んである。こうした研究の潮流は大きく2つに分けることができる。

- コヒーレンス時間の長い良質な量子ビットの作製と集積化などの「ハード」の研究
- 量子ビットを無駄なく、正確に操作するという量子操作の観点からの研究

両者とも、量子情報処理の実現に対して必要不可欠な研究である。前者については、近年の実験技術の進展によりいくつかの有望な候補となる系が提案されつつある。良質な量子ビットを作成することができたら、次の課題は量子ビット間に働く相互作用を高精度に決定することである。量子ビット間の相互作用が分からないと、所望の操作は不可能であり、量子計算を実行することができない。すべての量子ビットにアクセスすることができれば、非常に多くの情報を得ることができ、相互作用を推定することは容易である。しかしながら、場合によっては、端にある量子ビットにしかアクセスできない場合もあり得る。また、デコヒーレンスの原因となる外界とのインターフェイスを減らすために、端にある量子ビットのアクセスのみを許すことも考えられる。そのようなときに、いかにして量子ビット間の相互作用を推定するか？ というのが我々の問題である。

これまでも、制限されたアクセスの元で相互作用を推定する方法の提案はいくつかなされてきた [5-13]。 N 個の量子ビットからなるハイゼンベルクモデルや XXZ モデルの相互作用推定手法については、Burgarth らによってまとめられている [8,9]。 N 個の量子ビットの状態空間の次元は 2^N であるが、これらの系には良い量子数 (total S^z) が存在することを巧みに利用して、 N 次元

¹この原稿は、2010年11月4日～6日に京都大学基礎物理学研究所にて開催された基研研究会「量子科学における双対性とスケール」の研究会報告の原稿である。

の空間を考えることにより、ハミルトニアンを推定を行うことを可能にしている。この場合には、初期状態として良い量子数で特徴付けられるような状態を採用する必要がある。また、Di Francoらは N 個の量子ビットからなる 1 次元 XY モデルの相互作用推定を行った [10]。 XY モデルも良い量子数が存在するが、彼らはそれを用いずに（つまり、任意の初期状態から）、相互作用の推定が可能であることを示した。またごく最近、Burgarthらはいくつかの代表的な量子スピンモデルでのハミルトニアンを推定する手法を、準粒子描像の観点から系統的にまとめあげた [11]。

本研究の目的は NMR 量子コンピュータへの応用を想定して、サイトに依存する未知の横磁場が印加されたイジングモデルの相互作用の推定手法を考案することである。本研究では最も簡単な例として 3 スピンから成る系を考察した。2 章では考察したハミルトニアンと我々の提案する手法を述べ、時間発展の厳密な表式を示す。3 章では相互作用パラメータを推定する手法を述べる。4 章では本研究の結論と今後の展望について述べる。

2 ハミルトニアンと我々の提案する手法

我々は 3 つのスピンの系から成る、サイトに依存する横磁場や相互作用のある横磁場イジングモデルについて、端の 1 スピンの制御と測定のみにより相互作用と横磁場の値を推定する手法を考案した。我々が考察したハミルトニアンは

$$\mathcal{H}(\{h_i\}, \{J_i\}) = \sum_{i=1}^2 J_i \sigma_i^z \sigma_{i+1}^z - \sum_{i=1}^3 h_i \sigma_i^x \quad (1)$$

である。ここで、 σ_i^x や σ_i^z はそれぞれサイト i におけるスピン $S = 1/2$ のパウリ行列の x 成分、 z 成分である。このハミルトニアンは NMR 量子コンピュータのハミルトニアンに対応している [4, 14]。一般性を失わずに、各サイトにかかっている横磁場の方向を x 軸に取ることが可能である。後に述べるように、我々は 1 番目のスピンの x 方向の時間発展のみを検討するが、それは 2 番目や 3 番目の磁場の方向に依存しないことが以下によって分かる。

$$V^\dagger \sigma_1^x V = \sigma_1^x, \quad V = e^{-i(\alpha_2 \sigma_2^z + \alpha_3 \sigma_3^z)} \quad (2)$$

ここでユニタリ行列 V は 2 番目、3 番目の量子ビットを xy 平面で角度 α_2, α_3 だけ回転することに相当する。また以下では h_2, h_3 を正として考えることにする。

本研究の目的は、1 番目の量子ビットのみを操作・測定することが可能であるという条件下で相互作用のパラメータを推定することである。以下の 4 つの条件を仮定する。

- 初期状態として $|\uparrow\uparrow\uparrow\rangle$ を用意することが可能である。
- 1 番目の量子ビットのみ独立に操作することが可能である。
この仮定と上の仮定を用いて、 $(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)|\uparrow\uparrow\rangle$ という状態を用意することができる。
- 初期状態を用意したら、1 番目の量子ビットにかかっている磁場のみ制御することが可能である。2 番目、3 番目の量子ビットにかかる横磁場は未知のまま変化しないとする。

- 1 番目の量子ビットの x 方向の成分、すなわち、 σ_1^x のみ測定可能である。

我々の提案する手法は以下のようにまとめられる。

step 1 初期状態 $|\psi(0)\rangle$ を用意する。

step 2 1 番目の量子ビットの x 成分の時間発展の x 成分の期待値 $\langle\sigma_1^x(t)\rangle = \langle\psi(t)|\sigma_1^x|\psi(t)\rangle$ を測定する。ここで、 $|\psi(t)\rangle = e^{-i\mathcal{H}t}|\psi(0)\rangle$ である。

step 3 $\langle\sigma_1^x(t)\rangle$ のフーリエ変換 $\hat{\sigma}_1^x(\omega)$ を求める。

step 4 いくつかの h_1 について step 1 から step 3 を実行する。 $\hat{\sigma}_1^x(\omega)$ のピーク位置の h_1 依存性から $h_2, h_3, |J_1|, |J_2|$ を求める。

step 5 J_1 と J_2 の符号の組み合わせは 4 通り考えられる。その 4 通りの場合のそれぞれの $\langle\sigma_1^x(t)\rangle$ の計算結果と実測データとを比較して J_1 と J_2 の符号を決定する。

ここで、ハミルトニアン \mathcal{H} の固有値を $\{\epsilon_i\}_{1 \leq i \leq 8}$ とおく。具体的な表式は以下のように与えられる [1]。

$$\begin{cases} \epsilon_1 = \frac{\alpha}{2} + \frac{(\beta+\gamma)^2}{2} \\ \epsilon_2 = \frac{\alpha}{2} + \frac{(\beta-\gamma)^2}{2} \\ \epsilon_3 = \frac{\alpha}{2} - \frac{(\beta-\gamma)^2}{2} \\ \epsilon_4 = \frac{\alpha}{2} - \frac{(\beta+\gamma)^2}{2} \end{cases}, \quad \begin{cases} \epsilon_5 = -\epsilon_4 =: \epsilon_4 \\ \epsilon_6 = -\epsilon_3 =: \epsilon_3 \\ \epsilon_7 = -\epsilon_2 =: \epsilon_2 \\ \epsilon_8 = -\epsilon_1 =: \epsilon_1 \end{cases} \quad (3)$$

ただしここで、

$$\begin{aligned} \alpha &:= \sqrt{\frac{1}{3} \left(4S + \frac{B}{A} + \frac{A}{2^{1/3}} \right)}, \quad \beta := \frac{1}{3} \left(8S - \frac{B}{A} - \frac{A}{2^{1/3}} \right), \quad \gamma := \frac{16h_1h_2h_3}{\alpha}, \\ S &:= h_1^2 + h_2^2 + h_3^2 + J_1^2 + J_2^2, \quad A := \left[X + \sqrt{-256 \left(S^2 + 3\sqrt{\det\mathcal{H}} \right)^3 + X^2} \right]^{1/3}, \\ X &:= 16 \left[-S^3 + 9S\sqrt{\det\mathcal{H}} + 108h_1^2h_2^2h_3^2 \right], \quad B := 2^{7/3} \left[S^2 + 3\sqrt{\det\mathcal{H}} \right], \\ \sqrt{\det\mathcal{H}} &= h_1^4 + h_2^4 + (h_3^2 - J_1^2 + J_2^2)^2 - 2h_1^2(h_2^2 + h_3^2 - J_1^2 + J_2^2) + 2h_2^2(-h_3^2 + J_1^2 + J_2^2). \end{aligned} \quad (4)$$

である。また、式 (3) より明らかに、固有値は以下の関係を満たす。

$$\begin{cases} \epsilon_1 + \epsilon_4 = \epsilon_2 + \epsilon_3 \\ \epsilon_1 - \epsilon_3 = \epsilon_2 + \epsilon_4 \\ \epsilon_1 - \epsilon_2 = \epsilon_3 - \epsilon_4 \end{cases} \quad (5)$$

初期状態として、 $|\psi(0)\rangle = |\uparrow\uparrow\uparrow\rangle$ を用意する。時間発展演算子 $U(t)$ は

$$U(t) = \exp(-i\mathcal{H}t) = \sum_{j=1}^8 \exp(-i\epsilon_j t) \mathcal{P}_j, \quad \mathcal{P}_j = \prod_{k=1, k \neq j}^8 \frac{\mathcal{H} - \epsilon_k I}{\epsilon_j - \epsilon_k} \quad (6)$$

と表される。 \mathcal{P}_j は固有エネルギー ϵ_j に属する固有空間への射影演算子であり、 I は 8×8 の単位行列である。 $U(t)$ を用いて、時間 $t > 0$ における状態は $|\psi(t)\rangle = U(t)|\uparrow\uparrow\uparrow\rangle$ となる。

1 番目のスピンの x 成分の時間発展は

$$\langle \sigma_1^x(t) \rangle = \langle \psi(t) | \sigma_1^x | \psi(t) \rangle = C + \sum_{m=1, m>n}^4 [A_{mn} \cos(\epsilon_m + \epsilon_n)t + B_{mn} \sin(\epsilon_m - \epsilon_n)t] \quad (7)$$

となる。ここで C, A_{mn}, B_{mn} は $\{J_i\}, \{h_i\}$ の関数である。式 (7) より、 $\hat{\sigma}_1^x(\omega)$ は任意の m, n の組み合わせの $\omega = \epsilon_m \pm \epsilon_n$ でピークが現れうる。しかし以下の A_{mn}, B_{mn} は恒等的に 0 になるため、それに対応する m, n の組み合わせでは確実にピークが現れない。

$$A_{41} = A_{32} = B_{21} = B_{31} = B_{42} = B_{43} = 0 \quad (8)$$

また、 A_{mn}, B_{mn}, C について以下の和則が成立する。

$$C + \sum_{m=1, m>n}^4 (A_{mn} + B_{mn}) = \left\langle \uparrow\uparrow\uparrow \left| \left(\sum_{j=1}^8 \mathcal{P}_j \right) \sigma_1^x \left(\sum_{k=1}^8 \mathcal{P}_k \right) \right| \uparrow\uparrow\uparrow \right\rangle = \langle \uparrow\uparrow\uparrow | \sigma_1^x | \uparrow\uparrow\uparrow \rangle = 0 \quad (9)$$

ここで、完全性条件 $\sum_{j=1}^8 \mathcal{P}_j = I$ を用いた。

3 ハミルトニアンのパラメータ推定

ハミルトニアンのパラメータ $\{J_i\}, \{h_i\}$ を推定するために、我々は $\hat{\sigma}_1^x(\omega)$ のどのピークがどの $\epsilon_m \pm \epsilon_n$ に対応しているかを同定しなければならない。以降、数値データを示す際には、 $J_1 = 2, J_2 = 3, h_2 = 4, h_3 = 5$ とする。本研究で展開した議論は、パラメータに依存しないことを強調しておく。図 1 は、 A_{mn}, B_{mn} が恒等的に 0 にならない m, n に関する $\epsilon_m \pm \epsilon_n$ の h_1 依存性である。太線で示したのが $\epsilon_m + \epsilon_n$ であり、細い線で示したのが $\epsilon_m - \epsilon_n$ である。

図 1 から、中間の h_1 で $\epsilon_m + \epsilon_n$ や $\epsilon_m - \epsilon_n$ が交差することが分かる。そのため、ある 1 つの h_1 の値からだけでは、 $\hat{\sigma}_1^x(\omega)$ のどのピークがどの $\epsilon_m \pm \epsilon_n$ なのかを同定することはできない。しかし後に述べるように、 $|h_1| \gg h_2, h_3, |J_1|, |J_2|$ の場合や、 $|h_1| \ll h_2, h_3, |J_1|, |J_2|$ においては、 $\epsilon_m \pm \epsilon_n$ の決定は容易である。3.1 節では前者の場合、3.2 節では $h_1 = 0$ の場合について考察する。ただし $h_1 = 0$ の場合は、初期状態を $|\uparrow\uparrow\uparrow\rangle$ とすると、任意の時刻 t で $\langle \sigma_1^x(t) \rangle = 0$ となってしまう。そのためこの場合においては初期状態を変えなければならない。それについては 3.2 節で説明する。

原理的には $\hat{\sigma}_1^x(\omega)$ のピークの高さから $\epsilon_m \pm \epsilon_n$ を決定することは可能である。しかし以下の 2 点の理由から、現実的にはそれは困難であると言える。

- A_{mn}, B_{mn} の解析的表式は複雑であり、どのピークがどの $\epsilon_m \pm \epsilon_n$ に対応しているかを同定するには、面倒な数値的解析が必要である。
- 実際の実験においては、緩和の効果によりピークの幅は有限になる。緩和が ω に依存しない保証はないなどの測定上の問題がある。

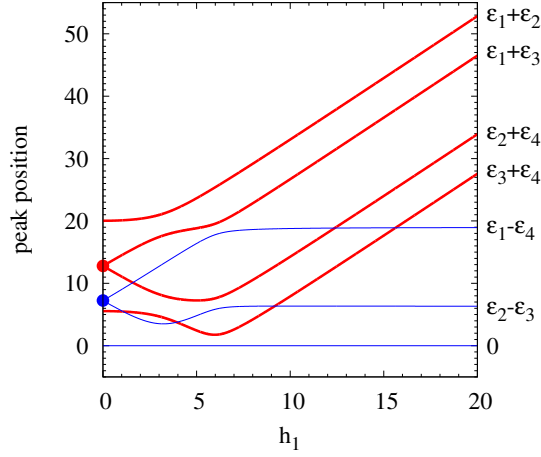


図 1: $J_1 = 2, J_2 = 3, h_2 = 4, h_3 = 5$ の場合の、 $\epsilon_m \pm \epsilon_n$ の h_1 依存性。太線が $\epsilon_m + \epsilon_n$ 、細い線が $\epsilon_m - \epsilon_n$ に対応する。丸は 3.2 節で取り扱う、 $h_1 = 0$ の場合の $\hat{\sigma}_1^x(\omega)$ のピーク位置である。

図 2 に、有限の値を取る A_{mn}, B_{mn}, C の h_1 依存性を示した。

A_{21}, A_{43}, C のみは ω に対して奇関数となっている。ただし、 A_{31} の h_1 が正の部分（負の部分）と A_{42} の h_1 が負の部分（正の部分）を合わせると奇関数となる。 B_{41} と B_{32} についても同様である。このような、 A_{mn}, B_{mn} の奇関数性が失われる原因はエネルギー固有値のラベル付けの仕方から来ている。また、図 2 より、先ほど述べた式 (9) が成立していることも確かめることができる。

3.1 $|h_1| \gg h_2, h_3, |J_1|, |J_2|$ の極限

ここでは $|h_1|$ が非常に大きい場合、すなわち、 $|h_1| \gg h_2, h_3, |J_1|, |J_2|$ の極限を考える。 $\hat{\sigma}_1^x(\omega)$ のピーク位置のうち、 $\epsilon_m + \epsilon_n$ ($1 \leq m, n \leq 4$) に対応するものは $h_1 \rightarrow \infty$ で線形に発散する。それに対して、 $\epsilon_m - \epsilon_n$ ($1 \leq n < m \leq 4$) に対応するものは一定値に収束する。 $|h_1| \gg h_2, h_3, |J_1|, |J_2|$ でのそれぞれの値の漸近形は、

$$\epsilon_1 + \epsilon_2 \rightarrow 2h_1 + \left[\sqrt{(h_2 + h_3)^2 + J_2^2} + \sqrt{(h_2 - h_3)^2 + J_2^2} \right] + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (10)$$

$$\epsilon_1 + \epsilon_3 \rightarrow 2h_1 + \left[\sqrt{(h_2 + h_3)^2 + J_2^2} - \sqrt{(h_2 - h_3)^2 + J_2^2} \right] + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (11)$$

$$\epsilon_2 + \epsilon_4 \rightarrow 2h_1 - \left[\sqrt{(h_2 + h_3)^2 + J_2^2} - \sqrt{(h_2 - h_3)^2 + J_2^2} \right] + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (12)$$

$$\epsilon_3 + \epsilon_4 \rightarrow 2h_1 - \left[\sqrt{(h_2 + h_3)^2 + J_2^2} + \sqrt{(h_2 - h_3)^2 + J_2^2} \right] + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (13)$$

$$\epsilon_1 - \epsilon_4 \rightarrow 2\sqrt{(h_2^2 + h_3^2) + J_2^2} + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (14)$$

$$\epsilon_2 - \epsilon_3 \rightarrow 2\sqrt{(h_2^2 - h_3^2) + J_2^2} + \mathcal{O}\left(\frac{1}{h_1}\right) \quad (15)$$

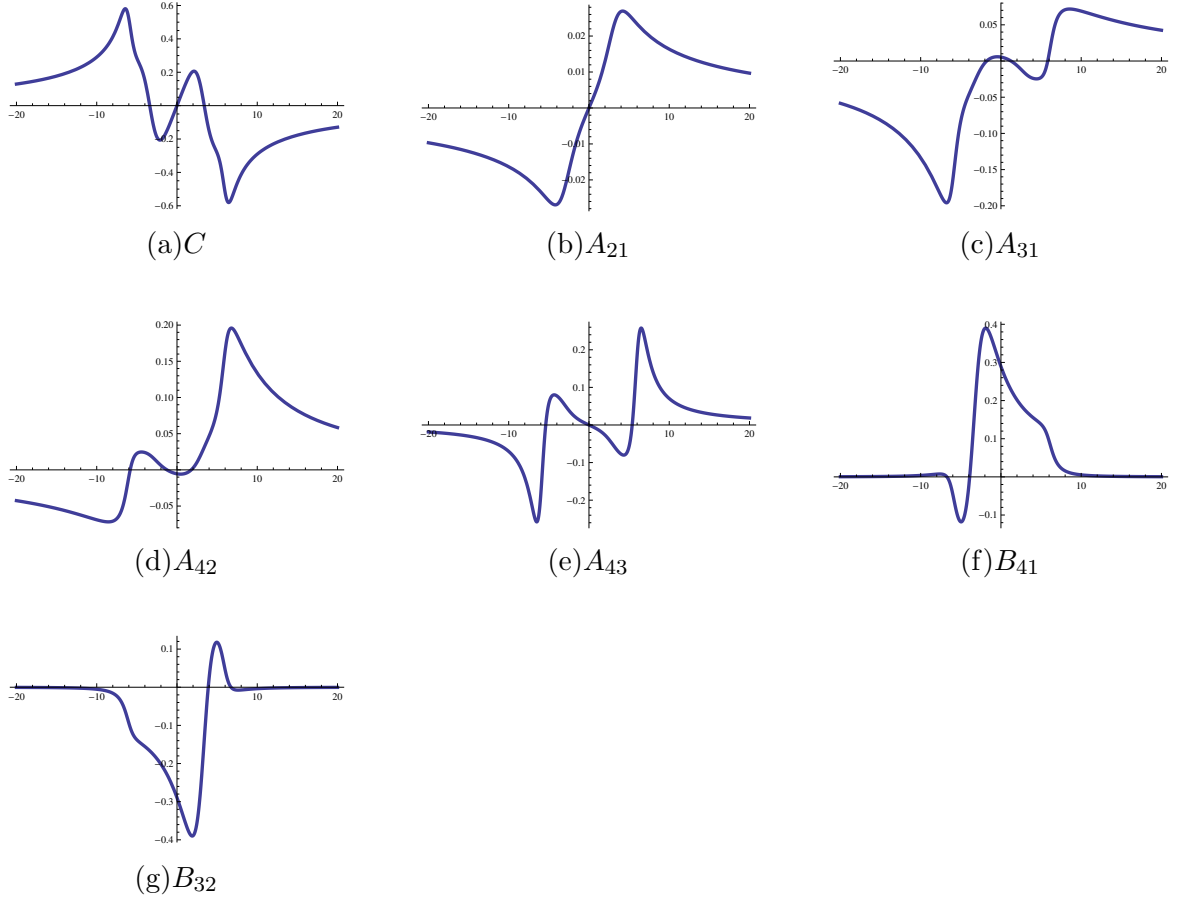


図 2: $J_1 = 2, J_2 = 3, h_2 = 4, h_3 = 5$ の場合の、 $\hat{\sigma}_1^x(\omega)$ のピークの高さの h_1 依存性。それぞれ (a) C 、(b) A_{21} 、(c) A_{31} 、(d) A_{42} 、(e) A_{43} 、(f) B_{41} 、(g) B_{32} に対応する。他のピークの高さは恒等的に 0 である (式 (8))。

これらの結果を用いると、十分大きい h_1 では $\hat{\sigma}_1^x(\omega)$ のピーク位置と、それに対応する $\epsilon_m \pm \epsilon_n$ を以下の式から同定することができる。

$$\epsilon_1 + \epsilon_2 > \epsilon_1 + \epsilon_3 > \epsilon_2 + \epsilon_4 > \epsilon_3 + \epsilon_4, \quad \epsilon_1 - \epsilon_4 > \epsilon_2 - \epsilon_3 \quad (16)$$

$\hat{\sigma}_1^x(\omega)$ のピーク位置と $\epsilon_m \pm \epsilon_n$ の対応関係を決定することができる、 $\sqrt{(h_2 \pm h_3)^2 + J_2^2}$ の値を求めることができる。我々はこの問題で 4 つの未知数、すなわち $h_2, h_3, |J_1|, |J_2|$ を求める必要がある、これだけでは条件式が足りない。そのため別の条件を考える必要がある。そこで最も簡単な $h_1 = 0$ の場合について 3.2 節で考察する。

$h_1 = 0$ の場合について考察する前に、式 (10) から (15) で与えられる漸近形が J_1 に依らない理由を簡単に説明する。以下のようにハミルトニアンを分割する。

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \quad \mathcal{H}_0 = -h_1 \sigma_1^x, \quad \mathcal{H}_1 = \sum_{i=1}^2 J_i \sigma_i^z \sigma_{i+1}^z - \sum_{i=2}^3 h_i \sigma_i^x \quad (17)$$

ここで非摂動ハミルトニアン \mathcal{H}_0 の固有ベクトルは、 $s_2, s_3 \in \{\uparrow, \downarrow\}$ として、 $1/\sqrt{2}(|\uparrow\rangle + |\downarrow\rangle) |s_2 s_3\rangle$ となる。1 次の摂動では固有値は J_1 に依存しない。このことが、式 (10) から (15) で与えられる

漸近形が J_1 に依らない理由である。またこの事実は 1 番目のスピンの時間発展からも定性的に理解することができる。 h_1 が他のパラメータに比べて十分大きいとする。このとき、1 番目のスピンの運動の特徴的な時間スケール τ は他の時間スケールに比べて非常に短くなる。そのため、相互作用 J_1 の効果は $t \gg \tau$ では平均化されてしまい、 J_1 を無視することができる。この議論は 3 スピン系だけではなく、一般の N 個のスピンの系に対しても適用可能である。

3.2 $h_1 = 0$ の場合

次に、3.1 節とは逆の極限、すなわち $h_1 = 0$ の場合について考える。この場合固有値は $\pm\epsilon_I, \pm\epsilon_{II}$ となり、それぞれ 2 重縮退している。ここで $\epsilon_I > \epsilon_{II} > 0$ とする。それぞれの表式は、

$$\epsilon_I = \sqrt{S_0 + 2\sqrt{C_0}}, \quad \epsilon_{II} = \sqrt{S_0 - 2\sqrt{C_0}}, \quad (18)$$

$$S_0 = h_2^2 + h_3^2 + J_1^2 + J_2^2, \quad C_0 = h_2^2 h_3^2 + h_3^2 J_1^2 + J_1^2 J_2^2 \quad (19)$$

で与えられる。

$h_1 = 0$ のとき、明らかに任意の $t > 0$ において

$$\langle \uparrow\uparrow\uparrow | U(t)^\dagger \sigma_1^x U(t) | \uparrow\uparrow\uparrow \rangle = 0 \quad (20)$$

が成立する。そのためこれまでとは別の初期状態を考える必要がある。ここで、 $h_1 = 0$ の場合には $|+\uparrow\rangle$ を初期状態として用意することにした。ただしここで $|+\rangle := 1/\sqrt{2}(|\uparrow\rangle + |\downarrow\rangle)$ である。このとき、1 番目のスピンの x 成分の時間発展は以下のように表される。

$$\langle \sigma_1^x(t) \rangle = \frac{h_2^2 h_3^2}{C_0} + A_{II,I}^0 \cos(\epsilon_I + \epsilon_{II})t + B_{II,I}^0 \sin(\epsilon_I - \epsilon_{II})t \quad (21)$$

ただしここで $A_{II,I}^0, B_{II,I}^0$ は、

$$A_{II,I}^0 := \frac{J_1^2 [h_2^2 (h_3^2 - J_2^2) - (h_3^2 + J_2^2) (h_3^2 - J_1^2 + J_2^2) + (J_2^2 + h_3^2) \epsilon_I \epsilon_{II}]}{2\epsilon_I \epsilon_{II} C_0} \quad (22)$$

$$B_{II,I}^0 := \frac{J_1^2 [h_2^2 (-h_3^2 + J_2^2) - (h_3^2 + J_2^2) (h_3^2 - J_1^2 + J_2^2) + (J_2^2 + h_3^2) \epsilon_I \epsilon_{II}]}{2\epsilon_I \epsilon_{II} C_0} \quad (23)$$

と定義した。ここで、 $J_1 = 2, J_2 = 3, h_1 = 0, h_2 = 4, h_3 = 5$ の場合、

$$\frac{h_2^2 h_3^2}{C_0} = 0.746269, \quad A_{II,I}^0 = 0.0242651, \quad B_{II,I}^0 = 0.229466 \quad (24)$$

となる。これらの値は

$$C + A_{II,I}^0 + B_{II,I}^0 = \langle +\uparrow | \sigma_1^x | +\uparrow \rangle = 1 \quad (25)$$

を満たす。

$h_1 = 0$ における $\hat{\sigma}_1^x(\omega)$ のピーク位置は $\omega = \epsilon_I \pm \epsilon_{II}$ であるから、式 (19) で定義される S_0, C_0 が分かる。以上の結果と、3.1 節で得られた結果とを合わせることで、4 つのパラメータ $h_2, h_3, |J_1|, |J_2|$ を決定することができる。

一度 $\hat{\sigma}_1^x(\omega)$ のピーク位置と $\epsilon_m \pm \epsilon_n$ が同定できたら、他の有限の h_1 における $\hat{\sigma}_1^x(\omega)$ のピーク位置を得ることによって、4つのパラメータ $h_2, h_3, |J_1|, |J_2|$ をより高精度に決定することができる。

また図2から、 $h_1 \rightarrow \infty$ で全てのピークの高さは小さくなるのが分かる。つまり、 h_1 がある程度大きくなるとピークが見えなくなることも十分予想できる。しかしながら、 $h_1 = 0$ の場合のピーク位置を特定し、その後いくつかの有限の h_1 の場合において $\hat{\sigma}_1^x(\omega)$ のピーク位置を得ることでパラメータを推定することが可能になる。

3.3 相互作用の符号の決定

以上の議論では、相互作用の符号を決定することはできなかつた。4つのパラメータ $h_2, h_3, |J_1|, |J_2|$ を決定するための全ての式が、 J_1, J_2 の2乗の関数となっているからである。図3は、それぞれの J_1, J_2 の符号の組み合わせに対する、 $|J_1| = 4, |J_2| = 5, h_1 = 1, h_2 = 6, h_3 = 7$ の場合の1番目のスピンの x 成分の時間発展である。上から順に $(\text{sign}(J_1), \text{sign}(J_2)) = (+, +), (-, +), (+, -), (-, -)$ である。初期状態は $|\psi(0)\rangle = |\uparrow\uparrow\uparrow\rangle$ とした。

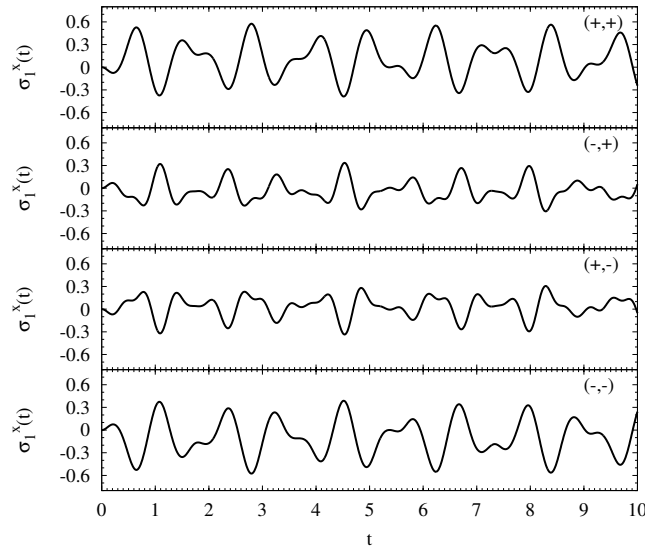


図3: $|J_1| = 2, |J_2| = 3, h_1 = 1, h_2 = 4, h_3 = 5$ の場合の、1番目のスピンの x 成分の時間発展を表している。初期状態は $|\psi(0)\rangle = |\uparrow\uparrow\uparrow\rangle$ とした。上から順に $(\text{sign}(J_1), \text{sign}(J_2)) = (+, +), (-, +), (+, -), (-, -)$ である。

相互作用 J_1 の符号が正（負）の場合、短時間で $\sigma_1^x(t)$ は負（正）になる。短時間極限で1番目のスピンに働く分子場は $(-h_1, 0, J_1)$ であることから、この振る舞いを説明することができる。 J_2 の符号の決定のためには、少なくとも $(1/|J_1| + 1/|J_2|)$ 程度の時間、 $\sigma_1^x(t)$ を測定する必要があるのは、直感的に明らかであろう。具体的には、 J_2 の符号は図3の4通りの計算結果と実験結果を比較することで決定できる。

また、図3より分かるように、 $\langle \sigma_1^x(t) \rangle$ は全ての相互作用の符号を反転すると正負が逆転する。これは以下のように理解できる。2番目のスピンを x 軸周りに π 回転させるユニタリ行列 $X_{2,\pi} = \exp(-i\sigma_2^x \pi/2)$ を導入すると、 $X_{2,\pi} \mathcal{H}(\{h_i\}, \{J_i\}) X_{2,\pi}^\dagger = \mathcal{H}(\{h_i\}, \{-J_i\})$ より、

$$\begin{aligned} & \left\langle \uparrow\uparrow\uparrow \left| e^{i\mathcal{H}(\{h_i\}, \{-J_i\})t} \sigma_1^x e^{-i\mathcal{H}(\{h_i\}, \{-J_i\})t} \right| \uparrow\uparrow\uparrow \right\rangle \\ &= \left\langle \uparrow\uparrow\uparrow \left| X_{2,\pi}^\dagger e^{i\mathcal{H}(\{h_i\}, \{J_i\})t} X_{2,\pi} \sigma_1^x X_{2,\pi}^\dagger e^{-i\mathcal{H}(\{h_i\}, \{-J_i\})t} X_{2,\pi} \right| \uparrow\uparrow\uparrow \right\rangle \\ &= - \left\langle \uparrow\uparrow\uparrow \left| e^{i\mathcal{H}(\{h_i\}, \{J_i\})t} \sigma_1^x e^{-i\mathcal{H}(\{h_i\}, \{J_i\})t} \right| \uparrow\uparrow\uparrow \right\rangle \end{aligned} \quad (26)$$

となるからである。

4 結論と今後の展望

本研究では、端にある量子ビットのみを操作し制御することで系全体のハミルトニアンのパラメータを推定する手法を考案した。我々はNMR量子コンピュータを想定して、3つの量子ビットから成る、サイト依存の相互作用と横磁場のある横磁場イジングモデルを例に取り上げた。横磁場イジングモデルには、ハイゼンベルクモデルや XXZ モデルのような良い量子数が存在しないので、それらに比べて相互作用の推定は格段に難しい。

我々の提案した相互作用の推定手法は以下のようにまとめられる。初期状態として、 $|\uparrow\uparrow\uparrow\rangle$ ($h_1 \neq 0$ の場合)、または $1/\sqrt{2}(|\uparrow\rangle + |\downarrow\rangle)|\uparrow\uparrow\rangle$ ($h_1 = 0$ の場合) を用意し、1番目の量子ビットの x 成分の時間発展 $\langle \sigma_1^x(t) \rangle$ を測定する。 $\langle \sigma_1^x(t) \rangle$ のフーリエ変換 $\hat{\sigma}_1^x(\omega)$ のピーク位置は、ハミルトニアン の正の固有値を ϵ_m, ϵ_n とすると、 $\{\epsilon_m \pm \epsilon_n\}$ に現れる。一般の有限の h_1 において、どのピークがどの $\{\epsilon_m \pm \epsilon_n\}$ に対応するかを同定するのは難しいが、 $h_1 = 0$ または $h_1 \rightarrow \infty$ においては、それが容易である。 $h_1 = 0$ から有限の h_1 まで連続的に $\hat{\sigma}_1^x(\omega)$ のピーク位置の変化を得ることができれば、それぞれのピークがどの $\{\epsilon_m \pm \epsilon_n\}$ に対応するかを特定することは可能である。また、 h_1 が大きい場合、ピークの高さが減少していくので、ピークを観測することができなくなる。そのことから、 $h_1 = 0$ から有限の h_1 まで連続的に $\hat{\sigma}_1^x(\omega)$ のピーク位置の変化を得ることは重要である。

以上のようにして、 $\hat{\sigma}_1^x(\omega)$ のピーク位置から $\{h_i\}, \{|J_i|\}$ を決定することができる。しかしながらこの段階では相互作用の符号を決定できない。1番目の量子ビットの x 成分の時間発展 $\langle \sigma_1^x(t) \rangle$ に関する全ての $\{|J_i|\}$ の符号の組み合わせに対する計算結果と、実験結果とを比較することで相互作用の符号を決めることが可能である。また、短い時間の $\langle \sigma_1^x(t) \rangle$ の振る舞いは J_1 の符号によるので、 J_1 の符号のみであれば4通りの計算結果を用いなくても決定することが可能である。

我々の提案する方法は、4つの条件のみを仮定している。これらは液体NMR量子コンピュータなどの多くの実験系で満たしているありふれた条件である。本研究で取り扱った、サイトに依存する横磁場イジングモデルはNMR量子コンピュータのハミルトニアンに対応している。それぞれの量子ビットのラーモア周波数を用いて、それぞれの量子ビットに対する回転系を考えると、各量子ビットにかかる横磁場を制御することは可能である。ただし、実際に本研究で展開した議論を相互作用推定に用いるには残された課題があることを以下に記す。本研究で採用した初期状

態は純粋状態であったが、NMR 量子コンピュータの場合、このような純粋状態を生成するためには、時間平均法または空間平均法を用いて擬純粋状態 (pseudo-pure state) を生成する必要がある [15–19]。擬純粋状態を生成するためにはハミルトニアンをすでに知っている必要がある。つまり、「卵が先か、鶏が先か」というジレンマに陥っているとも言える。本研究で提案した手法をデモンストレーションするためには、別の手法で量子ビット間の相互作用が分かっている系を用いて行う必要がある。

また本研究では、デコヒーレンスの効果についても考慮していない。これについては今後の課題である。

謝辞

Daniel Burgarth 博士、丸山耕司博士には有益なコメントを頂きました。また、Carlo Di Franco 博士には [10] を紹介していただきました。本研究は、文部科学省によるオープンリサーチセンター整備事業から支援を受けて実施されました。また田中は、科研費 No.21840021 及び No.22340111 から支援を受けました。本研究の予備計算に東京大学物性研究所のスーパーコンピュータを利用いたしました。基研研究会「量子科学における双対性とスケール」(YITP-W-10-14) の参加者との議論により、論文 [1] の研究を完成させることができました。この場を借りて感謝いたします。

参考文献

- [1] M. A. Fasihi, S. Tanaka, M. Nakahara, and Y. Kondo, *J. Phys. Soc. Jpn.* **80** (2011), 044002.
- [2] P. W. Shor, in *Proc. 35th Annual Symposium on Foundations of Computer Science* (IEEE Press, Los Alamitos, IL, 1994), 124.
- [3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).
- [4] M. Nakahara and T. Ohmi, *Quantum Computing From Linear Algebra to Physical Realizations* (CRC Press, Boca Raton, FL, 2008).
- [5] M. Mohseni and D.A. Lidar, *Phys. Rev. Lett.* **97** (2006), 170501.
- [6] R. Blume-Kohout, H.K. Ng, D. Poulin, and L. Viola, *Phys. Rev. Lett.* **100** (2008), 030501.
- [7] A. Bendersky, F. Pastawski, and J.P. Paz, *Phys. Rev. Lett.* **100** (2008), 190403.
- [8] D. Burgarth, K. Maruyama, and F. Nori, *Phys. Rev. A* **79** (2009), 020305.
- [9] D. Burgarth and K. Maruyama, *New. J. Phys.* **11** (2009), 103019.
- [10] C. Di Franco, M. Paternostro, and M. S. Kim, *Phys. Rev. Lett.* **102** (2009), 187203.

- [11] D. Burgarth, K. Maruyama, and F. Nori, *New. J. Phys.* **13** (2009), 013019.
- [12] Y. Shikano and S. Tanaka, arXiv:1007.5370.
- [13] Y. Shikano, S. Kagami, S. Tanaka, and A. Hosoya, to appear in *AIP Conference Proceedings*.
- [14] 近藤康、中原幹夫、谷村省吾、*日本物理学会誌*、**62** (2007), 365.
- [15] E. Knill, I. L. Chuang and R. Laflamme, *Phys. Rev. A* **81** (1998), 5672.
- [16] D. G. Cory, A. F. Fahmy and T. F. Havel, *Proc. Natl. Acad. Sci. USA* **94** (1997), 1634.
- [17] U. Sakaguchi, H. Ozawa and T. Fukumi, *Phys. Rev. A* **61** (2000), 042313.
- [18] N. A. Gershenfeld and I. L. Chuang, *Science* **275** (1997) 350.
- [19] Lieven M. K. Vandersypen, Costantino S. Yannoni, Mark H. Sherwood, and Isaac L. Chuang, *Phys. Rev. Lett.* **83** (1999), 3085.
- [20] P. D. Lax, *Linear Algebra and Its Applications*, (Wiley-Interscience, Hoboken, New Jersey, 2007).

量子意思決定論について

高知工科大学・理論物理学研究室 全卓樹 (Taksu Cheon)
Laboratory of Physics, Kochi University of Technology

はじめに

人間の意思決定における背理的な事象を「量子的」な確率を用いて記述しようとする試みが注目を集めている [1, 2, 3]。その代表例である「確実事象原則の破れの量子論」について、われわれの研究を中心に簡潔にまとめてみる。

心理学における確実事象原則の破れとは、20年ほど前に Tversky たち [4] によって発見されたもので、二つの与件 (c0)、(c1) それぞれのもとで、人間に二つの決定 (d0)、(d1) が可能な状況に関する、次のような現象である。すなわちもし与件 (c0) のもと (c1) のもとの両方にあつて、決定 (d1) が選択される確度が高いとすれば、与件が (c0) であるか (c1) であるか不明な場合も、「合理的」な判断では当然決定 (d1) が選択される確度が高いと期待されるが、ある種の状況では人間はそうでない選択をすることが実験的に示されるのである。

簡単な仮想例をあげると、(c0) 天気が悪い場合と (c1) 天気がよい場合で、登山クラブのメンバーが登山を決行 (d1) する確率が、それぞれ6割、8割とする。天気がよいか悪いか不明で半々の場合、単純に考えると登山決行 (d1) の確率は中間の7割になりそうであるが、実際にアンケートで調べると4割になる、といった具合である。

この例でわかるのは、不確定な事象に対する人間の反応は、必ずしも加重平均という単純な意味の「合理的」判断によつてはならない事であり、意思決定には安全を重視した「防御的」なバイアスがかかるような何らかのメカニズムが働いていると考えられる。その起源については、それが進化的に優位だったのだろうという推測ができるのみであり、そのメカニズムの詳細については、神経回路の解剖学的実験と理論的モデルの両面の研究の進展待つか無いだろう。おそらく現状で望み得るのは、確実事象原則の破れを記述する有効な現象論を確立し、それを性格分類や行動予測といった応用に供する事であろう。そのような要請に応えうるかもしれない一つの試みが、まさに以下に素描する量子意思決定論である。

相加平均、相乗平均、量子平均

選択肢二つの与件それぞれに対して、被験者が選択肢二つから決定を下せる心理学実験を想定する。与件 (c0) が与えられたとき被験者が決定 (d0)、(d1) を行う確率をそれぞれ $p(d0; c0)$ 、 $p(d1; c0)$ と書いてみる。さらに与件 (c1) のもとで決定 (d0)、(d1) を行う確率をそれぞれ $p(d0; c1)$ 、 $p(d1; c1)$ と書こう。これらの確率は当然 $p(d0; c0) + p(d1; c0) = 1$ 、 $p(d0; c1) + p(d1; c1) = 1$ という関係を満たす。いま与件 (c0)、(c1) の生起する確率がそれぞれ

れ P_{c0} 、 P_{c1} だとしよう。当然 $P_{c0} + P_{c1} = 1$ である。この状況で決定 (d0) がなされる確率 Q_{d0} 、そして決定 (d1) がなされる確率 Q_{d1} は

$$\begin{aligned} Q_{d0} &= p(d0; c0)P_{c0} + p(d0; c1)P_{c1}, \\ Q_{d1} &= p(d1; c0)P_{c0} + p(d1; c1)P_{c1}, \end{aligned} \quad (1)$$

と考えるのが、事象の確率の線形性という合理的予測を表している。この両確率の和は当然ながら上の条件から $Q_{d0} + Q_{d1} = 1$ となっている。これをここでは「古典的」結果とも呼ぶ事にする。

この式の一つの帰結は不等式

$$\begin{aligned} \text{Max} [p(d0; c0), p(d0; c1)] &\geq Q_{d0} \geq \text{Min} [p(d0; c0), p(d0; c1)], \\ \text{Max} [p(d1; c0), p(d1; c1)] &\geq Q_{d1} \geq \text{Min} [p(d1; c0), p(d1; c1)], \end{aligned} \quad (2)$$

で、これが「確実事象原則」の式による表現になっている。

生物の刺激に対する反応は一般に、刺激の強度に線形に応答するのではなく、強度の log に対して線形の応答となる事が多い。与件に対する判断に於いても、危急の場合大脳皮質に達しないで判断が行われて、そのような log に対して線形の応答があると推測する事は可能である。この場合二つの与件への平均的な応答は、相加平均でなく相乗平均

$$\begin{aligned} 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}}, \\ 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}, \end{aligned} \quad (3)$$

が関与すると考えることができる。実際の人間の判断はこのような「脊椎反応」(3) によって「大脳皮質の判断」(1) が補正されると推測してみよう。そうすると不定与件での決定 (c0)、(c1) の確率が二つの実数パラメータ α 、 β をもった式

$$\begin{aligned} Q_{d0} &= p(d0; c0)P_{c0} + p(d0; c1)P_{c1} + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}}\alpha, \\ Q_{d1} &= p(d1; c0)P_{c0} + p(d1; c1)P_{c1} + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}\beta, \end{aligned} \quad (4)$$

で与えられると考えることができる。ここで相加平均と相乗平均の大小関係

$$\begin{aligned} p(d0; c0)P_{c0} + p(d0; c1)P_{c1} &\geq 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}}, \\ p(d1; c0)P_{c0} + p(d1; c1)P_{c1} &\geq 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}, \end{aligned} \quad (5)$$

と当然の要請 $1 \geq Q_{d0} \geq 0$ 、 $1 \geq Q_{d1} \geq 0$ からパラメータの範囲 $1 \geq \alpha \geq -1$ 、 $1 \geq \beta \geq -1$ が定まる。そしてさらに基本的要請 $Q_{d0} + Q_{d1} = 1$ から二つのパラメータの間には

$$\sqrt{p(d0; c0)p(d0; c1)}\alpha = -\sqrt{p(d1; c0)p(d1; c1)}\beta, \quad (6)$$

という制限条件がかかることになり、 α と β は独立ではなく (4) は使いやすい表式とならない。いまこれが一つのパラメータだけに陽に依存する形にかける特別な二つの状況を考えることができる。それは $p(d0; c0) = p(d0; c1)$ 、 $p(d1; c0) = p(d1; c1)$ のとき、または $p(d0; c0) =$

$p(d1; c1)$ 、 $p(d1; c0) = p(d0; c1)$ ので、 $\alpha = -\beta$ ととれば $Q_{d0} + Q_{d1} = 1$ がうまく満たされ、不定与件下での確率は

$$\begin{aligned} Q_{d0} &= p(d0; c0)P_{c0} + p(d0; c1)P_{c1} + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \cos \theta, \\ Q_{d1} &= p(d1; c0)P_{c0} + p(d1; c1)P_{c1} - 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \cos \theta, \end{aligned} \quad (7)$$

と対称な形に書けるのである。この式では $\alpha = \cos \theta$ という円環的パラメータ θ による表記替えをおこなっている。相加平均、相乗平均の両方が出てくるこの (7) 式は、様々な文脈で「量子的干渉」のある非線形な確率和としてよく出てくるものである。 $\cos \theta = 0$ で第二項は消失するので、これが古典的な場合に対応している。第二項の存在のために、確実事象原則を表す不等式 (2) は一般には破られる。そのかわりに不等式

$$\begin{aligned} \sqrt{p(d0; c0)P_{c0} + p(d0; c1)P_{c1}} &\geq Q_{d0} \geq \sqrt{|p(d0; c0)P_{c0} - p(d0; c1)P_{c1}|}, \\ \sqrt{p(d1; c0)P_{c0} + p(d1; c1)P_{c1}} &\geq Q_{d1} \geq \sqrt{|p(d1; c0)P_{c0} - p(d1; c1)P_{c1}|}, \end{aligned} \quad (8)$$

が成り立つ事に注意しよう。式 (7) を用いて確実事象原則を破った心理学的実験を解析した例がいくつかあり、一定の現象論的成功を収めている。

一般の場合に適用できる量子平均

前説の (7) 式はわかりやすい形をしているが、辻褄を合わせるためには条件確率についての特別な仮定 $p(d0; c0) = p(d0; c1)$ または $p(d0; c0) = p(d1; c1)$ を必要とし、これはもちろん一般の条件に適用しようとするとき必ず無理が生じる。そこで (4) に立ち帰って、この右辺は Q_{d0} と Q_{d1} の比を与えているのだと解釈し直してみる。すなわち規格化定数 N を導入して

$$\begin{aligned} Q_{d0} &= N \left(p(d0; c0)P_{c0} + p(d0; c1)P_{c1} + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \alpha \right), \\ Q_{d1} &= N \left(p(d1; c0)P_{c0} + p(d1; c1)P_{c1} + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \beta \right), \end{aligned} \quad (9)$$

としてみる。確率の基本要請 $1 \geq Q_{d0} \geq 0$ 、 $1 \geq Q_{d1} \geq 0$ 、そして $Q_{d0} + Q_{d1} = 1$ をおくと、前の二つからは前節同様にパラメータの範囲 $1 \geq \alpha \geq -1$ 、 $1 \geq \beta \geq -1$ が定まり、三つ目から N が定まる。二つの円環パラメータ $\theta_0 = \arccos \alpha$ 、 $\theta_1 = \arccos \beta$ を導入すると結果は

$$\begin{aligned} Q_{d0} &= \frac{p(d0; c0)P_{c0} + p(d0; c1)P_{c1} + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \cos \theta_0}{1 + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \cos \theta_0 + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \cos \theta_1}, \\ Q_{d1} &= \frac{p(d1; c0)P_{c0} + p(d1; c1)P_{c1} + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \cos \theta_1}{1 + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \cos \theta_0 + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \cos \theta_1}, \end{aligned} \quad (10)$$

となる。この特殊な場合、すなわち $\sqrt{p(d0; c0)p(d0; c1)} \cos \theta_0 = -\sqrt{p(d1; c0)p(d1; c1)} \cos \theta_1$ が成立する場合に、(10) が (7) に帰着する事は一見してわかる。不定与件下での決定の確

率が相加平均と相乗平均の線形結合で与えられるとすると、必然的に二つの円環パラメータが必要であるという（結果を見れば当然な）結論になる。

一般の場合にも辻褄の合った結果を与える公式 (10) は、まだ例の数が少なく限定的ながら、実験結果の解析によって、その現象論的な有効性が確かめられている [3]。

この「一般化された量子平均」の式から、(8) の代わりにさらに緩い不等式

$$\begin{aligned}
 & \frac{\left(\sqrt{p(d0; c0)P_{c0}} + \sqrt{p(d0; c1)P_{c1}}\right)^2}{1 - \sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} - \sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}} \geq Q_{d0} \\
 & \geq \frac{\left(\sqrt{p(d0; c0)P_{c0}} - \sqrt{p(d0; c1)P_{c1}}\right)^2}{1 + \sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} + \sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}}, \\
 & \frac{\left(\sqrt{p(d1; c0)P_{c0}} + \sqrt{p(d1; c1)P_{c1}}\right)^2}{1 - \sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} - \sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}} \geq Q_{d1} \\
 & \geq \frac{\left(\sqrt{p(d1; c0)P_{c0}} - \sqrt{p(d1; c1)P_{c1}}\right)^2}{1 + \sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} + \sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}}}, \quad (11)
 \end{aligned}$$

が得られる。

量子平均の導出

ここまでの議論をみると、与件が不定な場合の被験者の決定を表す確率 (10) を a) 相加平均と相乗平均の線形和である、b) 確率論として辻褄が合ってる、という条件だけをたよりに推測によって導いただけのようにも見える。背後に神経回路のモデルがある訳ではない現状は、結局のところそういう段階だと言わざるを得ない。しかしこれを、このように完全にアド・ホックにではなく、通常確率論とは異なったヒルベルト・ベクトルによる確率記述で与えられる、という単一の仮定だけから導く事も可能である。式 (10)、そしてその特殊な限定版である式 (7) を「量子的」な表式と呼んだ理由も、それによって初めて明らかになる。

「なぜ意思決定の考察に量子論か？」という読者の疑問は当然で、有効な現象論的關係式 (10) が最初にそのように導かれたという経緯 [3] だけでは、もちろん答えになっていない。しかし量子論的確率は、古典的確率論を拡張する数学のなかで、いまのところ有効性がはっきり知られた唯一のものである。それは選択肢二つの二段階意思決定というここで扱われた問題を、選択肢数や段階数を一般の場合に拡張するための健全な出発点を与える。さらにこのように作り上げられた枠組みは意思決定の心理学以外にも適用対象を見つけるかもしれない。そして最後に贅言を弄すれば、人間の思考過程に決して量子論が関与して無いとする確たる証拠の無い現状では、そのような魅力的な可能性を最初から排除する必要は無いであろう。

今被験者の決定の選択肢を次元 2 のヒルベルト空間のベクトルの二つの基底 $|0\rangle$ と $|1\rangle$ で表してみる。さらに決定に先立つ与件も、これとは別の次元 2 のヒルベルト空間のベクトル

の二つの基底 $|0\rangle$ と $|1\rangle$ で表されると考えてみよう。与件が与えられる前の被験者の状態は次元 2×2 の直積ヒルベルト空間のベクトルでつぎのように表される。

$$|\Psi\rangle = |0\rangle |\psi^{(0)}\rangle + |1\rangle |\psi^{(1)}\rangle, \quad (12)$$

ここで決定を表すベクトルは

$$\begin{aligned} |\psi^{(0)}\rangle &= \psi_0^{(0)} |0\rangle + \psi_1^{(0)} |1\rangle, \\ |\psi^{(1)}\rangle &= \psi_0^{(1)} |0\rangle + \psi_1^{(1)} |1\rangle, \end{aligned} \quad (13)$$

であって、このなかの4つの量 $\psi_j^{(k)}$ 、($k = 0, 1$ および $j = 0, 1$) は複素数であって、与件 (ck) のもとでの決定 (dj) が行われる確率は $p(dj; ck) = |\psi_j^{(k)}|^2$ で与えられる。それゆえ当然 $k = 0, 1$ について $\psi_0^{(k)*} \psi_0^{(k)} + \psi_1^{(k)*} \psi_1^{(k)} = 1$ が要求される。

与件が $(c0)$ 、 $(c1)$ のどちらかに確定しない中間的確率的事象を

$$|K\rangle = \kappa^{(0)} |0\rangle + \kappa^{(1)} |1\rangle, \quad (14)$$

で表してみる。ここで $\kappa^{(k)}$ は条件 $\kappa^{(0)*} \kappa^{(0)} + \kappa^{(1)*} \kappa^{(1)} = 1$ を満たす複素数で、この絶対値の二乗が与件 (ck) が起きる確率 $P_{ck} = |\kappa^{(k)}|^2 = \langle k|K\rangle \langle K|k\rangle$ だと考えるのである。中間的な与件 $|K\rangle$ が与えられたという事態を、仮想的な観測者による観測行為の結果、任意の与件状態 $|\phi\rangle$ が状態 $|K\rangle$ に変化したものだとして想定してみる。この過程は非ユニタリー射影演算子

$$\mathcal{K}_\phi = \frac{1}{\sqrt{\langle \phi|K\rangle \langle K|\phi\rangle}} |K\rangle \langle K|, \quad (15)$$

による射影操作 $|\phi\rangle \rightarrow \mathcal{K}_\phi |\phi\rangle$ で記述される。与件と決定の直積空間の状態 $|\Psi\rangle$ に対しては、これを直積空間への射影に拡張したもの

$$\mathcal{K} = \frac{1}{\sqrt{\langle \Psi||K\rangle \langle K||\Psi\rangle}} |K\rangle \langle K|, \quad (16)$$

を考える。この \mathcal{K} は与件の状態の部分的観測に相当し、状態を $|\Psi\rangle \rightarrow |\Psi'\rangle = \mathcal{K} |\Psi\rangle$ と変化させる。すなわち

$$|\Psi'\rangle = \frac{1}{\sqrt{\langle \Psi||K\rangle \langle K||\Psi\rangle}} |K\rangle \langle K||\Psi\rangle, \quad (17)$$

であり、ここに登場する部分的行列要素 $\langle K||\Psi\rangle$ は

$$\begin{aligned} \langle K||\Psi\rangle &= \langle 0|(K||\Psi)\rangle |0\rangle + \langle 1|(K||\Psi)\rangle |1\rangle \\ &= (\kappa^{(0)*} \psi_0^{(0)} + \kappa^{(1)*} \psi_0^{(1)}) |0\rangle + (\kappa^{(0)*} \psi_1^{(0)} + \kappa^{(1)*} \psi_1^{(1)}) |1\rangle, \end{aligned} \quad (18)$$

で与えられる。この $|\Psi'\rangle$ の中に含まれる $|K\rangle|j\rangle$ の前の係数の絶対値の二乗が、中間与件 $|K\rangle$ の下での被験者の決定 (dj) の確率 Q_{dj} を表している。つまり

$$Q_{dj} = \langle j|(K|\Psi')\rangle = \frac{|\langle j|(K|\Psi)\rangle|^2}{|\langle 0|(K|\Psi)\rangle|^2 + |\langle 1|(K|\Psi)\rangle|^2}, \quad (19)$$

となり、これを書き換えると

$$Q_{dj} = \frac{p(dj; c0)P_{c0} + p(dj; c1)P_{c1} + 2\sqrt{p(dj; c0)P_{c0}p(dj; c1)P_{c1}} \cos \theta_j}{1 + 2\sqrt{p(d0; c0)P_{c0}p(d0; c1)P_{c1}} \cos \theta_0 + 2\sqrt{p(d1; c0)P_{c0}p(d1; c1)P_{c1}} \cos \theta_1}, \quad (20)$$

となって、前節の (10) が導かれるのである。ここで θ_j ($j = 0, 1$) は

$$\kappa^{(0)}\kappa^{(1)*}\psi_j^{(0)*}\psi_j^{(1)} = e^{i\theta_j}\sqrt{p(dj; c0)P_{c0}p(dj; c1)P_{c1}}, \quad (21)$$

によって定義される。前節で被験者の心理的特性を記述するものとして導入されたこの現象論的円環パラメータは、ここでの導出では量子位相に他ならないことが判明した訳である。そしてこれは $\cos \theta_j$ の入った「相乗平均」の項が量子干渉項に他ならないことを意味するのである。

考察

我々の中間事象の射影演算子 (16) による取り扱い、量子観測理論の言語では Lüders 射影公準 [5] を採用した事に相当する。もし同一の初期直積状態 $|\Psi\rangle$ のアンサンブルを考え、それに対して同一の中間的与件状態 $|K\rangle$ による測定を繰り返し行う、という操作を想定すれば、Lüders 射影公準は純粋状態 $\rho = |K\rangle\langle K|$ を用いた通常の von Neumann 射影公準 [6] に帰着する。仮にこの純粋状態を、同一の $\kappa^{(0)}$ と $\kappa^{(1)}$ をもつが位相がランダムな多くのベクトルからなる混合状態で置き換えるとすれば、(20) 中の干渉項は打ち消し合ってしまうであろう [7]。その場合は von Neumann 射影公準は古典的確率公式 (1) に移行するのである。

References

- [1] E. M. Pothos and J. R. Busemeyer, *A quantum probability explanation for violations of 'rational' decision theory*, Proc. Roy. Soc. B **276** (2009) 2171.
- [2] A. Yu. Khrennikov and E. Haven, *Quantum mechanics and violations of the sure-thing principle: The use of probability interference and other concepts*, J. Math. Psychol. **53** (2009) 378.
- [3] T. Cheon and T. Takahashi, *Interference and inequality in quantum decision theory*, Phys. Lett. **A375** (2010) 100.

- [4] A. Tversky and E. Shafir, *The disjunction effect in choice under uncertainty*, Psychol. Sci. **3** (1992) 30509.
- [5] G. Lüders, *Über die Zustandsänderung durch den Messprozess*, Ann. Phys. (Leipzig) **8** (1951), 322.
- [6] J. von Neumann, "Mathematical foundations of quantum mechanics", (Princeton Univ. Press, Princeton, N.J., 1955).
- [7] N. Gisin, *Quantum measurements and stochastic process*, Phys. Rev. Lett. **52** (1984), 1657.

Quantum algorithm for EXPTIME problem and Computational Complexity

S.Iriyama and M.Ohya

Department of Information Sciences, Tokyo University of Science

Abstract

We have studied a computational complexity of quantum algorithm for several years. A mathematical model of quantum algorithm, which is called a generalized quantum Turing machine(GQTM), is introduced by Ohya et al., and we discussed the relationships of language classes defined by it.

In this paper, we construct a quantum algorithm for the Pebble Game problem, which belongs to the class EXPTIME(i.e., it requires an exponential size of resources to solve), and discuss the computational complexity of it.

1 Introduction

We have studied on quantum algorithm for several years. Ohya, and Volovich discovered the quantum algorithm with the amplification process which is described chaos dynamics, and it is called the OMV quantum algorithm which can solve NP complete problem in polynomial time[1, 2, 3, 4]. We applied this algorithm to the other problems, for example, multiple alignment of amino acid sequence, Hamilton closed path problem, and protein folding problem[4]. Therefore we found that OMV quantum algorithm is useful for many problems in several fields to search the objects which satisfy the given conditions.

In this paper, we review mathematical foundations of quantum algorithm. A generalized quantum Turing machine is introduced by Ohya et al., and we can discuss the computational complexity of quantum algorithm rigorously by using it.

Then we show a quantum algorithm for EXPTIME problem, Pebble Game, and discuss the computational complexity of it.

2 Quantum Algorithm

A quantum algorithm is constructed by the following steps:

1. Prepare a Hilbert space
2. Construct an initial state
3. Construct unitary operators to solve the problem
4. Apply them for the initial state and obtain a result state
5. If necessary, amplify the probability of correct result
6. Measure an observable with the result state

In the first step, we define the Hilbert space depending on the problem. Let \mathbb{C}^2 be a Hilbert space spanned by $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, a normalized vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ on this space is called a qubit. Since we can use a superposition of $|0\rangle$ and $|1\rangle$ as an initial state vector, the quantum algorithm is more effective than classical one.

One can apply Hadamard transformation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

to create a superposition. For $|0\rangle$ and $|1\rangle$, it works as

$$\begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle \\ H|1\rangle &= \frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle. \end{aligned}$$

Hadamard transformation has a very important role in a quantum algorithm.

Here we introduce logical gates, which are NOT gate, C-NOT gate and CC-NOT gate. We call these gates fundamental gates. We can also construct AND and OR gate by considering the product of fundamental gates and some implementations. The NOT gate U_{NOT} is defined on a Hilbert space \mathbb{C}^2 as

$$U_{NOT} = |1\rangle\langle 0| + |0\rangle\langle 1|.$$

It works for an arbitrary qubit as

$$U_{NOT} (\alpha |0\rangle + \beta |1\rangle) = \alpha |1\rangle + \beta |0\rangle$$

C-NOT U_{CN} gate and CC-NOT U_{CCN} are given on two and three qubit Hilbert space as

$$\begin{aligned} U_{CN} &= |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U_{NOT} \\ U_{CCN} &= |0\rangle\langle 0| \otimes I \otimes I + |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes |1\rangle\langle 1| \otimes U_{NOT}, \end{aligned}$$

respectively. The unitary operator to solve the problem is constructed by these fundamental gates. Let Q be a set of states including an initial state q_0 and a set of final states $\{q_F\}$ and Σ a set of alphabets including a blank symbol $\#$. We denote a set of all infinite sequences of alphabets in Σ as

$$\Sigma^* = \Sigma \times \Sigma \times \dots = \Sigma^\infty.$$

Let $A \in \Sigma^*$ be a sequence of alphabets representing the tape state. Each tape alphabet has a unique index. We write the k -th tape alphabet a_k as $A(k)$.

GQTM

The GQTM M_{gq} is defined by the quadruplet $(Q, \Sigma, \mathcal{H}, \Lambda_\delta)$, where Λ_δ is a quantum transition function from $\mathcal{C} = Q \times \Sigma \times \mathbb{Z}$ to \mathcal{C} . Q and Σ are represented by a density operator on Hilbert spaces \mathcal{H}_Q and \mathcal{H}_Σ , which are spanned by canonical basis $\{|q\rangle; q \in Q\}$ and $\{|a\rangle; a \in \Sigma\}$, respectively. A tape configuration A is a sequence of elements of Σ represented by a density operator on the Hilbert space \mathcal{H}_Σ spanned by a canonical basis $\{|A\rangle; A \in \Sigma^*\}$. A position of tape head is represented by a density operator on the Hilbert space \mathcal{H}_Z spanned by a canonical basis $\{|i\rangle; i \in \mathbb{Z}\}$. Then a configuration of GQTM M_{gq} is described by a density operator ρ in $\mathcal{H} \equiv \mathcal{H}_Q \otimes \mathcal{H}_\Sigma \otimes \mathcal{H}_Z$. Let $\mathfrak{S}(\mathcal{H})$ be the set of all density operators in the Hilbert space \mathcal{H} .

Here, we define the transition function

$$\delta_1 : \mathbb{R} \times Q \times \Sigma \times Q \times \Sigma \times Q \times \Sigma \times \{0, \pm 1\} \times Q \times \Sigma \times \{0, \pm 1\} \rightarrow \mathbb{C}.$$

A quantum transition function is given by a quantum channel

$$\Lambda_{\delta_1} : \mathfrak{S}(\mathcal{H}) \rightarrow \mathfrak{S}(\mathcal{H}),$$

satisfying the following condition.

Definition 1 Λ_{δ_1} is called a quantum transition channel if there exists a transition function δ_1 such that for any quantum configuration $\rho = \sum_k \lambda_k |\psi_k\rangle \langle \psi_k|$, $|\psi_k\rangle = \sum_l \alpha_{k,l} |q_{k,l}, A_{k,l}, i_{k,l}\rangle$, $\sum_k \lambda_k = 1, \forall \lambda_k \geq 0$, $\sum_l |\alpha_{k,l}|^2 = 1, \forall \alpha_{k,l} \in \mathbb{C}$ it holds

$$\Lambda_{\delta_1}(\rho) \equiv \sum_{k,l,m,n,p,b,d,p',b',d'} \delta_1(\lambda_k, q_{k,l}, A_{k,l}(i_{k,l}), q_{m,n}, A_{m,n}(i_{m,n}), p, b, d, p', b', d') \\ \times |p, B, i_{k,l} + d\rangle \langle p', B', i_{m,n} + d'|$$

$$B(j) = \begin{cases} b & j = i_{k,l} \\ A_{k,l}(j) & \text{otherwise} \end{cases} \\ B'(j) = \begin{cases} b' & j = i_{m,n} \\ A_{m,n}(j) & \text{otherwise} \end{cases}$$

so that the RHS is a state.

We defined a LQTM(Linear Quantum Turing Machine) and unitary QTM (UQTM), see [10, 11]. Moreover, we can define a classical Turing machine as follows.

Definition 2 We call $M = (Q, \Sigma, \mathcal{H}, \Lambda_\delta)$ a classical Turing machine(CTM) iff Λ_{δ_c} is well defined on a subset S of state space \mathfrak{S} such that

$$S = \left\{ \rho; \rho \in \mathfrak{S}, \rho = |\psi\rangle \langle \psi|, |\psi\rangle = \sum_k \alpha_k |q_k, A_k, i_k\rangle, \sum_l |\alpha_k|^2 = 1, \forall \alpha_k \in \mathbb{C} \right\}$$

and δ_c satisfies the Turing machine conditions as follows.

- $\delta((q, a, i), (p, b, i + d)) = 0$ if $|d| > 1$ for all q, p, a, b, i .
- $\sum_{p,b,d} \delta((q, a, i), (p, b, i + d)) = 1$

2.1 Computation Process of GQTM

Let $M = (Q, \Sigma, \mathcal{H}, \Lambda_\delta)$ and $\rho_0 = |\psi_0\rangle \langle \psi_0|$ where $|\psi_0\rangle = |q_0, A, 0\rangle$, we call this state an initial state and A an input of M . The computation of GQTM proceeds,

applying Λ_δ to ρ_0 , till the processor state becomes $q_f \in \{q_F\}$, then it halts. This process is described by the products of Λ_δ as $\Lambda_\delta \circ \dots \circ \Lambda_\delta(\rho_0) = \rho_f$

Let $\mathfrak{S}_f(\mathcal{H}) = \{\rho | \rho \upharpoonright \mathfrak{S}(\mathcal{H}_Q) \in \{q_f\}\}$ be the set of final configurations, where \upharpoonright means the restriction of a state on $\mathfrak{S}(\mathcal{H}_Q)$. ρ is called a final state if ρ is in the form

$$\rho = \sum_k \lambda_k \rho_k + \sum_l \mu_l \sigma_l$$

$$\sum_k \lambda_k + \sum_l \mu_l = 1, \quad \forall \lambda_k, \mu_l \geq 0$$

where $\sigma_l \in \mathfrak{S}_f(\mathcal{H})$. We call $p = \sum_l \mu_l$ the halting probability.

3 Language Classes

There exist several classical language classes defined by a deterministic Turing machine.

Definition 3 Let M be a deterministic Turing machine such that halts for all input. For a length n of input, let $f(n)$ be a maximum length of tape cell of M . We define a space complexity of M as f .

Definition 4 *PSPACE* is the language class which is recognized by a deterministic Turing machine in a polynomial space.

Definition 5 *NPSPACE* is the language class which is recognized by a non-deterministic Turing machine in a polynomial space.

Definition 6 *EXPTIME* is the language class which is recognized by a deterministic Turing machine in an exponential time.

The following relation is known:

$$P \subseteq NP \subseteq PSPACE = NPSPACE \subseteq EXPTIME$$

4 Pebble Game

Pebble Game is the two players game using a play board and stones(pebbles). Players move one stone at a time along a given rule alternatively. A player wins when he moves a stone to the winning position on the board given by the rule,

or he loses when he cannot move any stones. We want to know whether there exist strategies such that a first mover can win every time. The computational complexity of this problem belongs to EXPTIME. Then we propose a quantum algorithm for the game.

4.1 Representation of a Game

Let \mathcal{I} be a set of players, \mathcal{M} a set of moves(or strategies) available to those players, and \mathcal{P} a set of payments for each combination of moves. A game G is given a triplet $(\mathcal{I}, \mathcal{M}, \mathcal{P})$. The set of moves is given by a rule of the game. Players choice their move in their turn alternatively, so then these choices are described by a sequence of moves. We denote this by a position p_i where a player $i \in \mathcal{I}$ has the move. A set of ordered pair (p_i, p_j) where p_i and p_j are positions such that $p_i \rightarrow p_j$ is a feasible move implies the rule of the game.

The game is finished when there does not exist any moves in someone's turn. When the game is finished, the payments are given to each players by a function of the position. In this study we assume that the game must be finished so that the length of position is finite.

We say a *player A wins* if the payments of A is greater than a player B. Then we consider the following problem:

Problem 7 *Does there exist a way such that a player A wins independently of how a player B plays?*

To answer this, a *winning position* of A was introduced by König.as a set of positions where A wins in finite moves in spite of moves of B.

A set of winning positions of A is constructed inductively from a set of positions where A wins by only one move. Let p is a winning position, q is also a winning position if there exists a move r_A of A such that for all moves r_B of B, it holds

$$q r_A r_B = p$$

Therefore, the Problem7 is equivalent as the following:

Problem 8 *Determine whether an initial position of the game is in a winning position of A?*

4.2 Pebble Game

Let n be a positive finite integer denoted by a size of Pebble Game. n -Pebble Game is given by a triplet $G_{Pebble} = (\mathcal{I}, \mathcal{M}, \mathcal{P})$ where $\mathcal{I} = \{A, B\}$, $\mathcal{M} =$

$\left\{ r = (r_1, r_2, r_3) \in \{1, \dots, n\}^3; r \text{ is an available move} \right\}$ and $\mathcal{P} = \{f : \mathcal{M}^k \rightarrow \{0, 1\}\}$.
 \mathcal{M} is a set of available moves depending on a situation of the board and positions of stones.

The rule of Pebble Game is the following:

- Prepare n lattices denoted by a board and put $m (< n)$ stones on nodes. Each nodes has a unique number $i = 1, \dots, n$.
- Players move one of stones alternatively if the stone so chosen can jump the neighbor stone.
- If a player puts a stone on the node n , he wins. If a player can not move any stones, he loses.

Let $x = (x_1, x_2, \dots, x_n)$ be a vector of Boolean variables denoted by a situation of board, where x_i has one to one correspondence with a node i :

$$x_i = \begin{cases} 0 & \text{no stone on a node } i \\ 1 & \text{node } i \text{ has a stone} \end{cases}$$

We prepare m stones on the board x_S arbitrary and call it an initial situation. And we call a finished situation x_f if there are no available moves.

A move r_i ($i \in \{A, B\}$) for a player i is represented by $(r_1, r_2, r_3) \in \{1, \dots, n\}^3$ where r_1 indicates a position of stone, r_2 the neighbor and r_3 a position of destination. If there is a stone on r_1 and r_2 and there is not a stone on r_3 , the move $r_i = (r_1, r_2, r_3)$ is available.

After one move, the situation of board is changed by

$$x_i = \begin{cases} \bar{x}_i & i = r_1 \text{ or } i = r_3 \\ x_i & \text{otherwise} \end{cases}$$

we denote $r_i(x)$ by a situation of board after a move r_i .

If there is no moves available or he can move a stone to the node n , the game is finished. Then we have a sequence of moves of each players. For a sequence of moves p_i ($i \in \{A, B\}$) = $(r_A^1, r_B^2, \dots, r_i^k)$ a payment for a player A is given by a function $f_A \in \mathcal{P}$:

$$f_A(p_i) = \begin{cases} 0 & i = B \\ 1 & i = A \end{cases}$$

and for a player B is by

$$f_B \in \mathcal{P} : f_B(p_i) = \begin{cases} 0 & i = A \\ 1 & i = B \end{cases}$$

The Pebble Game problem belongs to EXPTIME if m is not fixed[8]. We check whether an initial situation is a winning position of A for all number m of stones.

Problem 9 For all finished situations x_f determine whether there exists k such that

$$\exists r_A^k \forall r_B^{k-1} \dots \exists r_A^3 \forall r_B^2 \exists r_A^1 (x_S) = x_f$$

where x_S is an initial situation.

4.3 Computational Complexity of a Classical Algorithm for Pebble Game

In this study, we assume the following Oracle M_O :

$$M_O : \{x; x \text{ is a situation}\} \rightarrow \mathbb{N} \cup \{0\}$$

For a finished situation x_f , M_O outputs the time of it immediately, just one step. If a situation x is not a final situation, M_O outputs 0. Then a classical algorithm solving the problem is the following.

Step1 For all situations, we run the following steps.

Step2 Calculate time $k(x)$ for a situation x

Step3 If x is a final situation, construct a set W_i ($i = 1, 2, \dots, k(x)$) of winning situations:

$$W_i = \{y; \exists r'_A, \forall r_B, \exists r_A, r'_A r_B r_A y \in W_{i-1}\}$$

where $W_0 = \{x\}$

Step4 Check whether x_S belongs to $W_{k(x)}$. If no, go back to Step2.

The number of all situations 2^n . The upper bound of number of available moves is $4n$. Therefore the computational complexity of a classical algorithm $T_C(n)$ is

$$T_C(n) \sim (4n)^3 \times 2^n \sim \exp(n)$$

Even if we assume the Oracle, this problem belongs to EXPTIME.

5 Quantum Algorithm for Pebble Game

As we assume the Oracle M_O , we use a quantum Oracle U_{M_O} which works as same as M_O . Here, we construct the following quantum algorithm for Problem3:

Step1 Create a superposition of all situations.

Step2 For the superposition, apply Oracle U_{M_O} .

Step3 We construct W_i ($i = 1, 2, \dots, k(x_f)$) for the superposition.

Step4 If $x_S \in W_{k(x)}$, make the final qubit $|1\rangle$.

All situations are represented by a binary form, so then we can create a superposition of them using Hadamard transformation. Step3 is achieved by a product of unitary gates. In Step4, AND operation is constructed by unitary gates[9].

If final qubit of superposition is $|1\rangle$, there exists a way of winning. Using the chaos amplifier, we obtain the result.

6 Computational Complexity

Here, we define the computational complexity of the quantum algorithm as the total number of fundamental gates. Step1 is constructed by n Hadamard gates. Step2 contains only 1 Oracle. Step3 has the same complexity as the classical one. And Step4 requires n gates for taking AND calculation. The upper bound of $|W_{k(x)}|$ is $\binom{n}{m}$ where m is the number of pebbles.

Therefore, computational complexity $T_Q(n)$ of quantum algorithm of Pebble Game is given by

$$T_Q(n) \sim \left\{ n + 1 + (4n)^3 + \binom{n}{m} \right\} \times \left[\frac{5}{4}(n-1) \right] \sim poly(n)$$

7 Conclusion

The computational complexity $T_C(n)$ of a classical algorithm for n -Pebble Game with Oracle is

$$T_C(n) \sim (4n)^3 \times 2^n \sim \exp(n)$$

This is an exponential time of input size n . We constructed a quantum algorithm for this, the computational complexity $T_Q(n)$ is

$$T_Q(n) \sim \left\{ n + 1 + (4n)^3 + \binom{n}{m} \right\} \times \left[\frac{5}{4} (n-1) \right] \sim \text{poly}(n)$$

Using this quantum algorithm, we discuss the computational complexity of quantum algorithm of Protein Folding problem in the sequel paper.

References

- [1] M.Ohya and I.V.Volovich, *Quantum computing and chaotic amplification*, J. opt. B, **5**,No.6 639-642, 2003.
- [2] M.Ohya and I.V.Volovich, *New quantum algorithm for studying NP-complete problems*, Rep.Math.Phys., **52**, No.1,25-33 2003.
- [3] M.Ohya and I.V.Volovich, *Mathematical Foundation of Quantum Computers, Teleportations and Cryptography*, Springer, 2011
- [4] M.Ohya and N.Masuda, *NP problem in Quantum Algorithm*, Open Systems and Information Dynamics, 7 No.1 33-39, 2000.
- [5] L. Accardi and M.Ohya, *A Stochastic limit approach to the SAT problem*, Open systems and Information Dynamics, 11-3, 219-233, 2004
- [6] L.Accardi and R.Sabbadini, *On the Ohya–Masuda quantum SAT Algorithm*, Preprint Volterra, N. 432, 2000.
- [7] E.Bernstein and U.Vazirani, *Quantum Complexity Theory*, In Proc. 25th ACM Symp. on Theory of Computation, 11–20, 1993.
- [8] T.Kasai, A. Adachi, S. Iwata, *Classes of Pebble Games and Complete Problems*,SIAM J. Comput. Volume 8, Issue 4, pp. 574-586 (1979)
- [9] S.Iriyama and M.Ohya, *Rigorous Estimate for OMV SAT Algorithm*, Open Systems & Information Dynamics, 15, 2, 173-187, 2008
- [10] S.Iriyama, M.Ohya and I.V.Volovich (2006) *Generalized Quantum Turing Machine and its Application to the SAT Chaos Algorithm*, QP-PQ:Quantum Prob. White Noise Anal., Quantum Information and Computing, 19, World Sci. Publishing, 204-225

- [11] S.Iriyama and M.Ohya (2008) Language Classes Defined by Generalized Quantum Turing Machine, *Open System and Information Dynamics* 15:4, 383-396.
- [12] S.Iriyama and M.Ohya (2009) The problem to construct Unitary Quantum Turing Machine for compute partial recursive function, TUS preprint.
- [13] S.Iriyama and M.Ohya (2010) Quantum Algorithm for Pebble Game and Its Computational Complexity, TUS preprint.
- [14] K.Goto, S.Iriyama, M.Ohya and I.Yamato, Quantum Algorithm for Protein Folding, in preparation.

Discrete Time Quantum Walk with Incommensurate Coin¹

Yutaka Shikano^{♣,♠2} and Hosho Katsura^{♡,◇3}

[♣]Department of Physics, Tokyo Institute of Technology

[♠]Department of Mechanical Engineering, Massachusetts Institute of Technology

[♡]Department of Physics, Gakushuin University

[◇]Kavli Institute for Theoretical Physics, University of California Santa Barbara

Abstract

We study a class of discrete time quantum walks with inhomogeneous coins defined in [Y. Shikano and H. Katsura, Phys. Rev. E **83**, 031122 (2010)]. We also discuss the localization or delocalization of the discrete time quantum walk by the eigenvalue spectrum.

1 Introduction

Let us recapitulate the discrete time random walk in classical probability theory. First, we prepare the **particle**, which is located at the origin at the beginning, and the **coin**. We then repeat the following procedure t times:

1. Coin flip which has two possible outcomes (head or tail) with the same probability.
2. Shift to the left or right according to the outcome.

At the final step (t), we calculate the probability distribution $\text{Pr}(n; t)$ at the position n .

In a manner analogous to the discrete time random walk, we define the discrete time quantum walk (DTQW) [1]. First, we prepare the **quantum particle**, which is located at the origin at the beginning and is labeled by the position state as $|0\rangle$, and the **quantum coin** with the orthogonal basis; $|L\rangle = (1, 0)^{\mathbf{T}}$, $|R\rangle = (0, 1)^{\mathbf{T}}$, where \mathbf{T} is the transposition. We repeat the following procedure t times:

1. Quantum coin flip: This operator is generally defined as

$$\begin{aligned} C_t &= \sum_n [(a_{n,t}|n, L\rangle + c_{n,t}|n, R\rangle)\langle n, L| \\ &\quad + (d_{n,t}|n, R\rangle + b_{n,t}|n, L\rangle)\langle n, R|] \\ &= \sum_n \left[|n\rangle\langle n| \otimes \begin{pmatrix} a_{n,t} & b_{n,t} \\ c_{n,t} & d_{n,t} \end{pmatrix} \right] =: \sum_n \left[|n\rangle\langle n| \otimes \hat{C}_{n,t} \right], \end{aligned} \quad (1)$$

where $\hat{C}_{n,t} (\in U(2))$ stands for the coin operator at the position n and time t .

¹This proceeding is for the talk at YITP Research Meeting ‘‘Duality and Scale in Quantum-Theoretical Sciences’’ held at YITP, Kyoto university and is based on the work [2].

²e-mail: shikano@th.phys.titech.ac.jp

³email: hosho.katsura@gakushuin.ac.jp

2. Shift according to the outcome (head or tail): This operator is defined as

$$W = \sum_n (|n-1, L\rangle\langle n, L| + |n+1, R\rangle\langle n, R|). \quad (2)$$

It should be noted that these operators are unitary processes. The one-step time evolution of the quantum walk at the step t is defined by $U_t = WC_t$.

2 Main Theorems

In the following, we assume the spatially inhomogeneous coins as

$$\begin{aligned} C_t = C &:= \sum_n [(\cos(2\pi\alpha n)|n, L\rangle + \sin(2\pi\alpha n)|n, R\rangle)\langle n, L| \\ &\quad + (\cos(2\pi\alpha n)|n, R\rangle - \sin(2\pi\alpha n)|n, L\rangle)\langle n, R|] \\ &= \sum_n \left[|n\rangle\langle n| \otimes \begin{pmatrix} \cos(2\pi\alpha n) & -\sin(2\pi\alpha n) \\ \sin(2\pi\alpha n) & \cos(2\pi\alpha n) \end{pmatrix} \right], \end{aligned} \quad (3)$$

where α is a real number ($\alpha \in \mathbb{R}$) and corresponds to the inverse period of the coin operations.

The limit distribution for the DTQW with the inhomogeneous coin has been obtained:

Theorem 1 (Ref. [2]). *For any irrational $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and any rational $\alpha = \frac{P}{4Q} \in \mathbb{Q}$ where P and Q are relatively prime, the limit distribution of the inhomogeneous QW divided by any power of the time variable is localized at the origin:*

$$\frac{X_t}{t^\theta} \Rightarrow I \quad (t \rightarrow \infty), \quad (4)$$

where X_t is the random variable for the position at the t step, $\theta (> 0)$ is an arbitrary parameter, and “ \Rightarrow ” means convergence in distribution. Here, the limit distribution I has the probability density function $f(x) = \delta(x)$ ($x \in \mathbb{R}$), where $\delta(\cdot)$ is the Dirac delta function. Note that, the limit distribution I is independent of the parameter θ .

This means that the DTQW with the inhomogeneous coin in almost all cases of the inverse period of the coin operators is localized since the irrational numbers are dense in the real numbers.

3 Localization/Delocalization of Discrete Time Quantum Walk and Eigenvalue Spectrum

In this section, we discuss a criterion for the localization or delocalization from the viewpoint of the eigenvalue spectrum of the one-step operator WC . Analogous to the [In analogy with the?] discussion on the discrete Schrödinger equation, which is called the Ruelle-Amrein-Georgescu-Enss(RAGE) theorem [3], we obtain the following theorem:

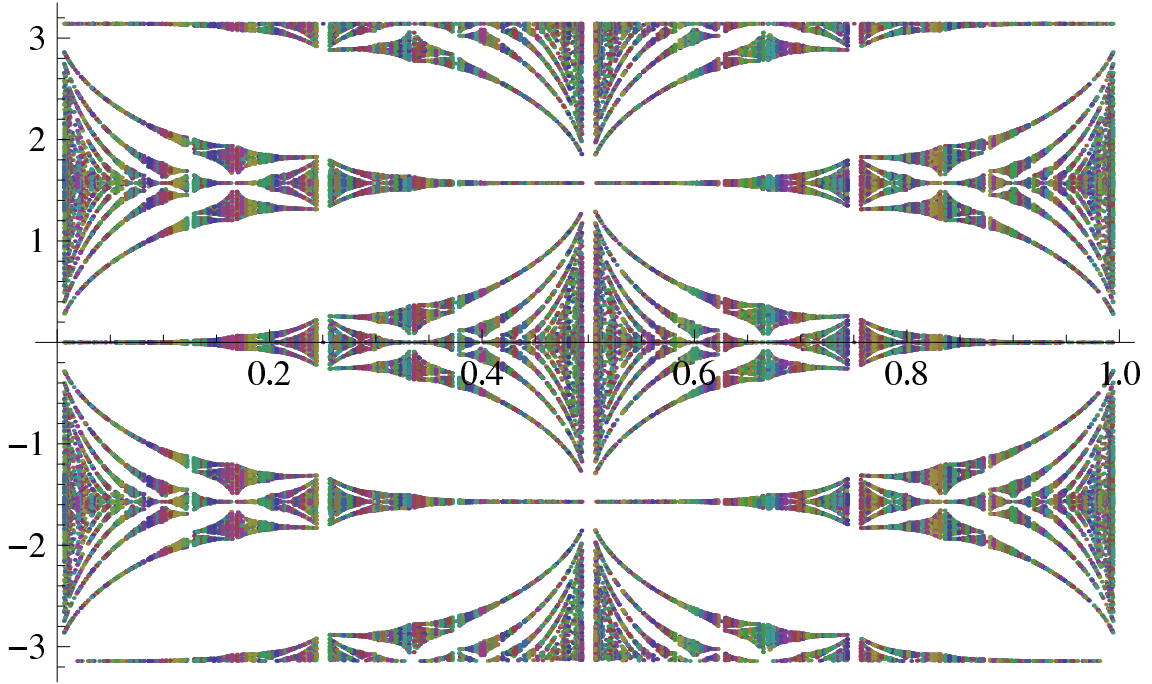


Figure 1: The self-similar and fractal structure in the DTQW with the inhomogeneous coin. Arguments of the eigenvalues of CW (vertical axis) are plotted as a function of the parameter $\alpha = \frac{P}{4Q}$ (horizontal axis) with $Q \leq 40$. Note that, P and Q must be relatively prime.

Theorem 2. *The distribution of the DTQW is not localized if the quantum walk operator WC has only continuous spectra and does not have embedded eigenvalues.*

This has been proved as a corollary in Ref. [4]. It is noted that the necessary and sufficient condition on the localization of the DTQW remains the open problem.

To study the eigenvalue spectrum of the DTQW with the inhomogeneous coin, we numerically plot the eigenvalue spectrum in the case of the rational numbers in Fig. 1. There holds the following theorem, which reflects the symmetries of the spectra in Fig. 1.

Theorem 3 (Refs. [2,5]). *For the eigenvalues of the one-step evolution operator WC , the following properties hold:*

(P1) *All the eigenvalues at α are identical to those at $1 - \alpha$.*

(P2) *For every eigenvalue λ , there is an eigenvalue λ^* .*

(P3) *For every eigenvalue λ , there is an eigenvalue $-\lambda$.*

(P4) *All the eigenvalues are simple, i.e., nondegenerate.*

(P5) *There are four eigenvalues $\lambda = \pm 1, \pm i$ for any $\alpha = \frac{P}{4Q} \in \mathbb{Q}$.*

(P6) *Every eigenvalue λ at $\alpha = \frac{P}{4Q} \in \mathbb{Q}$ corresponds to an eigenvalue $i\lambda$ at $\alpha + 1/2$.*

These properties are also used in Ref. [6].

4 Summary

We have introduced a class of the DTQWs with the spatially inhomogeneous coins. In almost all cases where the inverse period of the coins is irrational (or $\alpha = \frac{P}{4Q} \in \mathbb{Q}$), the wavefunction is localized. On the other hand, when α is rational and cannot be written as $\alpha = \frac{P}{4Q}$, the limit distribution still remains an open problem.

Acknowledgment

One of the authors (Y.S.) thanks Izumi Ojima, Andre Ahlbrecht, Reinhard F. Werner and Volkher Scholz for useful discussions. The authors thank the Yukawa Institute for Theoretical Physics at Kyoto University. Discussions during the YITP workshop YITP-W-10-14 on “Duality and Scale in Quantum-Theoretical Sciences” were useful to complete this work. Y.S. acknowledges the great hospitality and great support, especially from Margaret Sullivan, at Massachusetts Institute of Technology hosted by Seth Lloyd partially supported by JSPS Excellent Young Researcher Overseas Visit Program. Y.S. is supported by JSPS Research Fellowships for Young Scientists (Grant No. 21008624) and Global Center of Excellence Program “Nanoscience and Quantum Physics” at Tokyo Institute of Technology. H.K. is supported by the JSPS Postdoctoral Fellowships for Research Abroad and NSF Grant No. PHY05-51164.

References

- [1] A. Ambainis, E. Bach, A. Nayak, A. Vishwanath, and J. Watrous, in *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC'01)* (ACM Press, New York, 2001), pp. 37 - 49.
- [2] Y. Shikano and H. Katsura, *Phys. Rev. E* **83**, 031122 (2010).
- [3] B. Simon and M. Reed, *Methods of Modern Mathematical Physics Vol.III Scattering Theory* (Academic Press, Amsterdam, 1977).
- [4] A. Ahlbrecht, V. B. Scholz, and A. H. Werner, arXiv:1101.2298.
- [5] Y. Shikano and H. Katsura, arXiv:1104.2010 as the proceedings of QCMC 2010 to be published from AIP Conf. Proc.
- [6] H. Obuse and N. Kawakami, arXiv:1103.5545.

伏見関数と弱い測定

Husimi function and weak measurement

日大理工 市東 拓郎

1 はじめに

弱い測定を位置に対して行い、運動量を事後選択する。この観測によって得られる分布関数は伏見関数であり、この測定は Naimark 拡大の一つになっていることを示した。伏見関数はコヒーレント状態への遷移確率によって定義される量子系の分布関数である。この分布関数は正作用素値測度 (POVM) であるため対応した観測が存在する。これは系を拡大すればこの観測に対応する射影値測定に還元できる事が証明されているためである [1]。量子系の非可換な物理量の結合確率は量子系の性質である不確定性から議論に困難がつきまとう。位置、運動量それぞれの物理量は非可換であるために同時観測ができなく、分布関数を定義しようにも非可換であるために次の分布の値が等しくならない。

$$\mathrm{Tr} \hat{\rho} \delta(\hat{x} - x) \delta(\hat{p} - p) \neq \mathrm{Tr} \hat{\rho} \delta(\hat{p} - p) \delta(\hat{x} - x) \quad (1)$$

この非可換性が問題になる。ここで演算子 $\delta(\hat{x} - x) \delta(\hat{p} - p)$ を位置 x と運動量 p の対称 (Weyl ordering) にした演算子 $\hat{\omega}(x, p)$ を作る。この演算子から定義される分布関数 $W(x, p)$ を Wigner 関数と定義する [2]。

$$W(x, p) = \mathrm{Tr} \hat{\rho} \hat{\omega}(x, p) \quad (2)$$

$$\int dx dp \hat{\omega}(x, p) = \hat{I} \quad (3)$$

Wigner 関数は演算子の順序づけは問題にならないが演算子 $\hat{\omega}(x, p)$ が半正定値にならないため、Wigner 関数は必ず正值をとるとは限らず、分布関数としては不適當である。ただし、

式 (3) が成り立つという意味において総和は 1 となるので、全く分布関数の性質を満たさない訳でもない。また、別な対称化 (Anti-normal ordering) をした演算子より伏見関数が得られる。伏見関数は Wigner 関数とは違い値は全て正値をとる。伏見関数の詳細については後述する。以上は理論分布の話であるが、実際に非可換な物理量の測定理論の一つが弱い測定である [3]。先の Wigner 関数のように非可換な演算子の積は半正定値性を満たすとは限らないため、分布関数は半正値とは限らない。負値もとりうる分布は通常の観測では得られないが弱い測定は弱値としてこの値の観測を考えている。弱い測定は観測対象と観測装置を相互作用させる Von Neumann 測定を考える。一般に観測とは観測対象と観測装置を相互作用させ、観測装置を観測対象の状態に対応する状態に変化させ、それを観測する。通常、理想測定では観測装置の誤差は少なく測定値より系の状態が判別可能である。しかし、観測装置の分散が大きければ観測対象の状態は一回の測定で得られる情報が少なくなり、多数の観測値を集めなければ観測対象の状態が判別出来なくなるがその分、観測対象に対する観測による系の破壊が少なくなる。このような状況において、非可換な 2 つの物理量の同時測定に意味が出てくる。

2 伏見関数と弱い測定

先ほど演算子順序を選ぶことによって伏見関数が得られることに言及したが今回は伏見関数をコヒーレント状態から定義する。またここでは $\hbar = 1$ として議論を進める。コヒーレント状態を定義するため、まず次の消滅演算子 \hat{a} 、生成演算子 \hat{a}^\dagger を定義する。

$$[\hat{x}, \hat{p}] = i \quad (4)$$

$$\hat{a} = \frac{1}{\sqrt{2}}(\hat{x} + i\hat{p}) \quad (5)$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}(\hat{x} - i\hat{p}) \quad (6)$$

これらの演算子の交換関係は

$$[\hat{a}, \hat{a}^\dagger] = 1 \quad (7)$$

となる。この演算子の固有状態

$$\hat{a}|x, p\rangle = \frac{1}{\sqrt{2}}(x + ip)|x, p\rangle \quad (8)$$

をコヒーレント状態と定義する。

また、別の定義として消滅演算子の固有状態の一つである真空状態とハイゼンベルグ平面上の並進移動演算子 $\hat{D}(x, p)$ から次のように定義することも出来る。

$$\hat{a}|0\rangle = 0|0\rangle \quad (9)$$

$$|x, p\rangle = \hat{D}(x, p)|0\rangle \quad (10)$$

後者の定義に従えば真空状態の波動関数は位置、運動量ともに原点に中心を持つガウス分布であるため。コヒーレント状態の波動関数は位置、運動量に頂点を持つガウス分布になっていることが分かる。このコヒーレント状態の波動関数より次の POVM 演算子セットを定義する。

$$\{\hat{Q}(x, p)\}_{x, p} = \{|x, p\rangle\langle x, p|\}_{x, p} \quad (11)$$

この演算子セットより状態 $\hat{\rho}$ の伏見関数を定義する。

$$Q(x, p) = \text{Tr} \hat{\rho} \hat{Q}(x, p) \quad (12)$$

位置表示の波動関数から伏見関数への変換式は次のようになる。

$$Q(x, p) = \frac{1}{\sqrt{\pi}} \int dx' dx'' \exp \left\{ -\frac{1}{2}(x - x')^2 - \frac{1}{2}(x - x'')^2 - ip(x' - x'') \right\} \langle x' | \hat{\rho} | x' \rangle \quad (13)$$

伏見関数は POVM により構成されているため当然、次の確率としての性質を満たす。

$$\int \frac{dx dp}{2\pi} Q(x, p) = 1 \quad (14)$$

$$Q(x, p) \geq 0 \quad (15)$$

伏見関数が満たす他の性質として位置または運動量に関する周辺分布は次のようになる。

$$\int \frac{dp}{2\pi} Q(x, p) = \frac{1}{\sqrt{\pi}} \int dx' \exp\{-(x-x')^2\} \langle x' | \hat{\rho} | x' \rangle \quad (16)$$

$$\int \frac{dx}{2\pi} Q(x, p) = \frac{1}{\sqrt{\pi}} \int dp' \exp\{-(p-p')^2\} \langle p' | \hat{\rho} | p' \rangle \quad (17)$$

この伏見関数の位置に関する周辺分布をみると位置の波動関数をガウス分布でならしたものにしていることがわかる。この周辺分布を測定するには次のように位置に対して弱い測定をすればよい。ここからは簡単のため純粋状態のみを考える。まず、観測対象の系と観測装置（今回は観測装置が真空状態になっており波動関数がガウス分布になっていると仮定）を用意し二つの系を相互作用させる。相互作用ハミルトニアンは次のように設定する。

$$\hat{H}(t) = \delta(t-t')(\hat{x} \otimes \hat{I})(\hat{I} \otimes \hat{p}) \quad (18)$$

この時の時間 0 から t' までの時間発展演算子は

$$\hat{U} = \exp\left\{-i \int_0^{t'} dt \hat{H}(t)\right\} = \exp\left\{-i(\hat{x} \otimes \hat{I})(\hat{I} \otimes \hat{p})\right\} \quad (19)$$

と書ける。観測対象 $|\psi\rangle$ と観測装置の波動関数（真空状態） $\phi_0(x)$ の合成系が時間発展すれば

$$\hat{U} |\psi\rangle \phi_0(x) = \int dx' \langle x' | \psi \rangle \phi_0(x-x') |x'\rangle \quad (20)$$

となり、式 (20) に真空状態の式

$$\phi_0(x) = \frac{1}{\pi^{1/4}} \exp\left\{-\frac{1}{2}x^2\right\} \quad (21)$$

を当てはめ、絶対値の 2 乗をとれば

$$P(x, p) = |\hat{U} |\psi\rangle \phi_0(x)|^2 = \frac{1}{\sqrt{\pi}} \int dx' \exp\{-(x-x')^2\} |\langle x' | \psi \rangle|^2 \quad (22)$$

が得られ、伏見関数の位置の周辺分布と一致する。以上の議論は事後選択を考慮しない弱い測定である。次に位置の弱い測定の後運動量を事後選択する事を考える。弱い測定を

行った後の波束の収縮した状態 (20) に対して運動量を測定すれば次の式が得られる。

$$\langle p | \hat{U} | \psi \rangle \phi(x) = \int dx' \langle p | x' \rangle \langle x' | \psi \rangle \phi_0(x - x') \quad (23)$$

観測される分布関数はこの絶対値の 2 乗となるため

$$\frac{1}{\sqrt{\pi}} \int dx' dx'' \exp \left\{ -\frac{1}{2}(x - x')^2 - \frac{1}{2}(x - x'')^2 - ip(x' - x'') \right\} |\langle x' | \psi \rangle|^2 \quad (24)$$

が観測により得られるが、これは伏見関数そのものである。話を整理すれば位置を弱く測定しその後、運動量と事後選択すれば得られる分布関数は伏見関数と言うことである。この測定方法は弱い測定の一つであるが系を拡大し伏見関数の測定をしているため Naimark 拡大の一つになっている。

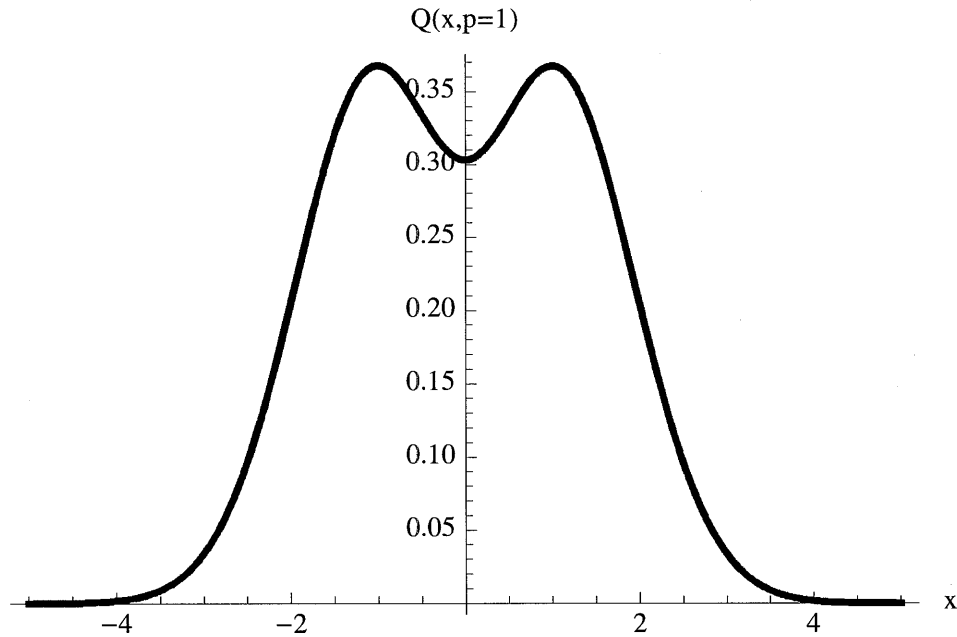


図 1: 1 光子状態の伏見関数の運動量 $p = 1$ に固定したときの位置 x の分布関数 (規格化はしていない)、弱い測定の言葉で言えば運動量 $p = 1$ を事後選択したときの位置 x の分布関数となる。

また今回は単純に弱い測定を行ったのみであるが観測装置の運動量の分散を十分小さくすることによって弱値を測定によって得ることが可能である (今回、弱値を観測する条件は本

題ではないため割愛する。詳しい議論は参考文献 [3] を参照)。この状態から得られる弱値は

$$\langle \hat{x} \rangle_\omega = \frac{\langle p | \hat{x} | \psi \rangle}{\langle p | \psi \rangle} \quad (25)$$

である。この式を変形すれば

$$\begin{aligned} \langle \hat{x} \rangle_\omega &= \frac{\langle \psi | p \rangle \langle p | \hat{x} | \psi \rangle}{|\langle p | \psi \rangle|^2} \\ &= \int dx \frac{x \langle \psi | \delta(\hat{p} - p) \delta(\hat{x} - x) | \psi \rangle}{|\langle p | \psi \rangle|^2} \end{aligned} \quad (26)$$

が得られ、運動量が p の時の条件付きの位置の期待値になる。この弱値と Wigner 関数の関係が気になるところであるが、この弱値は運動量が p の時の位置 \hat{x} の期待値を表しているが、式 (26) は運動量と位置の演算子順序に依存している為、直接この弱値と Wigner 関数と結びつけることは出来ない。また、伏見関数自体は位置、運動量の順によらないが弱値は順序による式になっている。

3 まとめ

今回、伏見関数と弱い測定の関係について議論した。結果としては測定装置に位置の測定（位置の弱い測定）を行い、観測対象の事後選択は行わなければ（観測対象は見ない）伏見関数の周辺分布が得られ、さらに事後選択を行えば伏見関数自体が得られる。また逆の立場からは伏見関数は運動量 $p = p'$ と固定すれば伏見関数 $Q(x, p = p')$ は運動量 $p = p'$ を事後選択した時の位置の弱い測定の分布関数となっている。これらの結果は伏見関数を観測する方法は Naimark 拡大の議論によって考えられるが、この議論が弱い測定や弱値の議論に関係している可能性を示している。また、伏見関数、Wigner 関数と弱値の関係性など細かい議論は省略したが、興味深い議論も出来るため今後、別な機会に伏見関数と弱値の関係も詳しく議論する予定である。

参考文献

- [1] A. S. Holevo, "Probabilistic and Statistical Aspects of Quantum Theory", North-Holland Publishing (1982)
- [2] U. Leonhardt, "Measuring the Quantum State of Light", Cambridge University press (1997)
- [3] Y. Aharonov, D. Rohrlich, "Quantum Paradoxes", WILEY-VCH (2005)

Algebra and analysis on ternary systems

堀口 俊、鈴木 理(日大文理)
前 祐行、郡山裕美、濃野聖晴(福岡教育大)

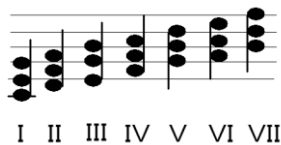
(概要)

この世界には3つの要素を与えると別の要素がひとつ定まる現象が多く見られる。例えば、音楽に見られる和音、素粒子論に見られるクォークによるバリオンの構成、RNAによる蛋白質の生成はその代表的な例である。本研究においては3つの要素からなる系(これを Ternary system とよぶ)を導入してその研究法を議論する。次の事柄を考えることが主要な目的である。(1) これに付随する代数(これを Ternary algebra とよぶ)を考えてその表現論を構成する。(2) Ternary Dirac 作用素を導入してこのクリフォード解析を与える。(3) 上記の現象、例えばクォーク閉じ込めの数理理論を構成する。

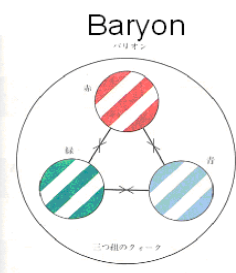
1. 研究の動機

まず、Ternary structure を考察する動機について述べる。

我々の身の回りには3つの要素に対してひとつの新たな要素を定めることにより成り立つ現象が多く見られる。これを Ternary phenomena とすることにする。ここでとりあげる現象を幾つか述べる。左より、和音、蛋白質の生成(コドン表)、クォークによるバリオンの生成である。これらの背後にある数理理論の構成をめざす。



1st base	2nd base				3rd base
	U	C	A	G	
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA end UAG end	UGU Cys UGC Cys UGA end UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G



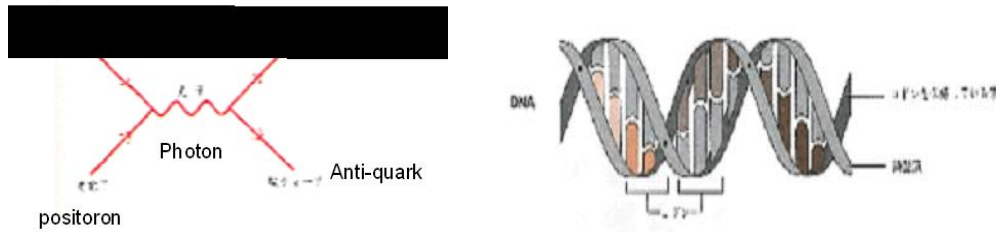
(Ternary number puzzle)

これらの現象を見ると(1) 和音は7の倍数個ある。(2) メッセージRNAは64種類からなっており、生成される蛋白質の種類は20である。(3) クォークは3世代から構成されており、各世代は3種類のクォークからなっている。

(問題 I)ここに現れる数のパズルはどのように説明されるか?

(Binary structure と Ternary structure の共存)

下図に素粒子論に見られる粒子と反粒子との双対性、DNAに見られる相補性を示した。



このように Binary structure と Ternary structure は共存している。そこで次の問題を提出することが出来る。

(問題 II) Binary structure と Ternary structure の共存する数理モデルを構成してクオーク・モデルあるいはDNA-RNAモデルを構成せよ。

2. 研究の方針

どのように研究を進めたらよいかよいのだろうか？また、現在 Ternary structure についてどのような研究がなされているのだろうか？先行研究は Prof. Kerner によるゲージ理論があるのみであるという過言ではない([1], [2])。ここでは離散群の対称性が中心になっており、連続群は取り扱われていない。従って考察にあたり、Ternary structure の基礎から構成しなくてはならない。ここでは Ternary system を導入し、このシステムを基本として数理理論の構築を行い、ゲージ理論にすすむ。最後に Ternary Clifford algebra を導入しそのクリフォード解析を構築する([3])。これをもとにクオーク閉じ込め理論の構成に挑む第一歩を目指す。

3. Ternary system

ここでは Ternary system を導入して基本事項をのべる。次節では特別なクラスをとり扱い composition algebra への応用を考える。

(Ternary system の定義)

まず、Ternary system を定義する。ここでは数列 $\{1, 2, \dots, M\}$ を与えこれから3つの元をとり、これらの元の作る順序集合を ternary element と言ひ、これらの集合を Ternary system ということにする。最も大きな Ternary system を Full ternary system とよぶことにする。

(Algebraic ternary system)

Binary algebra が定める Ternary system について述べる。これは最も小さい Ternary system と思われる。詳しくは次ぎの通である。 e_1, e_2, \dots, e_n の生成する実線形空間をと $R[e_1, e_2, \dots, e_n]$ とし積構造が定義されているとする。 $e_i e_j = e_k (i, j = 1, 2, \dots, n)$ となるときこの代数を mononic algebra と言ひ。この代数に対して $\{e_i, e_j, e_k\} (i, j = 1, 2, \dots, n)$ とおいて Ternary system が定義される。これを代数の定める Ternary system と言ひ。

(Ternary system の分類)

現在 Ternary system についての分類は行なわれていないといっても過言ではない。現実に現れる ternary 現象の記述を目指してその方向を模索する。ここでは形式的な分類を行なうことにとどめる。

(1) **Degenerate system** と **Non-degenerate system**: 要素 $\{i, j, k\}$ の各元 i, j, k がすべて異なるときこの system を non-degenerate といい、そうでないとき、つまりすくなくとも二つの元が等しくなる要素が可能となるとき Degenerate system という。次の節では Non-degenerate system をとりあつかう。Degenerate system の例は多く存在している:

例 1 (Full ternary) Full ternary system は degenerate system である。

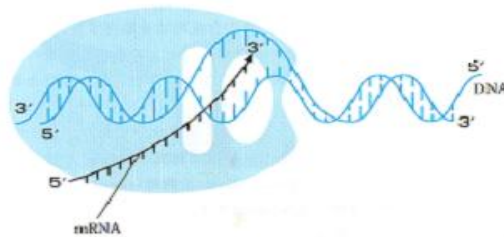
例 2 (Full matrix ternary)

$$e_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad e_2 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad e_3 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad e_4 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(2) **Exclusive system** と **Non-exclusive system**: 要素 $\{i, j, k\}$ のうち二つの文字を与えると残りの文字が唯一つ存在するとき、この system を Exclusive system といい、そうでないとき Non-exclusive system という。幾つか例を述べる

例 1 (Algebraic ternary system) 代数から生成される Ternary system は Exclusive system である。

例 2 (Generation of m-RNA) DNA から m-RNA が作られる仕方は Exclusive system を生成することがわかる。A, T, G, C の作る DNA の相補列はプロモーターにより拡張られここに相補元を作ることにより m-RNA が作られる。これは我々の言葉でいうと $\{i, i^*\}$ に対して $\{i, i^*, i^*\}$ を対応させると言う事ができる。ここで i^* は i の相補元である。



(Binary-Ternary coexistence system)

序文で述べたように Binary structure と Ternary structure が共存する system を Binary-ternary coexistence system あるいは単に Coexistence system という。そこで次の問題を考えることができる:

(問題 III) すべての Ternary system は Coexistence system に拡大することが可能か?

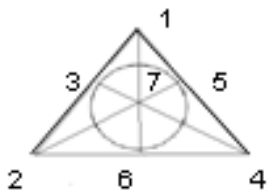
4. Complete system と perfect system

ここでは Binary structure に密接に関係する system に限って研究する。ここでは次の system に限定して考えることにする。(1) $\{i, j, k\}$ をその要素とするとき、 i, j, k はすべて異なるものとする。(2) また i, j, k の並べ替えも同一の要素とする。つまり order-free とする。(3) i と j を与えたとき element $\{i, j, k\}$ は順序に制限を与えたときただひとつであるとする。例えば $\{1, 2, \dots, 6\}$ について $\{1, 2, 3\}$ を採用すると $\{1, 2, 4\}, \{1, 2, 5\}$ … は許されない。さらに(4)すべての元 k に対して $\{i, j\}$ で $\{i, j, k\}$ を構成するもの(一つとは限らない)が存在する。この条件を満たす Ternary system を Complete ternary system という。(4) を次の条件(4)' どの k についても同数個の Ternary element を構成する、に置きかえたものを Perfect ternary system という。次の定理が成り立つことが予想される。証明も極めて簡単かと思われる。

定理 (1) M が奇数なら Complete ternary system が存在する。**(2)** M が 2 の冪から 1 を減じた数であるとき、Perfect ternary system となる。

注意 M が偶数のとき Exclusive system は Defect system になる。

Triangle realization



少し例を見るとこれが Quaternion, あるいは Octonion、Sederion と言ったある代数の系列を記述することがみとれる。実際、 $M=3$ では $\{1, 2, 3\}$ であり、 $M=7$ では $\{1, 2, 3\}, \{1, 4, 5\}$ … であり次のダイアグラムで記述されることが分かる。これは ternary element の構成規則 $\{i, j, k\}$ を代数演算に読み替えると Octonion の積表を記述することに他ならない。これより次の二つの重要な事実を見出すことになる：

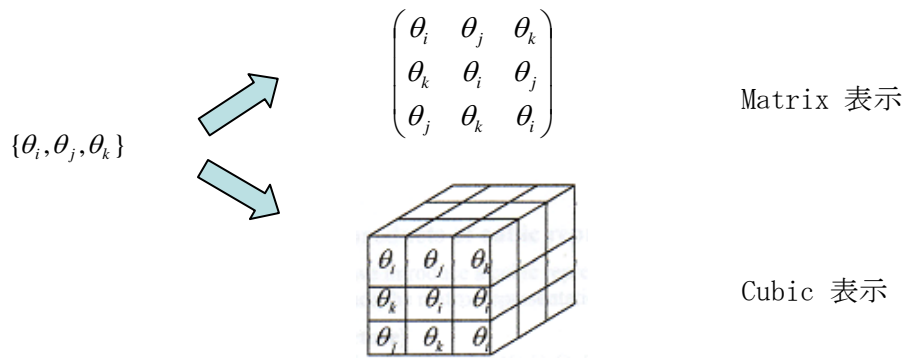
- (1) すべての Perfect ternary system にクリフォード型の代数が表現できる。
- (2) 従って時空間の記述がこの system で可能となる ([3])。

これはあとで Ternary analysis を用いて場の理論を展開するときに重要となる。そこで次のことがらを問題にすることが出来る：

(問題 IV) すべての Perfect ternary system はクリフォード型の代数の表現空間となるか？ これを用いて composition algebra は 4 種類しかないことが証明できないか？

5. Cubic の基礎

ここでは ternary element $\{\theta_i, \theta_j, \theta_k\}$ を matrix あるいは cubic 上に実現することを考える。ここでは 3 次の行列あるいは 3 次の cubic のみを考える：



(Cubic の座標表示)

ここでは ternary element の cubic 実現を考える。多くの例を見るとどうやら cubic に現れる type は次の2種類となるらしい。これらをまとめて cubic 表現という。

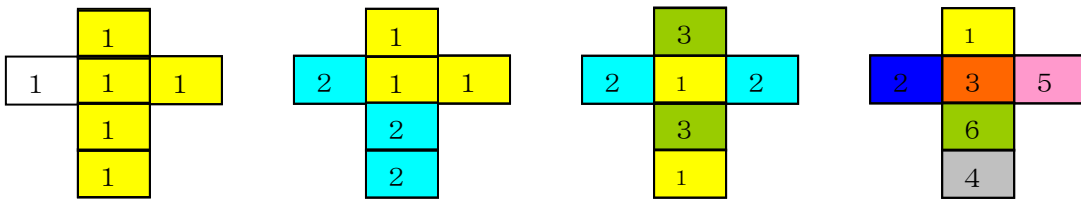
$$\begin{pmatrix} \theta_i & \theta_j & \theta_k \\ \theta_k & \theta_i & \theta_j \\ \theta_j & \theta_k & \theta_i \end{pmatrix} \qquad \begin{pmatrix} \theta'_i & \theta'_j & \theta'_k \\ \theta'_j & \theta'_k & \theta'_i \\ \theta'_k & \theta'_i & \theta'_j \end{pmatrix}$$

Diagonal type

Anti-diagonal type

cubic の表示について基本的な事項を述べる。3次の cubic は27個のセルからなる。以下セルのタイプと座標系を考える：

(1) Cell の type



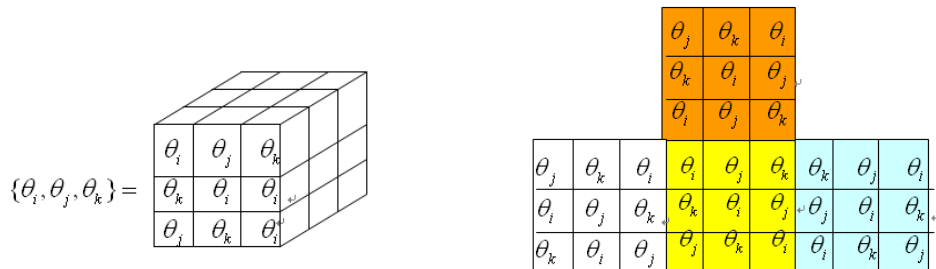
one color

two colors

three colors

six colors

(2) Coordinate



このとき cubic の座標 Q_{ijk} を下の様におく。一般の full ternary system もより多くのマス目を持った cubic に表すことが出来るかを問題にすることも出来る。

$$(Q_{ijk}) = \begin{array}{|c|c|c|c|c|c|c|c|c|} \hline 111 & 121 & 131 & 211 & 221 & 231 & 311 & 321 & 331 \\ \hline \theta_k & \theta_j & \theta_i & \theta_j & \theta_i & \theta_k & \theta_i & \theta_k & \theta_j \\ \hline 112 & 122 & 132 & 221 & 222 & 232 & 312 & 322 & 332 \\ \hline \theta_j & \theta_i & \theta_k & \theta_i & \theta_k & \theta_j & \theta_k & \theta_j & \theta_i \\ \hline 113 & 123 & 133 & 231 & 223 & 233 & 313 & 323 & 333 \\ \hline \theta_i & \theta_k & \theta_j & \theta_k & \theta_j & \theta_i & \theta_j & \theta_i & \theta_k \\ \hline \end{array}$$

6. Associative cubic algebra (Basic cubic algebra)

ここではひとまず Kerner に従って Ternary algebra の概念を導入する ([1])。この代数については定義があたえられているだけで実質的には未開発の分野と言える。

(Ternary algebra の導入)

線形空間 A で積構造 $m: A \otimes A \otimes A \rightarrow A$ を持つものを Ternary algebra と呼ぶ。次の条件をおいて考えることは自然である。次の問題を考察する：

(問題 V) $m(e_i, e_j, e_k) = \sum c_{ijk}^{(l)} e_l$ とおいて代数を決定せよ。ここで決定するとは標準的な積を定め一般的な積がどのようにしてそこに帰着するかその処方箋を記述することと理解する。

Binary product により積が定義されている、すなわち $m(X, Y, Z) = X \circ Y \circ Z$ となるとき Ternary algebra が pure であるとは $X \circ Y \notin A(X, Y \in A)$ となることとする。以下 pure な代数のみを考えることにする。以上で得られた ternary algebra を ternary system 上に実現する表現論を構成する。これに基づき ternary system を分類することは興味ある問題である。行列代数は $e_i \circ e_j = \sum c_{ij}^{(l)} e_l$ とおいて結合条件を満たすものとして得られる。これに対応するものを見つけることが課題である。現在その候補者は絞りきれていない。Prof. Kerner は次の積を標準的な積 (Standard product) と呼ぶことを提案している：

$$m(X, Y, Z) = \sum X_{ipq} Y_{pjr} Z_{prk}$$

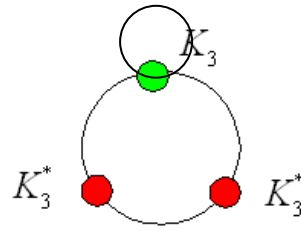
ここで $X = \sum X_{ijk} e_i \otimes e_j \otimes e_k$, $Y = \sum Y_{ijk} e_i \otimes e_j \otimes e_k$, $Z = \sum Z_{ijk} e_i \otimes e_j \otimes e_k$ である。行列代数に対応するものとして Cubic algebra, つまり cubic に実現できる代数を見出すことを考える。ここでは associative cubic algebra のみを取り扱う。

(Cubic algebra の導入)

Cubic type の元から生成される次の結合代数を考える：

$$\begin{array}{ccc} e_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & e_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} & e_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ e_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} & e_5 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} & e_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \end{array}$$

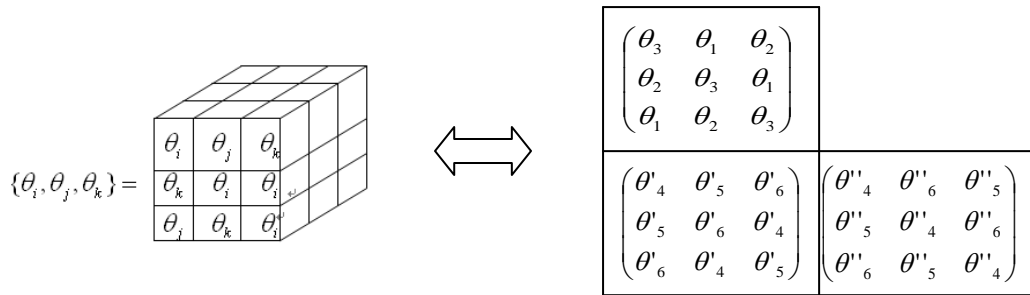
	e_1	e_2	e_3	e_4	e_5	e_6
e_1	e_1	e_2	e_3	e_4	e_5	e_6
e_2	e_2	e_3	e_1	e_5	e_6	e_4
e_3	e_3	e_1	e_2	e_6	e_4	e_5
e_4	e_1	e_3	e_2	e_4	e_6	e_5
e_5	e_2	e_1	e_3	e_5	e_4	e_6
e_6	e_3	e_2	e_1	e_6	e_5	e_4



左図の交換関係を満たす。 $e_i (i=1,2,3)$ の生成する subalgebra を K_3 と書く。これより Ternary algebra を導入することが出来る。 $e_i (i=4,5,6)$ の生成する線形空間を K_3^* とすると、

$$\begin{cases} K_3 K_3 \subseteq K_3, K_3^* K_3^* \subseteq K_3 \\ K_3^* K_3 \subseteq K_3^*, K_3 K_3^* \subseteq K_3^* \end{cases}$$

従って $K_3^* K_3^* \not\subseteq K_3^*$ であり、 $K_3^* K_3^* K_3^* \subseteq K_3^*$ となることに注意する。これは K_3^* は Ternary algebra が pure ternary algebra であることを示している。この代数は上記の cubic 表現で表すことが可能である。



ただし、このときひとつのセルの表面の座標は3面とも異なっていることに注意する。

7.Coexistence system の構成

序で述べたように多くの Ternary system においても Binary structure と Ternary structure とが共存していることを見た。ここではこのモデル構成を行なう。まず、Binary structure に実現される Ternary structure が存在することに注目する。次の定義をおく：

定義 Binary structure \hat{K} が Ternary sub algebra をもつとは次の分解 $\hat{K} = K \oplus K^*$ をもつ時である：

- (1) K, K^* は線形部分空間である。 (2) $\begin{cases} K_3 K_3 \subseteq K_3, K_3^* K_3^* \subseteq K_3 \\ K_3^* K_3 \subseteq K_3^*, K_3 K_3^* \subseteq K_3^* \end{cases}$

従って $K_3 * K_3 * \not\subseteq K_3 *$ であり、 $K_3 * K_3 * K_3 * \subseteq K_3 *$ であり Pure ternary algebra であることに注意する。 \hat{K} を Ternary algebra $K_3 *$ の Binary extension と呼び、 $K_3 *$ を \hat{K} の Ternary sub algebra と呼ぶことにする。

このような例は極めて多い。

例 1 (Full matrix algebra) Full matrix algebra $M(2^n)$ はこの例となっている。実際 $K_3 = R[e_1, e_2]$, $K_3^* = R[e_3, e_4]$, とおくと上の条件を満たしている。

例 2 (Quaternion, Octonion)

Quaternion

	e_1	e_2	e_3	e_4
e_1	e_1	e_2	e_3	e_4
e_2	$e_2 - e_1$	e_4	$-e_3$	
e_3	$e_3 - e_4 - e_1$	e_2		
e_4	e_4	$e_3 - e_2 - e_1$		

Octonions

	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_0	e_0	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_1	$e_1 - e_0$	$e_3 - e_2$	$e_5 - e_4 - \varepsilon_7$	ε_6				
e_2	$e_2 - e_3$	e_0	e_1	e_6	$\varepsilon_7 - e_4 - \varepsilon_5$			
e_3	e_3	$e_2 - e_1 - e_0$	$e_7 - \varepsilon_6$	$\varepsilon_5 - e_4$				
e_4	$e_4 - \varepsilon_5 - e_6 - e_7 - e_0$	ε_1	ε_2	e_3				
e_5	e_5	$\varepsilon_4 - e_7 - e_6 - e_1 - e_0 - \varepsilon_3$	ε_2					
e_6	e_6	ε_7	$e_4 - e_5 - e_2$	$e_3 - e_0 - \varepsilon_1$				
e_7	$e_7 - \varepsilon_6$	e_5	$e_4 - e_3 - e_2$	$\varepsilon_1 - e_0$				

例 3 (Associative cubic algebras) 我々の構成した associative cubic algebra は standard product の条件を満たしていない。しかし、non-associative algebra はこの積の条件を満たす？これらの例から次の問題が提出される：

(問題 VI) すべての Ternary algebra は Binary algebra に拡張されるか？

8. Ternary analysis の構成

Ternary analysis を構成する。Prof. Kerner により導入された Ternary Dirac algebra を従来の Binary Dirac algebra と共存させることにより Ternary analysis を構成している ([2])。最後に次の事柄をより扱う。これは素粒子物理学の最終目標の一つでもある。

(問題 VII) Ternary system 上に場の理論を構成してクオーク閉じ込め理論を構成することが出来るか？

(Binary analysis の構成)

場の量子論においては電子を記述するディラックの方程式 D 及び相対論から導かれるクライン・ゴルドンの方程式 Δ は基本的な方程式である：

$$i \frac{\partial}{\partial t} \psi = D \psi, \quad D \psi = \sum_{i=1}^4 \gamma_i \frac{\partial}{\partial x_i} \psi + m \psi, \quad D^* \psi = \sum_{i=1}^4 \gamma_i^* \frac{\partial}{\partial x_i} \psi + m \psi,$$

$$\frac{\partial^2}{\partial t^2} \phi = \Delta \phi, \quad \Delta \phi = \sum_{i=1}^4 \frac{\partial^2}{\partial x_i^2} \phi + m \phi$$

ここで $\gamma_i (i=1,2,3,4)$ はディラックの行列であり、次の交換関係をみます：

$$\gamma_i \gamma_j + \gamma_j \gamma_i = 2\delta_{ij} (i=1,2,3,4)$$

ディラックの方程式はクライン・ゴルドン方程式の平方をとるという考えから導かれることがわかる。つまり $DD^* = \Delta 1$ が成り立つ。これはこの世界の binary な構造を明確にあらわしている。

(Ternary analysis の構成)

次に Prof. Kerner に従って ternary Dirac algebra を導入してクオークの場の方程式、Ternary Klein-Gordon 方程式を導入する。

$$Q_i Q_j Q_k + Q_j Q_k Q_i + Q_k Q_j Q_i = 3\eta_{ijk} (i, j, k = 1, 2, 3)$$

に対して Ternary Dirac equation 及び Ternary Klein-Gordon equation を次のように定める：

$$\frac{\partial}{\partial t} \psi = \sum_{i=1}^3 Q_i \frac{\partial}{\partial x_i} \psi \quad \frac{\partial^3}{\partial t^3} \phi = \sum_{i=1}^3 \frac{\partial^3}{\partial x_i^3} \phi - 3 \frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} \phi$$

(Binary analysis と Ternary analysis)

Binary analysis と Ternary analysis との関係を考える。波動が外部に拡散されないという条件から閉じ込め条件を模索する。上記の方程式は次の行列式の量子化として与えられることに注意する：

$$X(\theta) = \begin{vmatrix} \theta_i & \theta_j & \theta_k \\ \theta_k & \theta_i & \theta_j \\ \theta_j & \theta_k & \theta_i \end{vmatrix} (= \theta_i^3 + \theta_j^3 + \theta_k^3 - 3\theta_i \theta_j \theta_k)$$

$$\begin{aligned} X(\theta) &= \theta_i^3 + \theta_j^3 + \theta_k^3 - 3\theta_i \theta_j \theta_k \\ &= (\theta_i + \theta_j + \theta_k)(\theta_i + \omega\theta_j + \omega^2\theta_k)(\theta_i + \omega^2\theta_j + \omega\theta_k) \end{aligned}$$

上式右辺の 2 次形式の標準形に注意する

補題 $F(\theta) = \theta_i^2 + \theta_j^2 + \theta_k^2 - \theta_i \theta_j - \theta_j \theta_k - \theta_k \theta_i$ を直交変換により標準形とすると $F(\theta') = \theta_i'^2 + \theta_i''^2$ と出来る

従ってフーリエ変換するとこれは 2 次元のラプラス作用素であり、対応する Klein-Gordon 方程式は波動方程式になる。このことは Binary analysis と Ternary analysis がどのように結びつくかをどのように考察すべきであることを教えているものと想像される。これらの詳しい結果は今後の研究課題である。最後にこの研究の道しるべとなる Binary analysis と Ternary analysis との対応を示す表を掲げておく。

	Binary analysis	Ternary analysis
Dirac matrix	$\gamma_i (i=1,2,3,4)$ $\gamma_i \gamma_j + \gamma_j \gamma_i = 2\delta_{ij} (i=1,2,3,4)$	$Q_i (i=1,2,3)$ $Q_i Q_j Q_k + Q_j Q_k Q_i + Q_k Q_j Q_i = 2\eta_{ijk} (i, j, k = 1,2,3)$
Dirac operator	$D\psi = \sum_{i=1}^4 \gamma_i \frac{\partial}{\partial x_i} \psi + m\psi$	$D\psi = \sum_{i=1}^3 Q_i \frac{\partial}{\partial x_i} \psi + m\psi$
Conjugate Dirac operator	$D^* \psi = \sum_{i=1}^4 \gamma_i^* \frac{\partial}{\partial x_i} \psi + m\psi$	$D^* \psi = Q_1 \frac{\partial}{\partial x_1} \psi + j Q_2 \frac{\partial}{\partial x_2} \psi + j^2 Q_3 \frac{\partial}{\partial x_3} \psi + m\psi$ $D^{**} \psi = Q_1 \frac{\partial}{\partial x_1} \psi + j^2 Q_2 \frac{\partial}{\partial x_2} \psi + j Q_3 \frac{\partial}{\partial x_3} \psi + m\psi$
Klein-Gordon operator	$\Delta \psi = \sum_{i=1}^4 \frac{\partial^2}{\partial x_i^2} \psi + m\psi$	$\frac{\partial^3}{\partial t^3} \psi = \sum_{i=1}^3 \frac{\partial^3}{\partial x_i^3} \psi - \frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} \psi$
Hurwitz condition	$ xy = x y $	$X(\Theta) = X(\theta)X(\theta')$
• • • •	• • • • •	• • • • •

参考文献

- [1] L. Vainerman and R. Kerner: On special classes of n-algebras, J. Math. Phys. 37-5 (1996), 2553-2565
- [2] R. Kerner: The cubic chessboard. Class Quantum Grav. 14 (1997) A-203-225
- [3] R. Kerner: Cubic and ternary algebras, ternary symmetries, and the Lorentz group, 数理解析研究所講究録 1705「量子科学における双対性とスケール」(2010), 134-146
- [4] R. Kerner and O. Suzuki: Internal symmetry groups of cubic algebras and ternary algebras, in preparation.
- [5] Hartle: Essential Genetics, A genetic perspective, Jonson-Battler Publisher, 2002, p 519

Enhanced binding for the generalized semi-relativistic quantum field model.

Toshimitsu TAKAESU
(Kyushu University)

1 Introduction

In this article, a system of semi-relativistic particles interacting with a massive Bose field is analyzed. The dynamics of the semi-relativistic particles is determined by the free relativistic Schrödinger operator. We consider the total Hamiltonian with a generalized interaction. By imposing generalized ultraviolet cutoff conditions on the Bose field, it is seen that the total Hamiltonian of the system is a self-adjoint operator on a Hilbert space. We are interested in spectral properties of the total Hamiltonian.

In the last decade, the spectral analysis for the quantum particles system interacting with Bose fields has been successfully analyzed. In [6], the detailed survey on the recent progress in this subject is given. The main problems of spectral analysis for quantum field models can be classified into (i) ground states, (ii) resonances, (iii) scattering theory, (iv) non-relativistic limits, scaling limits, and so on. In this article, we are mainly interested in analysis of the phenomena, called *enhanced binding*, which is related to the ground states problems. For a self-adjoint operator X with bounded from below, it is said that X has the ground state if the bottom of the spectrum of X is the eigenvalue of X . For quantum field system, it is said that there is the enhanced binding if the total Hamiltonian has the ground state even if the free Hamiltonian does not have the ground state. In this article, it is proven that under some assumptions including a binding condition, there is the enhanced binding in the interacting system. Historically, analysis of the enhanced binding for the quantum field model is initiated by Hiroshima-Spohn [7]. They analyze the non-relativistic QED model, and this model has also been analyzed in [2, 3, 4, 5]. Arai-Kawano [1] analyze the enhanced binding for the generalized quantum field model, and Hiroshima-Sasaki [8] investigate the Nelson model. Recently, the absence of the enhanced binding of the non-relativistic QED model has been investigated in [9].

2 Main Result

Let us define the state space and the total Hamiltonian for the semi-relativistic particles system interacting with a massive Bose field. The state space is given by $\mathcal{H} = L^2(\mathbf{R}_x^{dN}) \otimes \mathcal{F}_b(L^2(\mathbf{R}_k^d))$ where $\mathcal{F}_b(\mathcal{K})$ denotes the boson Fock space over the Hilbert space \mathcal{K} . We use a natural identification $\mathcal{H} \simeq \int_{\mathbf{R}^{dN}}^{\oplus} \mathcal{F}_b(L^2(\mathbf{R}_k^d)) d\mathbf{x}_1 \cdots d\mathbf{x}_N$ where $\int_X^{\oplus} d\mu$ denotes the fibre direct integral with base space $(X, d\mu)$. The total Hamiltonian is given by

$$H(\kappa) = H_p \otimes I + I \otimes H_b + \kappa H_I.$$

Here $H_p = \sum_{j=1}^N \left(\sqrt{-\Delta_j + M^2} - M \right)$ is the relativistic Schrödinger operator with the rest mass $M > 0$ and $H_b = d\Gamma_b(\omega)$ is the second quantization of $\omega \in C(\mathbf{R}^d)$ satisfying $\inf \omega(\mathbf{k}) > 0$. To assume the condition $\inf \omega(\mathbf{k}) > 0$ is called the massive condition. The interaction H_I is defined by $H_I = \sum_{j=1}^N \int_{\mathbf{R}^{dN}}^{\oplus} \phi(u_{\mathbf{x}_j}) d\mathbf{x}_1 \cdots d\mathbf{x}_N$ where $\phi(\xi) = \frac{1}{\sqrt{2}} (a(\xi) + a^*(\xi))$ denotes the field operator for $\xi \in L^2(\mathbf{R}_k^d)$, and $u_{\mathbf{x}}$ is a multiplication operator on $L^2(\mathbf{R}_x^d)$ satisfying the following condition :

(A.1) (Generalized ultraviolet-cutoff condition)

$$\sup_{\mathbf{x} \in \mathbf{R}^d} \int_{\mathbf{R}^d} |u_{\mathbf{x}}(\mathbf{k})|^2 d\mathbf{k} < \infty.$$

Let us set $H_0 = H_p \otimes I + I \otimes H_b$. Under the condition **(A.1)**, the interaction H_I is relatively bounded to $I \otimes H_b^{1/2}$. Then, H_I is relatively bounded to H_0 with infinitely small bond. Hence, from the Kato-Rellich theorem, $H(\kappa)$ is self-adjoint on $\mathcal{D}(H_0) = \mathcal{D}(H_p \otimes I) \cap \mathcal{D}(I \otimes \mathcal{H}_b)$ and essentially self-adjoint on any core of H_0 .

Our main interest is the ground state of the total Hamiltonian $H(\kappa)$. Here let us prepare for some notations. For an operator X , $\sigma(X)$ denotes the spectrum of X and set $E_0(X) = \inf \sigma(X)$. $\sigma_p(X)$ and $\sigma_{\text{ess}}(X)$ denote the point spectrum and essential spectrum of X , respectively. It is noted that, the ground state of H_0 does not exist, i.e. $E_0(H_0) \notin \sigma_p(H_0)$, since any external potentials are not turned on H_p .

To derive an potential of the particles from the interaction between the particles and the field, let us use the unitary transformation, called *dressing transformation* defined by

$$U(\kappa) = \exp \left(i\kappa \sum_{j=1}^N \int_{\mathbf{R}^d}^{\oplus} \pi \left(\frac{u_{\mathbf{x}_j}}{\omega} \right) d\mathbf{x}_1 \cdots d\mathbf{x}_N \right),$$

where $\pi(\xi) = \frac{i}{\sqrt{2}} (-a(\xi) + a^*(\xi))$ is the conjugate operator for $\xi \in L^2(\mathbf{R}_k^d)$. We introduce the following condition.

$$\begin{aligned} \text{(A.2)} \quad & \sup_{\mathbf{x} \in \mathbf{R}^d} \int_{\mathbf{R}^d} |\nabla u_{\mathbf{x}}(\mathbf{k})|^2 d\mathbf{k} < \infty, \quad \sup_{\mathbf{x} \in \mathbf{R}^d} \int_{\mathbf{R}^d} |\Delta u_{\mathbf{x}}(\mathbf{k})|^2 d\mathbf{k} < \infty, \\ & \left(\frac{\nabla f_{\mathbf{x}}}{\omega}, f_{\mathbf{y}} \right) \in \mathbf{R}, \quad \mathbf{x}, \mathbf{y} \in \mathbf{R}^d. \end{aligned}$$

Then from the canonical commutation relations, we have

$$U(\kappa)^{-1} H(\kappa) U(\kappa) = (H_p + \kappa^2 V_{\text{eff}}) \otimes I + I \otimes H_b + \delta H_p(\kappa), \quad (1)$$

where

$$V_{\text{eff}} = -\frac{1}{2} \sum_{j,l=1}^N \int_{\mathbf{R}^d} \frac{\overline{u_{\mathbf{x}_j}(\mathbf{k})} u_{\mathbf{x}_l}(\mathbf{k})}{\omega(\mathbf{k})} d\mathbf{k}, \quad (2)$$

$$\delta H_p(\kappa) = U(\kappa)^{-1} (H_p \otimes I) U(\kappa) - H_p \otimes I. \quad (3)$$

It is proven that $\delta H_p(\kappa)$ is relatively bounded with respect to H_0 in a similar way to ([10], Proposition 3.1).

Let us introduce the following assumptions :

(A.3) There exist constant $d > 0$ and $\tau > 0$ such that

$$\sup_{\mathbf{x} \in \mathbf{R}^d} |u_{\mathbf{x}}(\mathbf{k}) - u_{\mathbf{x}}(\mathbf{k}')| \leq c |\mathbf{k} - \mathbf{k}'|^\tau.$$

(A.4) (Binding Condition)

There exists constant $\kappa_* > 0$ such that for $0 < \kappa < \kappa_*$,

$$\inf \sigma_{\text{ess}} \left(H_p + \kappa^2 V_{\text{eff}} \right) \geq E_0 \left(H_p + \kappa^2 V_{\text{eff}} \right) + v(\kappa),$$

where $v(\kappa) > 0$ is positive real number.

Theorem (Ground states in massive condition)

Assume (A.1) - (A.4). Then $H(\kappa)$ has purely discrete spectrum in $[E_0(H(\kappa)), E_0(H(\kappa)) + m)$ for sufficiently small $\kappa > 0$. In particular the ground state of $H(\kappa)$ exist for sufficiently small $\kappa > 0$.

(Outline of the Proof) Let us set $H^U(\kappa) = U(\kappa)^{-1}H(\kappa)U(\kappa)$. By applying the methods of the finite volume approximation (refer to e.g. [1, 7]), we can construct the Hamiltonians $H_{\Lambda,\varepsilon}^U(\kappa)$ and $H_{\Lambda}^U(\kappa)$ such that (1) $H_{\Lambda,\varepsilon}^U(\kappa)$ has purely discrete spectrum in $E_0(H_{\Lambda,\varepsilon}^U(\kappa)), E_0(H_{\Lambda,\varepsilon}^U(\kappa)) + m)$ for sufficiently small $\kappa > 0$, (2) $H_{\Lambda,\varepsilon}^U(\kappa)$ and $H_{\Lambda}^U(\kappa)$ converges to $H_{\Lambda}^U(\kappa)$ and $H^U(\kappa)$, respectively, in norm resolvent sense. Then the proof is obtained.

References

- [1] A. Arai and H. Kawano, Enhanced binding in a general class of quantum field models, *Rev. Math. Phys.* **15** (2003) 387-423.
- [2] J. M. Barbaroux, H. Linde and S. Vugalter, Quantitative estimates on the enhanced binding for the Pauli-Fierz operator, *J. Math. Phys.* **46** (2005) 122103.
- [3] I. Catto, P. Exner and C. Hainzl, Enhanced binding revisited for a spinless particle in non-relativistic QED, *J. Math. Phys.* **45** (2004) 4174-4185.
- [4] T. Chen, V. Vougalter and S. A. Vougalter, The increase of binding energy and enhanced binding in non-relativistic QED, *J. Math. Phys.* **44** (2003) 1961-1970.
- [5] C. Hainzl, V. Vougalter and S. A. Vougalter, Enhanced binding in non-relativistic QED, *Commun. Math. Phys.* **44** (2003) 1961-1970.
- [6] F. Hiroshima, Perturbations of embedded eigenvalues in quantum field theory, *Sugaku expositions* **21** (2008) 177-207.
- [7] F. Hiroshima and H. Spohn, Enhanced binding through coupling to a quantum field, *Ann. Henri. Poincaré* **2** (2001), 1150-1187.
- [8] F. Hiroshima and I. Sasaki, Enhanced binding of an N particle system interacting with a scalar field I, *Math. Z.* **259** (2008) 657-680.
- [9] F. Hiroshima, H. Spohn and A. Suzuki, The no-binding regime of the Pauli-Fierz model, (preprint).
- [10] T. Takaesu, Scaling Limits for the System of Semi-Relativistic Particles Coupled to a Scalar Bose Field, to appear in *Lett. Math. Phys.* (2011).

Weak Measurement によるスピン相互作用の推定¹

田中 宗^A, 鹿野豊^{B,C}

^A 近畿大学 総合理工学研究科 量子コンピュータ研究センター

^B 東京工業大学 理工学研究科 基礎物理学専攻

^C マサチューセッツ工科大学機械工学科

量子情報処理を真に実現するためには、量子ビット間に働く相互作用ハミルトニアンを正確に決定する必要がある。我々は最も簡単な場合として、2個の量子ビットから成る系のスピン・スピン相互作用を実験的に推定する新しい方法を提案した [1,2]。我々の手法は、Yakir Aharonov らによって提案された weak measurement [3] を基礎としたものである。また提案した手法を用いて、ダイヤモンド中の NV 中心のハミルトニアン推定のデモンストレーションを数値的に行い、その有用性を確認した。

1 導入

ショアのアルゴリズム [4] が提案されて以来、量子情報処理の実現を目指した研究が多くなされてきている [5,6]。量子情報処理を実現するためには、与えられた問題を表現するハミルトニアンが正確に生成されているかを確認することが必要不可欠である。それができないと、注目している量子ビットに対して所望の操作をすることが不可能である。つまり、量子計算機を実現するための最小の条件として、2つのキュービットから成る系のスピン・スピン相互作用を正確に求める必要がある。

量子メモリーに代表される、量子情報を蓄積するデバイスは量子情報処理において必要不可欠である [7]。量子情報を蓄積するデバイスの重要かつ必要不可欠な特徴として、デコヒーレンスに対して頑強である、情報の読み書きが容易であるということが挙げられる。そのようなデバイスを作るためには、次のような2つの異なる量子ビットから成る系を作る必要がある。1つの量子ビットは操作が容易で、かつ情報の読み出しが簡単であるもの、もう1つの量子ビットはデコヒーレンスに対して頑強である、すなわち、情報を長時間そのまま蓄えておくことができるものである。また、片方の量子ビットからもう片方の量子ビットへ情報が転送できることも必要条件である。量子情報を転送するためには、2つの量子ビット間のスピン・スピン相互作用が完璧に分かっている必要がある。電子スピン共鳴 (ESR) のような通常ハミルトニアン推定手法においては [8]、ある実験結果から、2つの量子ビット間の相対位置を仮定した後、量子ビット間の相互作用を決

¹この原稿は、2010年11月4日～6日に京都大学基礎物理学研究所にて開催された基研研究会「量子科学における双対性とスケール」の研究会報告の原稿である。

定するのが普通である。ところが、量子ビットの相対距離を測定する際には、デコヒーレンスの影響、あるいは未知の実験的なエラーの効果を完全に除去することができない。そのため、量子ビット間に働く相互作用を完全に決定することはきわめて難しい問題である。

そこで我々は2つの量子ビットから成る、スピン・スピン相互作用を推定する新しい手法を開発した。またダイヤモンド中の NV(Nitrogen Vacancy) 中心の相互作用を例に、我々の提案する手法を数値計算によってデモンストレーションした。スピン・スピン相互作用を推定する手法はこれまでも数多く提案されてきているが [9–16]、我々の手法は weak measurement [3] と呼ばれる概念を基礎としており、その点において他の手法と一線を画する。

本題に移る前に、簡単に weak measurement について紹介しておこう。ターゲット系に対して weak measurement を実行すると、観測量 A の weak value を以下のように定義することができる。

$$\langle A \rangle_w := \frac{\langle f|A|i \rangle}{\langle f|i \rangle} \quad (1)$$

ただしここで、 $|i\rangle$, $\langle f|$ はそれぞれ初期状態と終状態を表す。Weak measurement の詳細な議論については、文献 [17, 19, 20] を参照されたい。Weak measurement によって、Hardy のパラドクス [21–24] や Leggett-Garg の不等式 [25] に代表されるような量子力学の基礎の理解をより深めることができるようになりつつある。また、量子力学の基礎論のみならず、「目に見える」新奇な現象が weak measurement で説明可能であることが分かりつつある [26–28]。本論文では、weak measurement の概念が、実は量子ビット間の相互作用推定という実用的な問題においても有用であることを示す。本研究は量子ビット間の相互作用推定における新しい手法というだけでなく、weak measurement の新しい展開という両側面がある。具体的に、weak measurement による相互作用推定手法の流れを次章で見たい。

2 Weak measurement による相互作用推定手法

本研究では、考察する系に対して以下の3つの条件を仮定する。

- 2つの量子ビットから成る系を用意する。1つのスピンを「ターゲットスピン」、もう1つのスピンを「プローブスピン」と呼ぶ。系全体のハミルトニアンは、

$$\mathcal{H}_{\text{tot}} = \mathcal{H}_t + \mathcal{H}_p + \mathcal{H}_{\text{int}} \quad (2)$$

とする。ここで第1項から第3項までそれぞれ、ターゲットスピン、プローブスピン、ターゲットスピンとプローブスピンの相互作用のハミルトニアンである。

- 1体のハミルトニアン \mathcal{H}_t , \mathcal{H}_p のダイナミクスは既知であるとする。
- ターゲットスピンを操作することにより、プローブスピンの観測ができるとする。両者のスピンの直接的な観測は必ずしも必要ではない。情報を1つのスピンからもう片方のスピんに転送した後、[29] で行われている方法でそのスピンを測定する。

近年の量子情報処理技術の進展により、以上の3つの仮定はいくつかの系で実現可能なものであり、それほど厳しい条件ではない。

これからターゲットスピンとプローブスピンの間の相互作用ハミルトニアンとして、以下のハミルトニアンを考えよう。相互作用は未知であるとする。

$$\mathcal{H}_{\text{int}} = \sum_{\mu, \nu \in \{x, y, z\}} g_{\mu\nu} (\sigma_t^\mu \sigma_p^\nu), \quad (3)$$

$$[g_{\mu\nu}] = \begin{pmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{pmatrix} =: (\mathbf{n}_x, \mathbf{n}_y, \mathbf{n}_z) \quad (4)$$

ここで σ_t^μ や σ_p^ν はそれぞれターゲットスピン、プローブスピンのパウリ行列である ($\mu, \nu = x, y, z$)。また $g_{\mu\nu}$ は対称であると仮定しよう。すなわち、 $g_{\mu\nu} = g_{\nu\mu}$ である。そうすると、独立な変数は6個となる。 $g_{\mu\nu}$ を推定する手法として、我々が提案したのは以下の通りである。

Step 1 始状態として、 $\Phi_1 = \rho_t \otimes \rho_p$ という状態を用意する。ただしここで、ターゲットスピン、プローブスピンの密度行列 ρ_t, ρ_p はそれぞれ

$$\rho_t = \frac{I + \mathbf{r}_t \cdot \vec{\sigma}_t}{2}, \quad \rho_p = \frac{I + \mathbf{r}_p \cdot \vec{\sigma}_p}{2} \quad (5)$$

で表される状態であるとする。 I は恒等演算子である。初期状態を用意している際には、相互作用 \mathcal{H}_{int} がかからないという仮定をおく。

Step 2 短い時間 δt だけ待つ。この段階で、未知のスピン・スピン相互作用による「弱い相互作用」がかかることになる。 δt の時間発展後の状態 Φ_2 は、

$$\Phi_2 = e^{-i\mathcal{H}_{\text{tot}}\delta t} \Phi_1 e^{i\mathcal{H}_{\text{tot}}\delta t} \quad (6)$$

と表される。この状態は、ターゲットスピンとプローブスピンが弱くエンタングルした状態である。ここで、ダイナミクスの中には1体のハミルトニアン $\mathcal{H}_t, \mathcal{H}_p$ からの寄与がある。

Step 3 ターゲットスピンの終状態 $\tilde{\rho}_t = \text{Tr}_p \Phi_2$ を量子状態トモグラフィを行うことで決定する。

Step 4 $\tilde{\mathbf{q}}$ 方向でプローブスピンを測定する。すなわち、

$$\text{Ex}(\mathbf{q} \cdot \vec{\sigma}_p) = \text{Tr}(P(\tilde{\mathbf{q}})\tilde{\rho}_p) \quad (7)$$

である。ただしここで、

$$\tilde{\rho}_p := \text{Tr}_t \Phi_2, \quad P(\tilde{\mathbf{q}}) := \tilde{\mathbf{q}} \cdot \vec{\sigma}_p \quad (8)$$

である。

Step 5 1体のハミルトニアン $\mathcal{H}_t, \mathcal{H}_p$ からの影響を除去する。そのために、 $\mathbf{r}_f, P(\mathbf{q})$ を以下のよう
に定義する。

$$e^{i\mathcal{H}_t\delta t}\tilde{\rho}_te^{-i\mathcal{H}_t\delta t} =: \frac{I + \mathbf{r}_f \cdot \vec{\sigma}_t}{2}, \quad e^{i\mathcal{H}_p\delta t}P(\tilde{\mathbf{q}})e^{-i\mathcal{H}_p\delta t} =: P(\mathbf{q}) \quad (9)$$

これは Trotter・鈴木分解からの帰結である [31,32]。

Step 6 制御可能な4つのパラメータ $(\mathbf{r}_i, \mathbf{p}, \delta t, \tilde{\mathbf{q}})$ を変化させて、Step 1から Step 5を繰り返す。

Step 7 Step 2と Step 3は weak measurement と見なせるから、Step 4で得られる期待値 $\text{Ex}(\mathbf{q} \cdot \sigma_p)$
は1次近似で表すことができる。すなわち、

$$\text{Ex}(\mathbf{q} \cdot \vec{\sigma}_p) \simeq \mathbf{q} \cdot \mathbf{p} + 2\delta t \sum_{\mu=x,y,z} \left[\frac{\{(\mathbf{q} \times \mathbf{n}_\mu) \cdot \mathbf{p}\}(\mathbf{r}_i + \mathbf{r}_f) \cdot \mathbf{e}_\mu}{1 + \mathbf{r}_i \cdot \mathbf{r}_f} + \frac{[\mathbf{n}_\mu \cdot \mathbf{q} - (\mathbf{n}_\mu \cdot \mathbf{p})(\mathbf{q} \cdot \mathbf{p})](\mathbf{r}_i \times \mathbf{r}_f) \cdot \mathbf{e}_\mu}{1 + \mathbf{r}_i \cdot \mathbf{r}_f} \right] \quad (10)$$

となる。 \mathbf{e}_μ は μ 方向 ($\mu = x, y, z$) の単位ベクトルである。ここで、weak value は

$$\langle \sigma_s \rangle_w = \frac{\mathbf{r}_i + \mathbf{r}_f + i(\mathbf{r}_i \times \mathbf{r}_f)}{1 + \mathbf{r}_i \cdot \mathbf{r}_f} \quad (11)$$

で与えられる。これはちょうど [33] の自然な拡張になっている。このことから、式 (10) は
weak value を自然に含む形になっていることが分かる。Step 6で複数の独立な制御パラメ
ータセットを用意することにより、式 (10) を用いて、スピン・スピン相互作用の6個の未知変
数 $[g_{\mu\nu}]$ を得ることができる。

6個の独立な線形方程式は

$$\xi = A^{-1}\zeta \quad (12)$$

のように表せることが分かる。 A は 6×6 の正方行列である。ここで、添え字 k は制御可能なパラ
メータセットのラベルとすると、

$$\xi := (g_{xx}, g_{yy}, g_{zz}, g_{xy}, g_{yz}, g_{zx})^T, \quad (13)$$

$$\zeta := \frac{[\text{Ex}(\mathbf{q}_k \cdot \sigma_p) - \mathbf{q}_k \cdot \mathbf{p}_k](1 + \mathbf{r}_{i,k} \cdot \mathbf{r}_{f,k})}{2\delta t_k}, \quad (14)$$

$$[A]_{j,k} := (\mathbf{p}_k \times \mathbf{q}_k)_\mu (\mathbf{r}_{i,k} + \mathbf{r}_{f,k})_\mu + [\mathbf{q}_k - \mathbf{q}_k(\mathbf{q}_k \cdot \mathbf{p}_k)]_\mu (\mathbf{r}_{i,k} \times \mathbf{r}_{f,k})_\nu \quad (15)$$

である。また、 T は行列の転置を表す。通常よく用いられる量子プロセストモグラフィの方法と比
較すると、我々の方法は、

- 2つの量子ビット間の相対位置を知らなくても良い [34]。
- 長時間ダイナミクスを必要としない、つまりデコヒーレンスに対してシビアではない [35]。
- Bell 測定をする必要がない [36]。

などの利点がある。

3 数値計算によるデモンストレーション

例として、ダイヤモンド中の NV 中心 [34] の相互作用推定に適用した数値計算結果を示す。ダイヤモンド中の NV 中心の場合、ターゲットスピンとプローブスピンはそれぞれ、NV の電子とそれに隣接する ^{13}C の各スピンである。数値計算によるデモンストレーションを行う限りにおいては、Step 5 から 1 体のハミルトニアン $\mathcal{H}_t, \mathcal{H}_p$ を無視してよい ($\mathcal{H}_t = \mathcal{H}_p = 0$ とした)。与えられた相互作用ハミルトニアンを用いて、Step 1 で用意した初期状態からの時間発展を計算した。 δt 後のターゲットスピンの終状態及び、プローブスピンの \mathbf{q} 方向での射影測定の期待値を計算した。ここで我々が与えた相互作用パラメータは

$$g_{\mu\nu} = \begin{pmatrix} 0.4 & -2.2 & -2.1 \\ -2.2 & 2.6 & -0.4 \\ -2.1 & -0.4 & 3.5 \end{pmatrix} \text{ [MHz]} \quad (16)$$

である。我々は表 1 のような 6 つの制御パラメータセットを用いて時間発展を計算した。

	\mathbf{r}_i	\mathbf{p}	\mathbf{q}	δt
(a)	$(0, 1, 0)^T$	$(0.77, 0.59, 0.25)^T$	$(0, 1, 0)^T$	0.294[μs]
(b)	$(0, 0, -1)^T$	$(0.18, -0.95, 0.25)^T$	$(-1, 0, 0)^T$	0.129[μs]
(c)	$(-1, 0, 0)^T$	$(0.48, -0.59, 0.65)^T$	$(0, 0, -1)$	0.085[μs]
(d)	$(0, 1, 0)^T$	$(0.25, 0.95, -0.18)^T$	$(-0.25, 0.95, 0.18)^T$	0.064[μs]
(e)	$(0, 1, 0)^T$	$(-0.25, 0.95, 0.18)^T$	$(-0.57, 0.59, -0.57)^T$	0.052[μs]
(f)	$(-0.25, 0.95, 0.18)^T$	$(0.29, 0.95, -0.10)^T$	$(-0.25, 0.95, 0.18)^T$	0.041[μs]

表 1: 本研究で用いたパラメータセット

これを用いて得られた結果を用いて得られた相互作用の「推定値」は、

$$\tilde{g}_{\mu\nu} = \begin{pmatrix} 0.56 & -2.43 & -2.22 \\ -2.43 & 3.31 & -0.22 \\ -2.22 & -0.22 & 3.48 \end{pmatrix} \text{ [MHz]} \quad (17)$$

となった。ここで誤差を以下のように定義した。

$$\langle g \rangle := \sum_{(\mu,\nu) \in \Omega} \frac{(\tilde{g}_{\mu\nu} - g_{\mu\nu})}{6}, \quad \sigma_g := \sqrt{\frac{\sum_{(\mu,\nu) \in \Omega} [(\tilde{g}_{\mu\nu} - g_{\mu\nu}) - \langle g \rangle]^2}{5}} \quad (18)$$

ここで $\Omega := \{(\mu, \nu)\} = (x, x), (y, y), (z, z), (x, y), (x, z), (y, z)$ である。誤差は $\langle g \rangle \pm \sigma_g = 0.11 \pm 0.36$ [MHz] となり、十分な精度が出ていることが分かった。

本研究でのデモンストレーションで、相互作用がかかっている時間 δt をどのように決めたか、について最後に触れる。図 1 は、式 (7) と式 (10) の右辺の差の絶対値 $\Delta \text{Ex}(\mathbf{q} \cdot \sigma_p)$ の時間発展を表している。短時間領域では $\Delta \text{Ex}(\mathbf{q} \cdot \sigma_p)$ は線形に振る舞うが、ある時間にくぼみ (矢印の部分) が生じている。これは式 (10) で無視した高次の項の影響である。今回の目的は提案した手法のデモンストレーションであることから、このくぼみの部分の時間を δt と採用した。

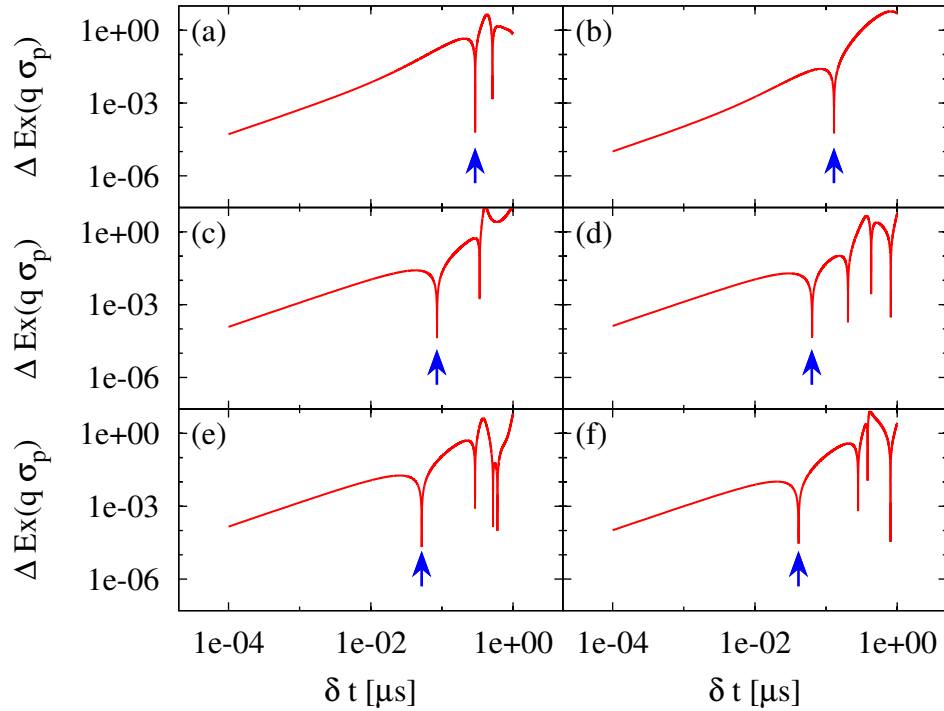


図 1: (a)-(f) はそれぞれ表 1 の (a)-(f) に対応している。矢印の位置を δt とした。

このようなパラメータの選び方は恣意的であるように思われるかもしれない。しかし実際に我々が提案した手法を用いて相互作用の推定を行う際には、多くのパラメータセットについて測定を行い、方程式を解いて推定値を求める。十分に多数のパラメータセットで測定を行い、推定値を求めるうちに、最終的にはその推定値が収束していくはずである。そのため、最終的にはここでデモンストレーションした程度の精度が出ることが分る。

4 結論と今後の課題

我々は 2 つの量子ビット間に働く相互作用を実験的に推定する新しい手法を提案した。この方法は weak measurement を基礎としており、相互作用推定の新しい手法の提案と言うだけでなく、weak measurement の実用的な応用例の 1 つを作ったとも言えるだろう。本手法では、(1) 2 つの量子ビットから成る、(2) 1 体のハミルトニアンが分かっている、(3) 片方の量子ビットを操作することにより、もう片方の量子ビットの観測ができる、の非常に緩い 3 つの条件のみを仮定している方法であり、色々な系に適用可能であると考えられる。この手法のメリットとして、(1) 2 つの量子ビット間の相対位置を前もって知っておく必要はない、(2) 長時間のダイナミクスを必要としない、(3) Bell 測定をする必要がない、という 3 点が挙げられる。

我々の提案した手法には残された課題がいくつかある。最も重要な問題としては、適切なパラ

メータセットをどのように選択するか、という問題である。いくつかのパラメータセットから得られた結果から、推定値を出し、その結果から別の適切なパラメータセットを選ぶ、いわば「学習理論」的なことができないか、を探るのは一つの課題としてあげられる。また、本手法は2つの量子ビットに限定した話である。スケーラビリティの観点から、この手法を多くの量子ビットからなる系に対しても適用可能な手法に昇華させることは大きな課題であると考えられる。これらの課題を解決することにより、本研究で提案した手法がより実用的になると考えられる。本手法の実験的な実現に期待したい。

謝辞

Yakir Aharonov 氏, Daniel Burgarth 氏, 細谷暁夫氏、小坂英男氏、丸山耕司氏、根来誠氏、James Rabeau 氏、玉手修平氏、Norman Yao 氏、吉野共広氏からは有益なコメントを頂きました。この場を借りて感謝いたします。本研究は、文部科学省によるオープンリサーチセンター整備事業から支援を受けて実施されました。また田中は、科研費 No.21840021 及び No.22340111 から支援を受けました。鹿野は科研費（特別研究員奨励費）No.21008624、学振優秀若手研究者海外派遣事業および東京工業大学 GCOE「ナノサイエンスを拓く量子物理学拠点」より支援を受けました。この場を借りて感謝申し上げます。

参考文献

- [1] Y. Shikano and S. Tanaka, arXiv:1007.5370.
- [2] Y. Shikano, S. Kagami, S. Tanaka, and A. Hosoya, to appear in AIP conference proceedings.
- [3] Y. Aharonov, D.Z. Albert, L. Vaidman, Phys. Rev. Lett. **60** (1988), 1351.
- [4] P. W. Shor, in Proc. 35th Annual Symposium on Foundations of Computer Science (IEEE Press, Los Alamitos, IL, 1994), 124.
- [5] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, U.K., 2000).
- [6] M. Nakahara and T. Ohmi, Quantum Computing From Linear Algebra to Physical Realizations (CRC Press, Boca Raton, FL, 2008).
- [7] D.P. DiVincenzo, Fortschr. Phys. **48** (2000), 771.
- [8] J.E. Wertz and J.R. Bolton, Electron Spin Resonance – Elementary Theory and Practical Applications (1986), (Chapman and Hall, New York)
- [9] M. Mohseni and D.A. Lidar, Phys. Rev. Lett. **97** (2006), 170501.

- [10] R. Blume-Kohout, H.K. Ng, D. Poulin, and L. Viola, *Phys. Rev. Lett.* **100** (2008), 030501.
- [11] A. Bendersky, F. Pastawski, and J.P. Paz, *Phys. Rev. Lett.* **100** (2008), 190403.
- [12] D. Burgarth, K. Maruyama, and F. Nori, *Phys. Rev. A* **79** (2009), 020305.
- [13] D. Burgarth and K. Maruyama, *New. J. Phys.* **11** (2009), 103019.
- [14] C. Di Franco, M. Paternostro, and M. S. Kim, *Phys. Rev. Lett.* **102** (2009), 187203.
- [15] D. Burgarth, K. Maruyama, and F. Nori, *New. J. Phys.* **13** (2011), 013019.
- [16] M.A. Fasihi, S. Tanaka, M. Nakahara, and Y. Kondo, *J. Phys. Soc. Jpn.* **80** (2011), 044002.
- [17] Y. Aharonov and L. Vaidman, in *Time in Quantum Mechanics, Vol.1*, edited by J.G. Muga, R. Sala Mayato, and I.L. Egusquiza (2008), (Springer, Berlin Heidelberg), 399.
- [18] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes* (2005), (Wiley-VCH, Weinheim).
- [19] Y. Shikano and A. Hosoya, *J. Phys. A* **43** (2010), 025304.
- [20] Y. Shikano and A. Hosoya, *Physica E* **43** (2011) 776.
- [21] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, *Phys. Lett. A* **301** (2002), 130
- [22] J.S. Lundeen, A.M. Steinberg, *Phys. Rev. Lett.* **102** (2009), 020404.
- [23] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, *New J. Phys.* **11** (2009), 033011
- [24] A. Hosoya and Y. Shikano, *J. Phys. A* **43** (2010), 385307.
- [25] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bartet, D. Vion, D. Esteve, and A.N. Korotkov, *Nature Phys.* **6** (2010), 442.
- [26] O. Hosten and P. Kwiat, *Science* **319** (2008), 787.
- [27] P.B. Dixon, D.J. Starling, A.N. Jordan, and J.C. Howell, *Phys. Rev. Lett.* **102** (2009), 173601.
- [28] D.J. Starling, P.B. Dixon, A.N. Jordan, and J.C. Howell, *Phys. Rev. A* **80** (2009), 041803.
- [29] M.V. Gurudev Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, A.S. Zibrov, P.R. Hemmer, and M.D. Lukin, *Science* **316** (2007), 1312.
- [30] P. Neumann, R. Kolesov, B. Naydenov, J. Beck, F. Rempp, M. Steiner, V. Jaxques, G. Balasubramanian, M.L. Markham, D.J. Twitchen, S. Pezzagna, J. Meijer, J. Twamley, F. Jelezko, and J. Wrachtrup, *Nature Phys.* **6** (2010), 249.

- [31] H.F. Trotter, Proc. Amer. Math. Soc. **10** (1959), 545.
- [32] M. Suzuki, Prog. Theor. Phys. **56** (1976), 1454.
- [33] S. Wu and K. Mølmer, Phys. Lett. A **374** (2009), 34.
- [34] L. Childress, M.V. Gurudev Dutt, J.M. Taylor, A.S. Zibrov, F. Jelezkko, J. Wrachtrup, P.R. Hemmer, and M.D. Lukin, Science **314** (2006), 281.
- [35] S.J. Devitt, J.H. Cole, L.C.L. Hollenberg, Phys. Rev. A **73** (2006), 052317.
- [36] M. Mohseni, A.T. Rezakhani, J.T. Barreiro, P.G. Kwiat, A. Aspuru-Guzik, Phys. Rev. A, **81** (2010), 032102.

確率過程における weak value とその応用

富田 博之

近畿大学総合理工学研究科オープンリサーチセンター

概要

確率過程におけるマスター方程式は、従属変数の変換により実時間のシュレディンガー方程式に変換される。これを用いると、2時点条件付確率が量子力学に類似の干渉効果をもつことが示される。以上はすでによく知られた話であるが、さらに2時点条件付確率を密度行列に拡張すれば非対角的な物理量を扱うことができ、量子力学で知られた weak value [1] と同様の特異な振り舞い——見かけ上の「負の確率」や期待値の異常増大など——を導くことができる。Weak value を理解する一助となろう。

1 マスター方程式の量子力学形式化

確率変数（の組）を x として、通常の条件付確率 $P(x, t | x_i, t_i)$, $t \geq t_i$ は以下の形のマスター方程式（前進方程式）に従う：

$$\begin{aligned} \frac{\partial}{\partial t} P(x, t | x_i, t_i) &= - \sum_{x'} W(x \rightarrow x') P(x, t | x_i, t_i) + \sum_{x'} W(x' \rightarrow x) P(x', t | x_i, t_i) \\ &= - \sum_{x'} L(x, x') P(x', t | x_i, t_i). \end{aligned} \quad (1)$$

ここで、 L は遷移確率 W と

$$L(x, x') = \left[\sum_{x''} W(x \rightarrow x'') \right] \delta_{x, x'} - W(x' \rightarrow x),$$

で関係づけられた行列であって、確率の保存則から

$$\sum_x L(x, x') = 0, \quad \forall x',$$

が成り立つため、 L は固有値 $\lambda_0 = 0$ の固有状態をもつ。これは定常状態、

$$P_0(x) = \lim_{t-t_i \rightarrow \infty} P(x, t | x_i, t_i),$$

に対応する。ここで $\phi_0(x) = \sqrt{P_0(x)}$ を用いて従属変数を

$$\psi(x, t | x_i, t_i) = \phi_0(x)^{-1} P(x, t | x_i, t_i), \quad (t \geq t_i) \quad (2)$$

と変換すれば, ψ は以下の「前進波動方程式」に従う:

$$\frac{\partial}{\partial t}\psi(\mathbf{x}, t) = -\sum_{\mathbf{x}'} H(\mathbf{x}, \mathbf{x}')\psi(\mathbf{x}', t). \quad (3)$$

ただし H は以下で与えられる:

$$H(\mathbf{x}, \mathbf{x}') = \phi_0(\mathbf{x})^{-1}L(\mathbf{x}, \mathbf{x}')\phi_0(\mathbf{x}'). \quad (4)$$

簡単のため当分の間, ψ 中の初期条件 (x_i, t_i) は省略している。 $\phi_0(\mathbf{x})$ は (3) の固有値 $\lambda_0 = 0$ の固有状態である。この変換のメリットは, もし H が対称 (エルミート)

$$H(\mathbf{x}, \mathbf{x}') = H(\mathbf{x}', \mathbf{x})$$

であれば, マスター方程式の固有値問題が扱いやすくなることである。この対称性は「詳細釣り合い条件」

$$P_0(\mathbf{x})W(\mathbf{x} \rightarrow \mathbf{x}') = P_0(\mathbf{x}')W(\mathbf{x}' \rightarrow \mathbf{x}),$$

または同じことであるが

$$L(\mathbf{x}, \mathbf{x}')P_0(\mathbf{x}') = L(\mathbf{x}', \mathbf{x})P_0(\mathbf{x}), \quad (5)$$

が成り立てば成り立ち, 広く期待できる条件である。この場合 H の固有値は実で, 定常状態の安定性から非負であり, ϕ_0 は基底状態である。

最も簡単な例は, 連続変数の 1 次元 Fokker-Planck 方程式,

$$\frac{\partial}{\partial t}P(x, t) = -\mathcal{L}[x]P(x, t), \quad \mathcal{L}[x] = -\frac{\partial}{\partial x} \left(F'(x) + \frac{\epsilon}{2} \frac{\partial}{\partial x} \right), \quad (6)$$

である。これは 1 次元ポテンシャル $F(x)$ の中でのブラウン運動を記述する。この場合の上記の変換は

$$\begin{aligned} P_0(x) &\propto \exp \left[-\frac{2}{\epsilon} F(x) \right], \quad \text{または} \quad \phi_0(x) \propto \exp \left[-\frac{1}{\epsilon} F(x) \right], \\ \mathcal{H}[x] &= \frac{1}{\epsilon} \left[-\frac{\epsilon^2}{2} \frac{\partial^2}{\partial x^2} + V(x) \right], \\ V(x) &= \frac{1}{2} \left[F'(x)^2 - \epsilon F''(x) \right], \end{aligned} \quad (7)$$

であり, 最終的には実時間の Schrödinger 方程式

$$-\epsilon \frac{\partial}{\partial t}\psi(x, t) = \left[-\frac{\epsilon^2}{2} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x, t),$$

が得られる。図 1 は, 準安定状態の確率論的崩壊定数が対応する Schrödinger 方程式の固有値 (第一励起状態 λ_1 : ほとんど λ_0 と縮退) で与えられることを示す例である。[2]

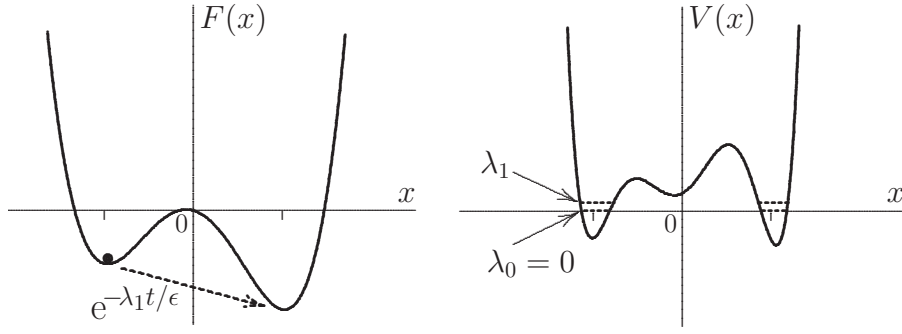


図 1: 準安定状態の確率論的崩壊過程の固有値問題

2 2時点条件付確率 (TPCP)

ここまでなら対称化することによりマスター方程式の固有値問題を慣れた量子力学の固有値問題で考えようというだけにすぎない。波動関数の干渉など量子力学独特の振舞いは2時点条件付確率 (Two Point Conditional Probability – TPCP)

$$P(\mathbf{x}, t | \mathbf{x}_i, t_i; \mathbf{x}_f, t_f), \quad t_i \leq t \leq t_f, \quad (; \text{は 'and' を表す}) \quad (8)$$

に初めて現れる。ここで, よく知られたベイズの関係式を導くのに用いられる恒等式,

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

とマルコフ性を繰り返し用いれば, TPCP は先の「波動関数」 ψ と以下で関係づけられる:

$$P(\mathbf{x}, t | \mathbf{x}_i, t_i; \mathbf{x}_f, t_f) = \frac{1}{\langle \psi_f | \psi_i \rangle} \bar{\psi}(\mathbf{x}, t | \mathbf{x}_f, t_f) \psi(\mathbf{x}, t | \mathbf{x}_i, t_i). \quad (9)$$

ここで $\bar{\psi}$ は先見的 (posterior) 条件付確率 $\bar{P}(\mathbf{x}, t | \mathbf{x}_f, t_f), t \leq t_f$ を用いて

$$\bar{\psi}(\mathbf{x}, t | \mathbf{x}_f, t_f) = \phi_0(\mathbf{x})^{-1} \bar{P}(\mathbf{x}, t | \mathbf{x}_f, t_f), \quad (10)$$

で定義される (後退) 波動関数であり, 以下の「後退波動方程式」に従う:

$$\frac{\partial}{\partial t} \bar{\psi}(\mathbf{x}, t) = \sum_{\mathbf{x}'} H^\dagger(\mathbf{x}, \mathbf{x}') \bar{\psi}(\mathbf{x}', t). \quad (11)$$

H^\dagger は H のエルミート共役 (ここでは単に転置) で, H がエルミート (対称) な今の場合, 前進・後退波動方程式で固有値と固有関数は共通である。(9) 式の分母は

$$\langle \psi_f | \psi_i \rangle = \sum_{\mathbf{x}'} \bar{\psi}(\mathbf{x}', t | \mathbf{x}_f, t_f) \psi(\mathbf{x}', t | \mathbf{x}_i, t_i), \quad (12)$$

であり, 前進・後退波動関数の重なりを表す量である。

ここで量子力学的手法を扱いやすくするため, ブラ・ケットベクトル

$$|\psi_i(t)\rangle = \{\psi(\mathbf{x}, t | \mathbf{x}_i, t_i)\}^\dagger, \quad \langle \psi_f(t)| = \{\bar{\psi}(\mathbf{x}, t | \mathbf{x}_f, t_f)\},$$

を導入すれば, (3), (11) は

$$\frac{\partial}{\partial t}|\psi_i(t)\rangle = -H|\psi_i(t)\rangle, \quad \frac{\partial}{\partial t}\langle\psi_f(t)| = \langle\psi_f(t)|H, \quad (13)$$

と書き表すことができる。そこで以後では H をハミルトニアンと呼ぶことにする。これを用いれば, 重なり積分 (内積) (12) は

$$\frac{\partial}{\partial t}\langle\psi_f|\psi_i\rangle = \langle\psi_f(t)|H|\psi_i(t)\rangle - \langle\psi_f(t)|H|\psi_i(t)\rangle = 0$$

を満たし, 途中の時間 t に依存しないことがわかる。さらに以下の極限的性質をもつ:

$$\left\{ \begin{array}{l} \text{(i)} \quad \lim_{t_f-t_i \rightarrow \infty} \langle\psi_f|\psi_i\rangle = 1, \\ \text{(ii)} \quad \lim_{t_f-t_i \rightarrow 0} \langle\psi_f|\psi_i\rangle = \phi_0(\mathbf{x}_i)^{-1}\phi_0(\mathbf{x}_f)^{-1}\delta_{\mathbf{x}_i,\mathbf{x}_f}. \end{array} \right. \quad (14)$$

また, TPCP による物理量の期待値 (Two Point Conditional Expectation – TPCE) はブラ・ケットを用いて

$$\langle A \rangle_{(i;f)} = \sum_{\mathbf{x}'} A(\mathbf{x}')P(\mathbf{x}', t|\mathbf{x}_i, t_i; \mathbf{x}_f, t_f) = \frac{\langle\psi_f(t)|A|\psi_i(t)\rangle}{\langle\psi_f|\psi_i\rangle}, \quad (15)$$

と表され, まさに量子力学の weak value の形になっていることがわかる。

このように, 前進・後退波動関数はそれぞれ閉じた線形発展方程式に従うため重ね合わせの原理が成り立つが, その積で与えられる TPCP は閉じた一つの線形発展方程式では記述できない。その結果, 重ね合わせの原理は一般に成り立たず, 波動関数の干渉に類似した振舞いが現れることが予想される。

2.1 例 1 : 1次元自由ブラウン運動

図 2 の左は初期条件 $\{x_i = \pm a \text{ at } t_i = 0\}$ のときの通常の条件付確率

$$P(x, t|x = \pm a \text{ at } t_i = 0),$$

である。太い実線は各ゲート $x = \pm a$ から出発した場合の確率であり, 細い点線は全確率である。この場合は確率の重ねあわせ (足し算) が成り立っている。

右はこれに終端条件 $\{x_f = \pm a \text{ at } t = t_f\}$ を加えた TPCP,

$$P(x, t|x = \pm a \text{ at } t_i = 0 \text{ and } t_f),$$

で, 確率の重ねあわせ (足し算) は成り立っておらず, 太い破線で示された波動関数の「干渉部分」が現れている。

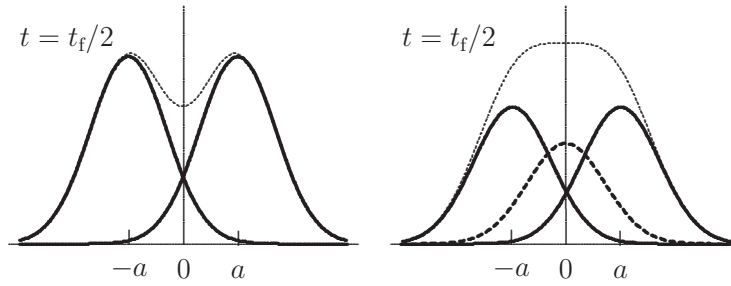


図 2: 1次元自由ブラウン運動における干渉効果 (右が TPCP)

2.2 例 2 : 確率論的 Ising モデル (2 スピン系)

交換相互作用をもつ 2 つの確率論的古典的 Ising スピン対 ($\sigma = \pm 1$)

$$\mathbf{x} = (\sigma_1, \sigma_2), \quad E(\mathbf{x}) = -J\sigma_1\sigma_2.$$

を考える。確率変数 \mathbf{x} を $(1, 1), (1, -1), (-1, 1), (-1, -1)$ の順序で並べることにし, マスター方程式 (1) の遷移確率を

$$W = \begin{pmatrix} 0 & 1 & 1 & 0 \\ p^2 & 0 & 0 & p^2 \\ p^2 & 0 & 0 & p^2 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \text{または} \quad L = \begin{pmatrix} 2p^2 & -1 & -1 & 0 \\ -p^2 & 2 & 0 & -p^2 \\ -p^2 & 0 & 2 & -p^2 \\ 0 & -1 & -1 & 2p^2 \end{pmatrix}, \quad (p = e^{-\beta J}) \quad (16)$$

で与える。これは定常状態 (温度 $\beta = 1/k_B T$ の熱平衡状態) において詳細釣り合いの条件,

$$e^{-\beta E(\mathbf{x})} W(\mathbf{x} \rightarrow \mathbf{x}') = e^{-\beta E(\mathbf{x}')} W(\mathbf{x}' \rightarrow \mathbf{x}),$$

を満たしている。平衡分布

$$P_0(\mathbf{x}) = \frac{1}{2(1+p^2)} (1, p^2, p^2, 1), \quad \phi_0(\mathbf{x}) = \frac{1}{\sqrt{2(1+p^2)}} (1, p, p, 1),$$

を用いれば, ハミルトニアンは以下のように交換相互作用と横磁場をもつ「量子スピン系」の形になる:

$$\begin{aligned} H &= \begin{pmatrix} 2p^2 & -p & -p & 0 \\ -p & 2 & 0 & -p \\ -p & 0 & 2 & -p \\ 0 & -p & -p & 2p^2 \end{pmatrix} \\ &= (1+p^2) \sigma_0 \otimes \sigma_0 - (1-p^2) \sigma_z \otimes \sigma_z - p (\sigma_0 \otimes \sigma_x + \sigma_x \otimes \sigma_0), \end{aligned} \quad (17)$$

ここで σ_x と σ_z は通常のパウリ演算子, また $\sigma_0 = I_2$ である。固有値と固有ベクトルは簡単に求められ,

$$\left\{ \begin{array}{l} \lambda_0 = 0, \lambda_1 = 2p^2, \lambda_2 = 2, \lambda_3 = 2(1+p^2) \\ |0\rangle = \frac{1}{\sqrt{2(1+p^2)}} [|\uparrow\uparrow\rangle + p|\uparrow\downarrow\rangle + p|\downarrow\uparrow\rangle + |\downarrow\downarrow\rangle] \\ |1\rangle = \frac{1}{\sqrt{2}} [|\uparrow\uparrow\rangle - |\downarrow\downarrow\rangle] \\ |2\rangle = \frac{1}{\sqrt{2}} [|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \\ |3\rangle = \frac{1}{\sqrt{2(1+p^2)}} [p|\uparrow\uparrow\rangle - |\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle + p|\downarrow\downarrow\rangle] \end{array} \right. \quad (18)$$

である。 $|0\rangle (= |\phi_0\rangle)$ が基底状態である。ここではラベルとして $\sigma = \pm 1$ の代わりによく使われる \uparrow, \downarrow を用いた。この固有ベクトル系を用いて量子力学と全く同じ手順で, 任意の初期・終端条件に対する状態ベクトル, $|\psi_i(t)\rangle, \langle\psi_f(t)|$ を計算することができる。時間 t が実であることだけが異なる。

Weak value の場合と同じく, 初期条件と終端条件が異なる場合に興味ある振舞いが見られる。例えば, $\{\mathbf{x}_i = (\uparrow\uparrow) \text{ at } t = 0 \text{ and } \mathbf{x}_f = (\downarrow\downarrow) \text{ at } t = t_f\}$, すなわち

$$P(\mathbf{x}, 0) = (1, 0, 0, 0), \quad \bar{P}(\mathbf{x}, t_f) = (0, 0, 0, 1),$$

したがって

$$|\psi_i(0)\rangle = \sqrt{2(1+p^2)} |\uparrow\uparrow\rangle, \quad \langle\psi_f(t_f)| = \sqrt{2(1+p^2)} \langle\downarrow\downarrow|$$

であるとしよう。通常通り固有ベクトル展開により

$$|\psi_i(t)\rangle = |0\rangle + \sqrt{1+p^2} e^{-\lambda_1 t} |1\rangle + p e^{-\lambda_3 t} |3\rangle, \quad (19)$$

$$\langle\psi_f(t)| = \langle 0| - \sqrt{1+p^2} e^{-\lambda_1(t_f-t)} \langle 1| + p e^{-\lambda_3(t_f-t)} \langle 3|,$$

$$\langle\psi_f|\psi_i\rangle = 1 - (1+p^2) e^{-\lambda_1 t_f} + p^2 e^{-\lambda_3 t_f} (> 0), \quad (20)$$

が得られる。図3に TPCP が示されているが, 容易に納得できる結果であり何ら不審な様子は見られず, 確率は常に非負である。

これに対して, 基底ベクトルをスピン状態 $\{|\mathbf{x}\rangle = |\sigma_1\sigma_2\rangle\}$ ではなくハミルトニアン H の固有状態 $\{|k\rangle, k = 0, 1, 2, 3\}$ にとったとき, 奇妙な振る舞いが現れる。この場合にはさらに簡単に各状態の「実現確率」を計算することができ, 以下となる:

$$\left. \begin{array}{l} P(0, t) \\ P(1, t) \\ P(2, t) \\ P(3, t) \end{array} \right\} = \frac{1}{\langle\psi_f|\psi_i\rangle} \times \left\{ \begin{array}{l} \langle\psi_f(t)|0\rangle \langle 0|\psi_i(t)\rangle \\ \langle\psi_f(t)|1\rangle \langle 1|\psi_i(t)\rangle \\ \langle\psi_f(t)|2\rangle \langle 2|\psi_i(t)\rangle \\ \langle\psi_f(t)|3\rangle \langle 3|\psi_i(t)\rangle \end{array} \right. = \frac{1}{\langle\psi_f|\psi_i\rangle} \times \left\{ \begin{array}{l} 1 \\ -(1+p^2) e^{-\lambda_1 t_f} (< 0) \\ 0 \\ p^2 e^{-\lambda_3 t_f} \end{array} \right. \quad (21)$$

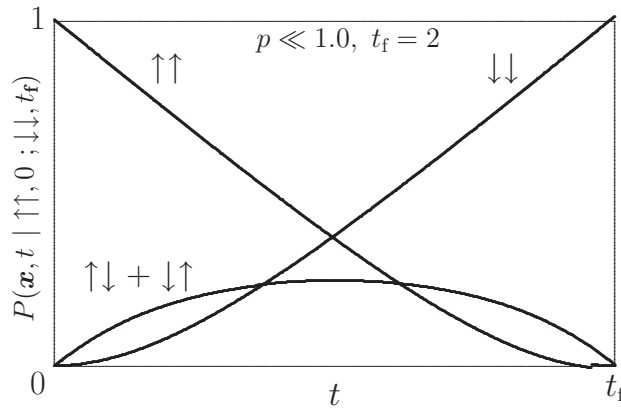


図 3: 確率論的古典 Ising スピン対の TPCP

これは各状態への射影演算子 $|k\rangle\langle k|$ (固有値は 0 と 1) の期待値であるにもかかわらず、「負の確率」が現れている。もちろん

$$\sum_{k=0}^3 P(k, t) = 1,$$

が成り立っていることは明らかである。

この結果として、ある種のオブザーバブルの期待値 (TPCE) に異常増大が現れることがある。例えば

$$M_x = \frac{1}{2}(\sigma_x \otimes \sigma_0 + \sigma_0 \otimes \sigma_x), \quad (22)$$

は固有値が $(-1, 0, 0, 1)$ であるにもかかわらず、 p および t_f が十分小さいとき、

$$\langle M_x \rangle_{(i:f)} = \frac{1}{\langle \psi_f | \psi_i \rangle} \left[\frac{2p}{1+p^2} (1 - p^2 e^{-\lambda_3 t_f}) - \frac{1-p^2}{1+p^2} (e^{-\lambda_3 t} + e^{-\lambda_3(t_f-t)}) \right] > 1,$$

となる範囲が存在し、 $p^2 t_f \ll 1$ のとき

$$\langle M_x \rangle_{(i:f)} \gg 1$$

と異常増大を示す。この異常性は分母の重なり積分から来ており、 $x_i \neq x_f$ の場合、(14) の (ii) の性質から一般的に予想されることである。

これに対して、初期状態と終端状態が同じ場合 (例えば $x_i = x_f = (\uparrow\uparrow)$)、量子力学における普通の期待値と同じで何ら異常性は現れない。

図 4 の左にこの両方の場合の期待値の振る舞いが示されている。図の右は、別の物理量 $A = \sigma_x \otimes \sigma_x$ (固有値は ± 1) の期待値である。こちらは保存量であるから時間にはよらない。 $p^2 t_f$ が小さくなるにつれて TPCE が異常増大していることが見てとれよう。

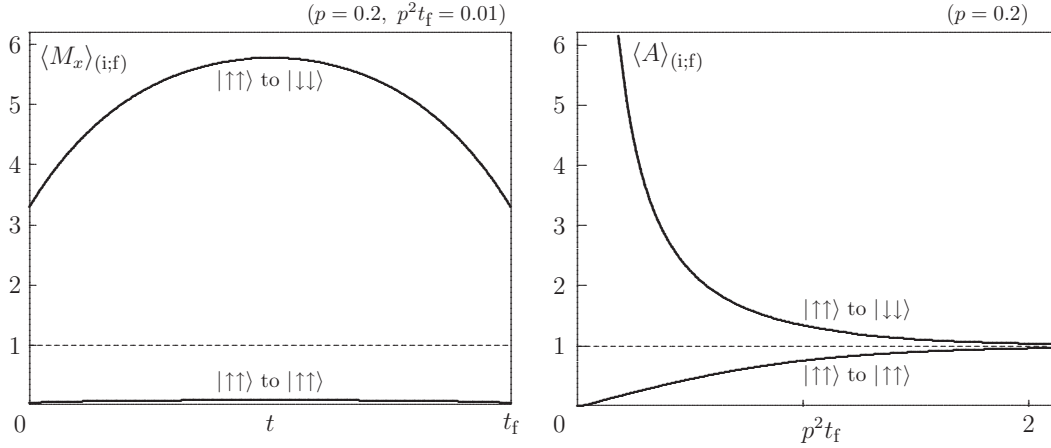


図 4: 物理量 M_x , A の期待値：異常および通常の振舞い

3 密度行列への拡張

上で見た 2 つの量 M_x , A はともにスピン状態表示では非対角的な量であって、対応する古典的物理量はない。したがって通常の TPCP だけでは期待値は計算できない。実はすでにこれを拡張した「密度行列」

$$\begin{aligned} \rho_{(i,f)}(t) &= \frac{1}{\langle \psi_f | \psi_i \rangle} |\psi_i(t)\rangle \langle \psi_f(t)| \\ &= \frac{1}{\langle \psi_f | \psi_i \rangle} \sum_{\mathbf{x}, \mathbf{x}'} \bar{\psi}(\mathbf{x}', t | \mathbf{x}_f, t_f) \psi(\mathbf{x}, t | \mathbf{x}_i, 0) |\mathbf{x}\rangle \langle \mathbf{x}'| \end{aligned} \quad (23)$$

を導入しているのである。重なり積分 $\langle \psi_f | \psi_i \rangle$ の定義, (12) により

$$\text{Tr } \rho_{(i,f)}(t) = \frac{1}{\langle \psi_f | \psi_i \rangle} \sum_{\mathbf{x}} \bar{\psi}(\mathbf{x}, t | \mathbf{x}_f, t_f) \psi(\mathbf{x}, t | \mathbf{x}_i, 0) = 1$$

が成り立っていることは明らかであり、「物理量」 A の期待値 (TPCE) は

$$\langle A \rangle_{(i,f)} = \text{Tr } \rho_{(i,f)}(t) A$$

で求められる。もちろん A が対角的な場合は古典的 TPCE に一致する。

この「密度行列」は普通、古典的確率過程では扱われてこなかった概念であるが、これ自体は古典的確率過程の範囲内 (例えばモンテカルロシミュレーション) で計算できる量である。実際、波動関数 ψ , $\bar{\psi}$ の定義と定常性および時間反転対称性

$$\bar{\psi}(\mathbf{x}', t | \mathbf{x}_f, t_f) = \psi(\mathbf{x}', t_f | \mathbf{x}_f, t) \left(= \phi_0(\mathbf{x}')^{-1} P(\mathbf{x}', t_f - t | \mathbf{x}_f, 0) \right), \quad (t \leq t_f) \quad (24)$$

すなわち,

$$\bar{P}(\mathbf{x}', t | \mathbf{x}_f, t_f) P_0(\mathbf{x}_f) = P(\mathbf{x}_f, t_f | \mathbf{x}', t) P_0(\mathbf{x}') = P(\mathbf{x}', t_f | \mathbf{x}_f, t) P_0(\mathbf{x}_f), \quad (25)$$

を用いれば，密度行列は

$$\rho_{(i;f)}(t) = \frac{1}{\langle \psi_f | \psi_i \rangle} \sum_{\mathbf{x}, \mathbf{x}'} (\phi_0(\mathbf{x}') \phi_0(\mathbf{x}))^{-1} P(\mathbf{x}', t_f - t | \mathbf{x}_f, 0) P(\mathbf{x}, t | \mathbf{x}_i, 0) |\mathbf{x}\rangle \langle \mathbf{x}'|, \quad (26)$$

で与えられる。この式を用いれば，例えばモンテカルロシミュレーションを実行する場合に，「終端条件を満たさない試行を捨てる」ということはやらなくてもよい。初期条件だけ与えてどんどん走らせ，すべてのデータを蓄積すればよい。

4 まとめ

時間発展が実であり，波動関数が常に実で非負であることを除けば，古典的確率過程を量子力学と同等の形式で記述することができ，特に2時点条件付確率には量子力学独特の振る舞いに類似の振る舞いが現れる。今回は，例として量子力学で知られた weak value と類似の異常性が現れる系を示した。この場合の weak value は，例2のように固有値問題が正確に解けて固有ベクトル展開で計算できない場合でも，例えばモンテカルロシミュレーションによって計算可能である。

量子力学における weak value は，「弱い測定」(weak measurement) という量子力学の観測理論における重要な基本的概念と密接な関係をもっているが，これの古典的対応物は目下のところ見出せていない。

参考文献

- [1] Y. Aharonov, D. Z. Albert and L. Vaidman, Phys. Rev. Letters, **60**, 1351, (1988).
- [2] H. Tomita, A. Ito and H. Kidachi, Prog. Theor. Phys. **56**, 786, (1976).

Duality of Linearity and Nonlinearity to Reveal Irreversibility in Transport Phenomena

Masuo Suzuki (TUS)
東京理科大学 鈴木増雄

要約: 不可逆性の問題は物理学の中で最も基本的な研究テーマの一つである。ここでは、輸送現象に焦点を絞って最近の自分自身の研究結果を報告する。¹⁻³⁾ 前回は、電気伝導という典型的な輸送現象におけるエンタピー生成が密度行列の対称成分 $\rho_s(t) (= \rho_0 + \rho_2(t) + \dots)$ によって記述されることを見つけた。この事実に基づいてフオン・ノイマン方程式を対称成分 $\rho_s(t)$ と反対称成分 $\rho_a(t) = \rho_1(t) + \rho_3(t) + \dots$ とに分離し、対称成分の時間発展方程式の系の内部で発生した熱を外部に取り出す緩和項をつけ加える定式化を行い、定常状態

を記述することに成功した。¹⁾ この“対称性分離の密度行列定式化”は、外からエネルギーを供給する効果を密度行列を用いて定式化することも可能にすることがわかった。²⁾ この定式化により、内部エネルギー $\langle H_0 \rangle_t$ も電界エネルギー $\langle H_E \rangle_t$ ($= -e \sum_j \langle I_{j,t} \rangle \cdot E$) も時間に依らない定常状態が密度行列で表現され、なおかつエントロピー生成 $(dS/dt)_{irr}$ ($= \int_E \cdot E / T = \int_E E^2 / T$) が正しく

与えられる理論が確立された。²⁾ この定式化を一般化するために、熱伝導の力学的定式化を行った。^{2,3)} (ラッティジャーの理論よりも直接的に久保公式を導ける新しい定式化である。)

さらに、非線形輸送現象を一般的に扱うため、久保タイプとスバーレフタイプの二つの拡張を行い、両者の関係を調べた。²⁾ 前者は、確率分布 $p(t)$ に対する統計的揺動理論であり、後者は、ハミルトニアンに対する力学的揺動理論である。後者の一般論で

密度行列の肩の部分を外場の一次まで展開した近似をすればスバーレフ理論⁴⁾になり, それをさらに密度行列そのものについて一次近似すると久保理論⁵⁾になる。

1. 非平衡応答の二つのタイプの定式化

この節では, 外部とのエネルギーのやりとりを顕微鏡には考慮しない取り扱いの一般的な構造を調べることにする。系のハミルトニアン $\mathcal{H}(t)$ を $\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_1(t)$ として, フォン・ノイマン方程式

$$i\hbar \frac{\partial \rho(t)}{\partial t} = [\mathcal{H}(t), \rho(t)]; \quad (1.1)$$

の解 $\rho(t)$ を, 初期値 $\rho_0 = \exp(-\beta \mathcal{H}_0) / Z_0$ の下に二通りの方法で表すことにする (ただし, $Z_0(\beta) = \text{Tr} \exp(-\beta \mathcal{H}_0)$)。 $\mathcal{H}_1(t) = -A \cdot F(t)$ として, 外力が時間 t に依存する場合には, 解 $\rho(t)$ は時間順序つき指数演算子で表され, 記述が煩雑になるので, 定式化の構造を説明することに主眼を置いて,

$\mathcal{F}_1(t) = \mathcal{F}_1$ と時間によらない場合を考察する。

1-1 久保タイプの定式化

時刻 $t = t_0$ まで系は ρ_0 で記述される平衡系にあり, $t = t_0$ から, 外力 \mathcal{F}_1 が (断熱因子 $e^{\varepsilon t}$; $\varepsilon \rightarrow +0$ で) 加わるものとする。 $\Delta\rho(t) = \rho(t) - \rho_0$ は, $i\hbar \frac{\partial}{\partial t} \Delta\rho(t) = [\mathcal{H}_0, \Delta\rho(t)] + [\mathcal{H}_1, \rho_0]$ (12) の解 (散密解) として次のように与えられる²⁾:

$$\begin{aligned} \Delta\rho(t) &= \frac{1}{Z(\beta)} \int_{t_0}^t e^{-\beta \mathcal{H}_0(t-t')} \int_0^\beta \dot{A}(-i\hbar\lambda)(t-t) d\lambda dt' \cdot F e^{\varepsilon t'} \\ &= \frac{1}{Z(\beta)} \int_{t_0}^t dt' \int_0^\beta d\lambda e^{-(\beta-\lambda)\mathcal{H}_0(t-t')} j(t-t) e^{-\lambda\mathcal{H}_0(t-t)} \cdot F e^{\varepsilon t'} \quad (1.3) \end{aligned}$$

ここで, $j = \dot{A} = \frac{1}{i\hbar} [A, \mathcal{H}_0]$ および $Q(t) = e^{-\frac{t\mathcal{H}_0}{i\hbar}} Q e^{\frac{t\mathcal{H}_0}{i\hbar}}$ (ただし, $\dot{A}(-i\hbar\lambda) = e^{\lambda\mathcal{H}_0} \dot{A} e^{-\lambda\mathcal{H}_0}$ と定義した。) 明らか, F の I 次までの線形近似では, $\mathcal{H}_0(t-t)$ は, \mathcal{H}_0 とし, また $j(t-t) = \exp(\frac{t-t'}{i\hbar} \mathcal{H}_0) j \exp(-\frac{t-t'}{i\hbar} \mathcal{H}_0)$ とおけるので, 久保の $\rho_1(t)$ に帰着する。任意の F に対し,

$$\begin{cases} \mathcal{J}_F = \langle j \Delta\rho(t) \rangle = \sigma_F \cdot F; \quad (t_0 \rightarrow -\infty \text{ とし } t) \\ \sigma_F = \int_0^\infty dt e^{-\varepsilon t} \int_0^\beta d\lambda \langle j(-i\hbar\lambda) j_F(t) \rangle_0; \quad j_F(t) \equiv e^{-\frac{t\mathcal{H}_0}{i\hbar}} j e^{\frac{t\mathcal{H}_0}{i\hbar}} \quad (1.4) \end{cases}$$

この解はすでに谷¹⁾によって議論されている。

1-2 スーパーレフタイプの定式化

もう一つの定式化は, (1.1)の形式解を

$$\rho(t) = \frac{1}{Z_0(\beta)} \exp\left(-\beta e^{\frac{t-t_0}{i\hbar}} \rho_0 e^{-\frac{t-t_0}{i\hbar}} \rho\right) \\ \equiv \frac{1}{Z_0(\beta)} \exp\left(-\beta(\rho_0 + \mathcal{R}(t))\right) \quad (1.5)$$

と表し, $\mathcal{R}(t)$ を求め, $\Delta\rho(t) \equiv \rho(t) - \rho_0$ を
 $\mathcal{R}(t)$ で表す方法である⁶⁾が工夫すれば,
流れ J_F は, 次のように与えられる²⁾:

$$J_F = \lim_{t_0 \rightarrow -\infty} \mathcal{R}(t) = - \int_0^\infty e^{-\epsilon t} j_F(t) \cdot F dt, \quad (1.6)$$

$$J_F = \mathcal{O}_F \cdot F; \quad \mathcal{O}_F = \int_0^\infty dt e^{-\epsilon t} \int_0^1 d\mu \langle j; j_F(t) \rangle_{\mu F}, \quad (1.7)$$

$$\langle j; j_F(t) \rangle_{\mu F} = \int_0^\beta d\lambda \text{Tr} e^{-(\beta-\lambda)(\rho_0 + \mu R_F)} j e^{-\lambda(\rho_0 + \mu R_F)} \frac{j_F(t)/Z_0(\beta)}{F/Z_0(\beta)} \quad (1.8)$$

この三重積分形式は対称性の観点から²⁾ (1.8)
興味深い²⁾。当然のことではあるが, 久保⁷⁾
の一般式(1.4)から直接的にスーパーレフタイプの
一般式(1.7)を導くこともできる²⁾。しかし, 上の導出

法からもわかる通り、2つの定式化の物理的意味合いは非常に異なることに注意された。それは近似をするときに顕著に現れる。まず2つの表式の大きな違いは、久保タイプでは密度行列そのものの時間積分(convolution-type)になっているが、スバーレフタイプでは指数演算子の肩の中($R(t)$)が時間積分になっていることである。これは、物理的には、前者が確率分布の擾動に対応し、後者が力学的擾動に対応していることを表している。

1-3 定式化の数学的問題: 非有界性

スバーレフタイプの一般的な表式(1.5)および(1.6)を見ると、断熱因子 $e^{-\epsilon t}$ の下で $t_0 \rightarrow -\infty$ の極限をとり、 $R(t)$ の極限 ρ_F が存在して(1.6)式のように時間 t に依らなければ、定常密度行列 $\rho_{st} = e^{-\beta(\rho_0 + \rho_F)}$ が求まることになる。しかし、仮りに、 ρ_0 と ρ_1 に対する適当な条件の下に ρ_F が存在しても、

それは一般に上だけでなく下にも非有界である。例えば、電気伝導の場合には、 ρ_F は電流演算子 $j = \dot{A}$ したがって運動量演算子に比例するので下にも非有界である。幸いに、 ρ_0 がその平方に比例する（下に有界で上に非有界な）項を含むので、 $\rho_0 + \rho_F$ は下には有界となり、 ρ_F が存在する場合には ρ_{st} も求まることになる。熱伝導に対する筆者の力学的定式化では、後で述べるように、 $\rho_H = -A_H \cdot E_T$ が ρ_0 よりも強く下に非有界となるため、 $\rho_0 + \rho_F$ が下に非有界となり、 ρ_F が存在したとしても、 ρ_{st} は存在しない。しかし、いずれの場合でも密度行列を場の外場で有限次まで展開した場合には、平衡演算子 ρ_0 の因子が全体として収束性を保障して有限次の応答係数は有限に求まることになる²⁾。これは、非平衡系における摂動的定式化（特に線形理論）の大きな利点の一つである²⁾（ただし、最近

グラフ中で流行のディラックのハミルトニアンで ρ_0 が表される場合には、電気伝導でも $\rho_0 + \rho_F$ が下に有界になるとは限らないことに注意したい。ディラックのハミルトニアン ρ_D の平方を粒子数 または体積で割ったものを ρ_0 とすれば、磁場中の非相対論の問題に帰着し、 $\rho_0 + \rho_F$ が下に有界な問題になる。))

2. 熱伝導の力学的定式化

電気伝導のような力学的ハミルトニアンを用いて外力の効果が表れる場合と違って、熱伝導に対する線形応答理論を作るには現象論と組み合わせる方法が用いられてきた。この節では、熱エネルギーの差を重力ポテンシャルで表現するラテンジャー¹⁾の“力学的定式化”より、さらに直接的に熱伝導に対する応答をミクロに与える定式化を提唱する^{2,3)}。臨界現象のくり込み群の理論をウィルソク⁴⁾が作ったときに出発点にしたランダウ-ギンツブルグ-ウィルソクハミルトニアンと同様に、温度勾配 $\nabla\beta(r)$ を把える

ことが出来る程度に充分大きな、しかも局所的な領域で平均化された(粗視化され、くり込まれた)ハミルトニアン $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$ を考える。ここで、 \mathcal{H}_1 が、くり込まれた粒子に対する“熱界”(thermal field) $E_T = \nabla\beta \cdot \mathbf{r} / \beta_0$ (ただし、 β_0 は特徴的な $\beta \equiv 1/k_B T$ の値であり、 E_T は長さの逆数の次元を持つ)とそれに共役な演算子 A_H とを用いて、

$$\mathcal{H}_1 = -A_H \cdot E_T \quad (2.1)$$

と表せるように A_H を求める。 \mathcal{H}_0 が

$$\mathcal{H}_0 = \sum_j h_j; \quad h_j = \varepsilon_j + u_j, \quad \varepsilon_j = \frac{1}{2m} p_j^2, \quad u_j = u(\mathbf{r}_j, \{\mathbf{r}_k\})$$

と表せるとする。 u_j は相互作用エネルギーを表す (2.2)

演算子で、 $\{k\}$ は粒子 j と相互作用している粒子の座標全体を表す。このとき、 A_H は

$$A_H = \sum_j \{h_j \mathbf{r}_j\} = \frac{1}{2} \sum_j (h_j \cdot \mathbf{r}_j + \mathbf{r}_j h_j) \quad (2.3)$$

で与えられることがわかる。^{2,3)} これに対応して、

熱流演算子 \dot{J}_H は、 $\dot{v}_j = \dot{\mathbf{r}}_j$ を用いて、

$$\dot{J}_H = \dot{A}_H = \frac{1}{i\hbar} [A_H, \mathcal{H}_0] = \sum_j \{h_j \dot{v}_j\} + \frac{1}{2i\hbar} \sum_{ij} \left\{ (\mathbf{r}_i - \mathbf{r}_j) [u_j(\mathbf{r}_j, \{\mathbf{r}_k\}), \varepsilon_i - \varepsilon_j] \right\} \quad (2.4)$$

で与えられることがわかる。^{2,3)} こうして、熱伝導の定式化が可能となる。熱流の期待値 $J_H = \langle \dot{J}_H \rangle$ は

$$J_H = \kappa E_T ; \quad \kappa = \int_0^\infty dt \int_0^\beta d\lambda e^{-\epsilon t} \langle \dot{J}_H \dot{J}_H(t+i\hbar\lambda) \rangle_0 \quad (2.5)$$

と線形近似⁵⁾ではスドレ公式⁵⁾で与えられる。非線形近似^{6,7,9)}は(1.4)と(1.7)で与えられる。²⁾

こうして、時間に依らない非平衡ハミルトンが少なくとも二つはあり、その一般性が増し、論文1のエントロピー生成²⁾と定常状態の定式化¹⁾の意義が増してきた。

3. エネルギー供給を伴う系のフォック/マツダ方程式

論文Iで定式化した対称性分離の密度行列理論^{1,3)}では、外部からのエネルギー供給の効果は取り入れられていない。そのため、 $\langle E \rangle_t$ は一定にできるが、 $\langle E \rangle_t$ は時間と共に減少する。この節では、エネルギーを外部から系に供給する効果を取り込んだ密度行列の時間発展方程式を定式化する。

*1) ここでは、 E_t までを"考える系"とみて、 E_t を維持する(電源を)働きを密度行列の時間発展方程式に取り込む定式化を試みる。

詳細は省略するが、その場合でも エンタロピー生成は

$$\left(\frac{dS}{dt}\right)_{irr} = \frac{1}{T} \text{Tr}(\tilde{D}_d(\rho_0)) \rho(t) = \frac{J_F \cdot F}{T} = \frac{\sigma_F F^2}{T} \quad (3.1)$$

で与えられる。ここで、 \tilde{D}_d は $\tilde{D}_d(\rho_0) = [\rho_0, \rho_0]_{\text{flow}}$ で定義される“動的微分演算子”である。¹⁾

4. 物理的考察

4-1. 久保タイプとスバーレフタイプの定式化と不可逆性

前節でも述べた通り、久保タイプの定式化は $\rho(t) = \rho_0 + \Delta\rho(t)$ とアンサンブルに関する擾動理論の形式になっているので、不可逆性という統計物理の本質を扱い易くしている。それに対して、スバーレフタイプの定式は、密度行列 $\rho(t)$ を平衡分布の形に似せて $\rho(t) = \exp(-\beta(\mathcal{H}_0 + \mathcal{R}(t)))$ とおいて、 $\mathcal{R}(t)$ に関する擾動理論の形式すなわち リップマン-シュウィンガーの量子擾動理論 (ハミルトニア形式) になっているので、不可逆性の取り扱いについてはより間接的になっている。えっとも無限次まで取り込めば等価な理論であるが、有限次までで近似するとその物理的意味が異なる。後者の方が非線形効果は扱い易い。

4-2. 熱伝導の力学的定式化の意義

従来の方法¹⁰⁾でも熱伝導度の κ_H の表式(2.5)は導けるが、その表式に現れる熱流演算子 j_H の表式が現象論的であったり、近似的であったりして今まですっかりしなかつたように思われる。

電気伝導と熱伝導が形式的にも同様の統計力学的取り扱いが可能となり、不可逆性の本質を表すエントピー生成^{2,3)}についても、電気伝導の場合と全く同様に密度行列の2次(一般には $\rho_s(t)$)の微分から

$$\left(\frac{dS}{dt}\right)_{irr} = \frac{\kappa E_T^2}{T} > 0 \quad (4.1)$$

と与えられることが結論される。^{2,3)}

さらに、任意の温度勾配に対してもエントピー生成は

$$\left(\frac{dS}{dt}\right)_{irr}(E_T) = \frac{\kappa(E_T) E_T^2}{T} > 0 \quad (4.2)$$

と与えられる。ここで、 $\kappa(E_T)$ は非線形熱伝導度である。

参考文献

- 1) M. Suzuki, *Physica A* **390** (2011) 1904 (論文I) および
鈴木増雄, 『線形応答とエントロピー変化(生成)-可逆性の熱統計学的解明』
(数研講究録1705(2010年))
- 2) M. Suzuki, *Physica A* (準備中). (数研講究録1705(2010年))
- 3) M. Suzuki, First-Principle Derivation of Entropy Production in
Transport Phenomena, Proceedings of STATPHYS-KOLKATA VII,
J. of Phys. Conf. ser. (IOP, UK) edited by B. K. Chakrabarti (2011).
- 4) D.N.Zubarev, 1971 *Nonequilibrium Statistical Mechanics* (Nauka).
- 5) R.Kubo, 1957 *J.Phys.Soc.Jpn.* **12** 570.
- 6) M. Suzuki, *Prog. Theor. Phys.* **100** (1998) 475 (量解析).
- 7) K. Tani, *Prog. Theor. Phys.* **32** (1964) 167.
- 8) J.M. Luttinger, 1964 *Phys.Rev.* **135A** 1505. および
A.Einstein, 1912 *Ann.Physik* **38** 443.
R.C.Tolman and P.Ehrenfest, 1930 *Phys.Rev.* **36** 1791.
H.Mori, 1958 *Phys.Rev.* **112** 1829.
- 9) S.J.Miyake and R.Kubo, 1962 *Phys.Rev.Lett.* **9** 62.
H.Nakano, 1993 *Int.J.Mod.Phys.* **B7** 2397.
See also M.Ichiyanagi, 1986 *Prog.Theor.Phys.* **76** 37.
K.Kawasaki and J.D.Gunton, 1973 *Phys.Rev.* **A8**, 2048.
M.Suzuki, 1975 *Prog.Theor.Phys.* **53** 1657; *ibid* **55** (1976) 383, 1064,
J.Phys. Soc. Jpn. **69** (2000) Suppl. A.
T.Fujii, 2007 *J.Phys.Soc.Jpn.* **76** 044709.
T.S.Komatsu, N.Nakagawa, S-I.Sasa and H.Tasaki, 2009 *J.Stat.Phys.* **134** 401,
S.Yuan, M.I.Katsnelson and H.De Raedt, 2009 *J.Phys.Soc.Jpn.* **78** 094003.
S-H.Chong, M.Otsuki and H.Hayakawa, 2010 *Phys.Rev.* **E81** 041130
S.Nakamura et al., 2010 *Phys.Rev.Lett.* **104** 080602, and references cited therein.
K.Saito and Y.Utsumi, 2008 *Phys.Rev.* **B78** 115429
- 10) R. Kubo, M. Yokota and S. Nakajima,
J. Phys. Soc. Jpn. **12** (1957) 1203.

共形不変な電磁放射場が満たす“双対”の幾何力学構造について

佐久間弘文*、小嶋泉**

* (独) 海洋研究開発機構横浜研究所

** 京都大学数理解析研究所

概要

電磁気学の基礎式として知られる **Maxwell** 方程式は、変位電流という数学的に整合的な量を導入する事により、電場と磁場を統一する電磁場としての表現形式を与える事に成功した。ここでは、まず、流体力学的な古典場の表現方法を用いて、ヌル測地線と関連付けられる渦（幾何）力学の存在を議論し、それが真空中を伝わる電磁波と同型である事を示す。電磁波に関する **Lorentz gauge** の中に現れるスカラー場は **Sugawara form** の **mass-less scalar field** と見做す事ができ、このスカラー場を上記の渦力学を構成する一要素と見做す事によって、電磁場とそれに同型の渦の場を力学的に結合できる事を示す。更には、この渦（2階の交代テンソル）のテンソル積から作られる4階のテンソルは **Riemann** の曲率テンソルが満たす性質を全て満たす事を証明し、それ故に、電磁場に加わったこの双対構造としての新たな自由度は、真空中において重力場のエネルギーと曲率を運ぶ自由度を与え得る事も示す。すなわち、このスカラー場は、電磁場と重力場放射を結合するという意味において、電場と磁場を結びつけた変位電流と同様の役割をになう存在であることを議論する。

目次

1. はじめに
2. ヌル測地線と渦力学
3. 電磁場との比較
4. 偏光に対応する量及び時空の基底ベクトルを与える渦幾何力学
5. **Riemann** の曲率テンソルとの関係
6. 双対性に基づく電磁・重力エネルギー放射
7. 付録
8. 参考文献

1. はじめに

古典電磁場と流体力学とは、場の理論として少なからぬ類似的側面を持つ為に、その点に関する研究については古くから数多く行われ、枚挙に遑が無い状態であると言えるが、電磁気学を独自の視点から“連続体力学”の一つとして再構築できる事を示した今井[1]の仕事をまずここに言及しておく事は、以下に展開する電磁波の相対論的流体力学モデルの物理的動機付けに対しその意義は大きいと思われる。本小論文は、研究会のテーマである、「量子科学における双対性とスケール」に対して、これまでに報告されていない著者独自の視点から、4次元時空の因果律を規定する上で重要な役割を担う（4次元時空のみで）共形不変な電磁場放射とそれに付随するヌル測地線に注目し、電磁場と重力場放射の統一的記述の試みという未知の課題に対する一つの整合的モデルを、電場と磁場との双対性をヌル測地線の幾何学を用いて数学的に拡張する事により示すものである。

また、量子論を古典論と比べた場合、その特質を際立たせるものとして、物質の波動と粒子の二重性というものがある。量子論とはもちろん異なる文脈ではあるが、この「波動」と「粒子」という概念それ自体は、流体という古典力学的対象の中でも自然に「波動」と「流体粒子」として共存している。特に、量子論の教科書の初めには、古典論と対比する形での **Einstein – de Broglie** 関係式が出て来て、エネルギーの離散化と共に、1粒子当たりのエネルギー（運動量）が振動数（波数）に比例する事が示される。前者はもちろん量子論に固有な事柄ではあるが、後者は古典論からもある程度類推される1例として、この小論文の主題を議論する為に導入した流体力学モデルをここに紹介する。紙面の制約上、数学的展開の詳細は、必要に応じ、その概要を末尾の付録に記載したが、一部は割愛した。この割愛部分の詳細に興味のある方は、文献の一つに挙げた現在準備中の著者の論文を参考にして頂きたい。

2. ヌル測地線と渦力学

測地線上の座標を x^a 、パラメータ s を測地線に沿った長さを表すとすると、

$$x^a = x^a(s) ; \quad U^a \equiv dx^a/ds ; \quad ds^2 = g_{ab} dx^a dx^b \quad (1a, b, c)$$

測地線の式は

$$d^2 x^a / ds^2 + \Gamma_{bc}^a (dx^b / ds) (dx^c / ds) = 0 \quad (1d)$$

となるが、(1b) で定義される U^a の共変成分表示を使うと、(1d) は

$$U^b \nabla_b U_a = 0; \quad \nabla_b U_a \equiv \partial_b U_a - \Gamma_{ab}^c U_c \quad (\text{Levi-Civita connection}) \quad (1e, f)$$

となり、(1e) は更に以下の様に変形される。

$$U^b \nabla_b U_a = U^b (\nabla_b U_a - \nabla_a U_b) + \nabla_a (U^b U_b / 2) = 0 \quad (1g)$$

(1c) により、(1g) の右辺第二項に現れるベクトル U^b の長さは ± 1 となり、ヌル測地線の場合は s とは異なる適当なパラメータを用いると、0 となる事が示るので、最終的に (1g) は、本小論文における “canonical form” と呼ぶべき以下の形を取る。

$$S_{ab} U^b = 0; \quad S_{ab} \equiv \nabla_b U_a - \nabla_a U_b = \partial_b U_a - \partial_a U_b = -S_{ba} \quad (2a, b)$$

(1g) と (2a) から、長さが一定の傾度ベクトルは測地線を表す一つの解である事が容易に見て取れる。

次に、考察する対象を測地線から理想流体へと大きく移す事にする。理想流体の相対論的支配方程式は、Landau[2] によると、

$$\zeta_{ab} v^b = T \nabla_a \eta; \quad \zeta_{ab} \equiv \nabla_b (w v_a) - \nabla_a (w v_b) \quad (3a, b)$$

となり、ここに、 T は絶対温度、 η と w は、それぞれ **specific entropy** 及び **enthalpy**、(3b) の左辺で定義される交代テンソルは、相対論的渦度テンソルである。一般に、流体力学においては、エントロピー分布が一定の順圧流体とそうでない傾圧流体という分類が行われ、前者の場合の支配方程式は、(2a) と同型になる事がわかる。また、この場合、流体力学の **H(Hamiltonian)**-構造の研究において、速度場を以下の形で表現する

$$v_a = \lambda \nabla_a \phi + \nabla_a \chi \quad (4)$$

Clebsch Parametrisation (CP) は、Lamb[3] による流体力学の古典的教科書以来数多くの研究がなされていて、以下に見る様に、本小論文においても **CP** は非常に重要な役割を担うものである。しかし、ここでは式 (4) の右辺の第 2 項は本質的なものではないので、以下この項を省略したものを **CP** と呼ぶことにする。式 (1g, 2a) において、測地線の式は傾度ベクトルによって記述される事を見たが、**CP** を使う

と傾度ベクトルではない表現も存在する事を示す事ができる。

その事を示す為に、電磁放射場の幾何光学的な記述を考える。すなわち、光の伝搬をヌル測地線に沿って動く無数の古典的粒子と考える。このモデルにおいては、mass-less の Klein Gordon 方程式：

$$g^{ab}\nabla_a\nabla_b\phi=0 \quad (5)$$

を満たすスカラー場 ϕ を用いて以下に定義するヌルベクトルによって“光粒子”の速度ベクトルが記述できる。

$$C_b \equiv \nabla_b \phi ; \quad C^b C_b = 0 \quad (6a, b)$$

ここで、更に CP で表現されるベクトル

$$U_a = \lambda \nabla_a \phi \quad (7)$$

を考え、それをヌル測地線が与える条件式 (1g, 2a) に代入し、(6b) を利用すると、

$$S_{ab}U^b = 0 \rightarrow C^b \nabla_b \lambda = 0 : C^b L_b = 0 \quad (L_b \equiv \nabla_b \lambda) \quad (8a, b)$$

となる事が容易に導かれる。(ここでは二つのベクトル C^b と L^b が一次独立な非自明な場合のみを考える) 上記の二つの条件式 (6b, 8b) は、以下の様に、幾何的な直交条件という意味づけと、CP に現れる 2 変数 ϕ と λ はヌル測地線に沿って移動する“光粒子”のラグランジュ的ラベルと見做す事ができるという二つの解釈を許すものとなっている。

直交条件	流体力学的解釈 (ラグランジュラベル)
$C^b C_b = 0$	$C^b \nabla_b \phi = 0$
$C^b L_b = 0 \rightarrow L_b$ は space-like	$C^b \nabla_b \lambda = 0$

式 (1g, 2a) の導出に関して、既に述べた事ではあるが、長さが一定の傾度ベクトルは測地線の式を自動的に満たすので、式 (7) に現れる CP の変数の一つである λ は空間的 (space-like) 測地線群のパラメータと見做す事ができる。すなわち、計量に関する signature を (+1, -1, -1, -1) とすると、

$$g^{ab}\nabla_a\lambda\nabla_b\lambda=L_bL^b=-1 \quad (9)$$

となり、(9) は傾度ベクトルの大きさに関する条件であるのに対し、(8a, b) はその方向に関する条件であるので、両者は矛盾する事なく指定でき、(9) は後述する測地線の渦力学から導かれるエネルギー・運動量テンソルの保存に関して重要な役割を果たす事が示される。以上の議論を簡単にまとめると、ヌル及び空間的測地線の表示が与える二つのパラメータを式 (7) の CP のペアーパラメータと解釈して定義される速度場からは、ヌル測地線に関する渦力学を構築できるという事になる。

3. 電磁場との比較

電磁場を記述する交代テンソル F^{ab} は、4元ベクトルポテンシャル A^b を用いて、

$$F_{ab} \equiv \nabla_b A_a - \nabla_a A_b \quad (10a)$$

と定義され、エネルギー・運動量テンソルの一般形は

$$T_a^b = -F_{ad} F^{bd} + g_a^b F_{cd} F^{cd} / 4 \quad (10b)$$

となるが、ここで問題にしている電磁放射場に関しては、右辺の第2項はゼロとなり、これは電場と磁場の振幅は同じ大きさであるという事に対応している。3次元のベクトルとして定義される電場と磁場をそれぞれ \tilde{E} と \tilde{M} として表現すると、

$$\tilde{E} = (F_{01}, F_{02}, F_{03}); \quad \tilde{M} = (F_{23}, F_{31}, F_{12}) \quad (11a, b)$$

となり、電磁放射場のエネルギー・運動量を表現する Poynting vector P^b は

$$\tilde{P} \equiv \tilde{E} \times \tilde{M}; \quad P_b P^b = 0 \quad (11c)$$

で定義され、 F_{ab} と P^b は3頁の式 (2a, b) に関して述べた “canonical form” を満たす。

$$F_{ab} P^b = 0 \quad (12a)$$

(2a) 及び (12a) の左辺を交代行列と列ベクトルの積と見做すと、ベクトル P^b が非自明なゼロでない成分を持つ為の条件として、交代行列の行列式がゼロであるという帰結が導かれる。この条件は、電磁放射場においては電場と磁場が直交しているという良く知られた条件： $\tilde{E} \perp \tilde{M}$ に対応するものであり、(2a) に対しても全く同様な条件

$$S_{01}S_{23} + S_{02}S_{31} + S_{03}S_{12} = 0 \quad (12b)$$

が導かれる。この条件 (12b) は、**CP** の定義式である式 (7) が (12b) と等価である以下の恒等式

$$S_{ab}S_{cd} + S_{ac}S_{db} + S_{ad}S_{bc} = 0 \quad (12c)$$

を自動的に満たす事からも導かれる。これは、**Hugget & Tod [4]** において言及されている以下の定理の帰結によるものである。

$$S_{01}S_{23} + S_{02}S_{31} + S_{03}S_{12} = 0 \Leftrightarrow S_{ab} \text{ は二つのベクトルの外積} \quad (12d)$$

実際、**CP** の定義式 (7) から得られる S^{ab} は、 C^a と L^a との外積で定義される。

次に、電磁放射場のエネルギー・運動量テンソルに準じて、**CP** の渦力学に対する同様のテンソルを以下の様に定義する。

$$\hat{T}_a^b = -S_{ad}S^{bd} = -(C_a L_d - L_a C_d)(C^b L^d - L^b C^d) = -(L_d L^d)C_a C^b = C_a C^b \quad (13a)$$

上の式変形の最後では空間的な測地線の式 (9) を用いた。**Klein Gordon** 方程式を表す式 (5) はヌル測地ベクトル C^a の発散がゼロであるという事を意味しているのでこの事実を使うと、電磁放射場と同様に、上記のテンソルの発散がゼロである事が容易に見て取れる。すなわち、

$$\nabla_b \hat{T}_a^b = -\nabla_b (S_{ad}S^{bd}) = -S_{ad} \nabla_b S^{bd} = (\nabla_b C^b)C_a + C^b \nabla_b C_a = 0 \quad (13b)$$

Maxwell の電磁理論においては、真空中における (10b) の発散がゼロとなる根拠は、真空中に電流が存在しない条件

$$J^d = \nabla_b F^{bd} = 0 \quad (13c)$$

である。(13b) の左から3番目の式においても同様な条件を採用する事はできるが、それは (13b) がゼロとなる充分条件ではあるが必要充分条件ではない。ここでは紙面の制約上詳細な議論を展開する事はできないが、この点に関する議論の要点を以下にまとめる事とする。まず、代数的見地のみからすると (13b) を満たすテンソル S^{bd} の発散は必ずしもゼロではなく、一般に

$$\nabla_b S^{bd} = aC^d + K^d \quad (14a)$$

という形に書ける。ここに、 a はスカラー、 K^d は空間的なベクトルで、 C^d が満足する直交関係 (2a) と同じ関係を満たすものである。しかし、更に K^d が満たすべき条件を詳細に調べると最終的には、少し制約された以下の形

$$\nabla_b S^{bd} = aC^d \quad (14b)$$

が得られる事が示される (Sakuma and Ojima[5])。 (14b) の左辺の添え字 d に関する発散は恒等的にゼロとなるので、その事及び簡単な計算より、以下の2式が導かれる。

$$C^d \nabla_d a = 0 ; \quad C^d \nabla_d S_{ab} = 0 \quad (15a, b)$$

(15a) は特に、スカラー a が CP の2変数 ϕ と λ のある関数である事を示唆するものであり ((6a, 8a) を参照)、(15b) はテンソル量 S^{ab} もヌル測地ベクトル C^d に沿って移流されている事を示すものである。議論のこの段階に至って、次に指摘する事実は、電磁放射場と CP のヌル測地渦力学を比較する上で特に重要となる。それは、式 (7) に現れる変数 λ について、 λ を ϕ と λ の任意関数として定義される新しい $\Lambda = \Lambda(\lambda, \phi)$ で置き換えても上で展開した議論 (例えば条件 (8a, b)) はそのままの形で有効であるという事である。 $a = a(\lambda, \phi)$ と λ を $\Lambda = \Lambda(\lambda, \phi)$ で置き換えても理論の本質部分是不変であるという二つの事から、適当に選択した $\Lambda = \Lambda(\lambda, \phi)$ に対しては、電磁放射場と同様に (14b) の右辺をゼロとする事が可能であり、この証明は末尾の付録に示す通りである。以下の議論において λ は、その様に調整されたものとする。

4. 偏光に対応する量、並びに時空の基底ベクトルを与える渦幾何力学

以上で示した様に、CP を通して定義されるヌル測地線の渦力学は電磁放射場と同型な構造を有している。従ってそれは、電磁波にとって重要な偏光という性質も当然含むものになっているはずである。実際、以下に見る様に、CP の渦力学における“偏光”は時空の幾何的構造を渦力学と結び付ける上で非常に重要な役割を担うものとなっている。ここでは、問題の本質的部分を簡明に説明する事を目的として、平坦な時空を進む平面波を考え議論するが、用いる数学は ϕ に対しては波動を初等的な三角関数で表したものであり、また λ に対しては時空間座標の一次関数であるので (詳しくは参考文献[5]に記載)、細かな式の展開は省き、重要な結果を図で示す事にする。

まず、座標系として Lorentz 系 : $(x^i; 0 \leq i \leq 3)$ をとり、平面波は x^1 の方向へ進んでいるものとする。これまでの議論で導入された CP に関する二つのベクトルのうち C^a の空間成分は x^1 の正の方向へ、 L^a は双対性の図 1 に示す x^1 と x^3 とで張る平面内に存在する配置になっているものとする。式 (2a) と電磁放射場の式 (11a, b) を参照にして、図においては二つ

のベクトル C^a と L^a の共変空間成分の外積を \tilde{M}_s と表して、図における負符号は、この値が導入した Lorentz 系では負になる事を示している。 C^a と L^a の一組が表す波動は、電磁場の場合の平面偏光に相当するものとなっている。円偏光は、これとは異なるある平面波を元の波に重ね合わせる事で得られる。図 1 で示されている * が付いた量が、円偏光を作る為のもう一つの平面波である。* C^a は C^a と空間成分としては同じベクトルではあるが、時間方向に位相が 1/4 波長ずれているベクトルで、* L^a は、 x^1 と x^2 とで張られる平面内の空間的なベクトルである。量子力学的には、この左回り、あるいは右回りの円偏光が光子のスピンと対応している。この意味において、CP の渦力学においても、この円偏光の状態を平面偏光に比べより基本的な状態と見做せば、その状態が時空の次元と同じ上記の四つの“基底ベクトル”で表現できる事は興味深い事実である。また、この四つのベクトルは以下の、直交関係を満たす事が示される事も付け加えておく。

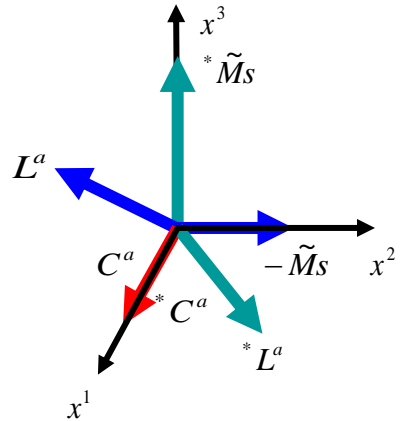


図 1 : 双対性

* $C_a C^d = 0$; $L_d C^d = 0$; * $L_d C^d = 0$; * $C_d L^d = 0$; * $C_d * L^d = 0$; * $L_d L^d = 0$ (16)

CP の渦力学において特に注目すべき特質は、(13a) で与えられるエネルギー・運動量テンソルの二価表現

$$\hat{T}_a^b = -S_{ad} S^{bd} = \rho C_a C^b \tag{17}$$

である。ここに、“密度” ρ はセクション 3 の最後で説明した変換 $\lambda \rightarrow \Lambda$ によって導入されるスカラー量である。電磁放射場の Poynting vector P^b に準じて CP の渦力学においても同様のベクトル量 Π^b を考える事ができる。ここでは議論を簡単

にするため、双対性の図 1 の中で表される一組のペアー (C^a, L^a) から生成される平面偏光の波について考え、それから式 (11c) に準じて導かれる Π^b のスピノール表現を定義する以下の式[4]でその成分を表すと、 $\Pi^2 = \Pi^3 = 0$ なので

$$\sqrt{2}\Psi(\Pi^a) \equiv \begin{pmatrix} \Pi^0 + \Pi^3 & \Pi^1 + i\Pi^2 \\ \Pi^1 - i\Pi^2 & \Pi^0 - \Pi^3 \end{pmatrix} = \begin{pmatrix} (S_{03})^2 & S_{03}(-S_{31}) \\ S_{03}(-S_{31}) & (S_{03})^2 \end{pmatrix} \quad (18a)$$

となり、また、 $|S_{03}| = |S_{31}|$ なのでエネルギー・運動量テンソルは上式と同じ形の

$$\hat{T}^{ab} = \begin{pmatrix} (S_{03})^2 - S_{03}S_{31} & \\ -S_{03}S_{31} & (S_{03})^2 \end{pmatrix} = \begin{pmatrix} \rho C^0 C^0 & \rho C^0 C^1 \\ \rho C^0 C^1 & \rho C^1 C^1 \end{pmatrix} \quad (18b)$$

となる。上の二つの式は、 Π^b の未定定数を含まないスピノール二価表現と考えられ、この二価表現に現れる四つの成分：

$$(\sqrt{\rho}C^0, \sqrt{\rho}C^1, -S_{31}, S_{03}) \quad (19)$$

は互いに直交する時空の基底ベクトルとなっている。(19)において、第 1、2 成分のペアーと、第 3、4 成分のペアーとでは、見かけ上表式は異なるが、“密度” ρ はベクトル L^a の大きさと解釈されるので ($\lambda \rightarrow \Lambda$ に対する (9) の一般化)、その意味で前者は、 C^a と L^a との (外) 積で定義される後 (式 (13a) 参照) と同形であるという事が、スピノール二価表現 (18b) の意味するところである。以上の関係をまとめると

測地線の幾何学的側面 ← 幾何・力学対応 → CP の渦力学的側面

$$(C_a, L_a, {}^*C_a, {}^*L_a)$$

$$(\sqrt{\rho}C^0, \sqrt{\rho}C^1, -S_{31}, S_{03})$$

という図式が認められ、(19) は 4 次元リーマン擬多様体の局所直交基底を与える underlying structure (ベクトル L^a の方向は一意的には定まらない) としての (16) から外積の幾何学を通じて定義される渦力学となっていて、それ自体が (16) と同様に時空の直交基底を与えるという“渦幾何・力学”構造を備えている。最後に式 (17) の二価表現の最後の“粒子的”表現に着目し、密度 ρ を単位体積における“粒子数”と形式的に翻訳すると、上で考えた単色平面波に対しては、一波長当たり且つ一粒子当たりのエネルギーおよび運動量は波の振動数と波数に比例するという Einstein – de Broglie 関係の古典版が出る事を指摘しこのセクションを結ぶ。

5. Riemann の曲率テンソルとの関係

重力場放射との対応を調べる為に、以下に定義される4階のテンソルを考える。

$$\hat{S}_{abcd} \equiv S_{ab}S_{cd} \quad (20a)$$

テンソル S_{ab} の交代性より直ちに、

$$\hat{S}_{bacd} = -\hat{S}_{abcd} \quad ; \quad \hat{S}_{abcd} = -\hat{S}_{abcd} \quad ; \quad \hat{S}_{cdab} = \hat{S}_{abcd} \quad (20b, c, d)$$

が導かれ、また (12c) から直接的に第一ビアンキの恒等式

$$\hat{S}_{abcd} + \hat{S}_{acdb} + \hat{S}_{adbc} = 0 \quad (20e)$$

が得られる。次に、第二ビアンキの恒等式の成否を確かめる為に以下の量を考える。

$$J_{abcde} \equiv \nabla_a (\hat{S}_{bcde}) + \nabla_b (\hat{S}_{cade}) + \nabla_c (\hat{S}_{abde}) \quad (20f)$$

(20f) に (20a) を代入し、恒等式 $\nabla_a S_{bc} + \nabla_b S_{ca} + \nabla_c S_{ab} = 0$ を利用すると

$$J_{abcde} = (S_{bc} \nabla_a + S_{ca} \nabla_b + S_{ab} \nabla_c) S_{de} \quad (20g)$$

を得る。 S^{ab} の Hodge dual $*S^{ab}$ [4] を考え、局所 Lorentz 系で $*S_{(L)}^{ab}$ は $S_{ab}^{(L)}$ を用いて表現できる事を使うと

$$(*S_{(L)}^{ab}) = \begin{pmatrix} 0 & S_{23}^{(L)} & S_{31}^{(L)} & S_{12}^{(L)} \\ -S_{23}^{(L)} & 0 & S_{03}^{(L)} & -S_{02}^{(L)} \\ -S_{31}^{(L)} & -S_{03}^{(L)} & 0 & S_{01}^{(L)} \\ -S_{12}^{(L)} & S_{02}^{(L)} & -S_{01}^{(L)} & 0 \end{pmatrix} \quad (21a)$$

となる。ここで θ を任意のスカラー場を表すものとして、その傾度ベクトルと (21a) との内積を作ると

$$(*S_{(L)}^{ab} \partial_b \theta) = \begin{pmatrix} + \{S_{23}^{(L)} \partial_1 \theta + S_{31}^{(L)} \partial_2 \theta + S_{12}^{(L)} \partial_3 \theta\} \\ - \{S_{23}^{(L)} \partial_0 \theta + S_{30}^{(L)} \partial_2 \theta + S_{02}^{(L)} \partial_3 \theta\} \\ - \{S_{31}^{(L)} \partial_0 \theta + S_{10}^{(L)} \partial_3 \theta + S_{03}^{(L)} \partial_1 \theta\} \\ - \{S_{12}^{(L)} \partial_0 \theta + S_{20}^{(L)} \partial_1 \theta + S_{01}^{(L)} \partial_2 \theta\} \end{pmatrix} \quad (21b)$$

となり、各成分は統一的に $\pm (S_{bc}^{(L)} \partial_a \theta + S_{ca}^{(L)} \partial_b \theta + S_{ab}^{(L)} \partial_c \theta)$ と表現される。 S_{ab} の定

儀式 (2b) を C_a と L_a とで表した: $S_{ab} = C_a L_b - L_a C_b$ を使うと、(21b) から

$$\left(S_{bc}^{(L)} \partial_a + S_{ca}^{(L)} \partial_b + S_{ab}^{(L)} \partial_c \right) \phi = 0 ; \left(S_{bc}^{(L)} \partial_a + S_{ca}^{(L)} \partial_b + S_{ab}^{(L)} \partial_c \right) \lambda = 0 \quad (21c, d)$$

が得られ、この式と (15b) をここで導入した局所 Lorentz 系で評価した式

$$C^d \partial_d S_{ab}^{(L)} = 0 \rightarrow S_{ab}^{(L)} = S_{ab}^{(L)}(\lambda, \phi) \quad (21e, f)$$

を (20g) 代入することにより、最終的な結論

$$J_{abcde} \equiv \nabla_a (\hat{S}_{bcde}) + \nabla_b (\hat{S}_{cade}) + \nabla_c (\hat{S}_{abde}) = 0 \quad (22)$$

を得る。(20b, c, d)、(20e) 及び (22) から (20a) で定義された 4 階のテンソルは Riemann の曲率テンソルが満たす性質を全て満たす事が示される。

6. 双対性に基づく電磁・重力エネルギー放射

真空中における電磁波の伝搬は、4 元ベクトルポテンシャル A^a を用いて

$$0 = \nabla_b F^{bc} = -g^{cd} \nabla_d (\nabla_e A^e) + (g^{de} \nabla_d \nabla_e A^c + R_d^c A^d) \quad (23a)$$

と表される。通常 (23a) においては、Lorentz gauge: $\nabla_e A^e = 0$ を課して

$$0 = g^{de} \nabla_d \nabla_e A^c + R_d^c A^d \quad (23b)$$

とするが、ここでは、もう一つの可能性として、

$$0 = (g^{de} \nabla_d \nabla_e A^c + R_d^c A^d) ; \nabla_b F^{bc} = -g^{cd} \nabla_d \phi ; \phi \equiv \nabla_e A^e \quad (24a, b, c)$$

を考察する。(24b) については、 F^{bc} の交代性から添え字 c についての発散がゼロとなる為に、(24b, c) で定義された ϕ は式 (5) と同じ Klein Gordon を満たす。従って前節の CP の渦力学で導入した ϕ を (24c) で定義される電磁場の縦モードと同定する事が可能となる。このスカラー場 ϕ からは式 (17) で示す力学系が構築され、その構造は量子論における Sugawara form に付随する mass-less scalar field や相対論/Twistor 理論の研究にける shear-free null congruence[6] の概念と密接に関係するものである。(24c) による ϕ の同定は、電磁放射場の F^{ab} と S^{ab} とを結合させ、 F^{ab} に対応する Poynting vector と共に、 S^{ab} に対応する新たなエ

エネルギー・運動量を運ぶ量としての **Poynting vector** の存在が許容される事になる。前節で見た様に、後者は **Riemann** の曲率テンソルと同様の性質を満たすため、この新たな自由度は重力場放射を担うものとの解釈ができる。変位電流が電場と磁場を結び付け、真空中を伝わる電磁波を予測した様に、(24c) と **CP** の渦力学は電磁・重力場放射の存在を予測する。この仮説の重要な物理的側面は、4元ベクトルポテンシャル A^a に付随するスカラー場 ϕ の実在性である。この点に関しては、本小論文の内容と必ずしも同様な文脈で議論されているわけではないが、外村等[7]による **AB** 効果の実験的証明は、これまでの **Heaviside** の解釈から抜け出し A^a の実在性の可能性を示す大きな一歩であり、その意味において、ここで展開された仮説と関係する意義深いものである。

7. 付録

式 (7) に対する λ を $\Lambda = \Lambda(\lambda, \phi)$ で置き換えた以下の式を考える。

$$U_a = \lambda \nabla_a \phi = \lambda C_a \quad \rightarrow \quad \check{U}_a = \Lambda(\lambda, \phi) C_a \quad (\text{A1})$$

$\check{L}_d \equiv \nabla_d \Lambda(\lambda, \phi)$ として、直交条件 (8b) : $C^b L_b = 0$ に対応するもの考えると、

$$C^d \check{L}_d = C^d \nabla_d \Lambda = C^d \left[(\partial_\lambda \Lambda) L_d + (\partial_\phi \Lambda) C_d \right] = 0 \quad (\text{A2})$$

となり、同様な条件を得る。混合テンソル $S^b_d = C^b L_d - L^b C_d$ に対しても同様に、 $\check{S}^b_d = C^b \check{L}_d - \check{L}^b C_d$ を考えて、その発散を計算すると

$$\nabla_b \check{S}^b_d = \nabla_b (C^b \check{L}_d - \check{L}^b C_d) = C^b \nabla_b \check{L}_d - (\nabla_b L^b) C_d - \check{L}^b \nabla_b C_d \quad (\text{A3})$$

となるが、 \check{L}_d は傾度ベクトルなので $\nabla_b \check{L}_d = \nabla_d \check{L}_b$ が成り立ち、これと (A2) を用いると (A3) は最終的に以下の様に変形される。

$$\nabla_b \check{S}^b_d = -(\nabla_b L^b) C_d - 2\check{L}^b \nabla_b C_d \quad (\text{A4})$$

また、

$$\begin{aligned} \nabla_b \check{L}^b &= \nabla_b \left\{ (\partial_\lambda \Lambda) L^b + (\partial_\phi \Lambda) C^b \right\} = (\partial_\lambda \Lambda) \nabla_b L^b + \left\{ (\partial_{\lambda\lambda}^2 \Lambda) L_b + (\partial_{\lambda\phi}^2 \Lambda) C_b \right\} L^b \\ &= + \left\{ (\partial_{\lambda\phi}^2 \Lambda) L_b + (\partial_{\phi\phi}^2 \Lambda) C_b \right\} C^b = (\partial_\lambda \Lambda) \nabla_b L^b + (\partial_{\lambda\lambda}^2 \Lambda) L_b L^b \end{aligned} \quad (\text{A5})$$

と変形され、最後の項に条件式 (9) を代入した (A5) を (A4) 代入して、

$$\nabla_b \check{S}_d^b = -(\partial_\lambda \Lambda) \{ (\nabla_b L^b) C_d + 2L^b \nabla_b C_d \} + (\partial_{\lambda\lambda}^2 \Lambda) C_d \quad (\text{A6})$$

得るが、(A6) において $\Lambda = \lambda$ とすると

$$\nabla_b S_d^b = -\{ (\nabla_b L^b) C_d + 2L^b \nabla_b C_d \} \quad (\text{A7})$$

となるが、これは式 (14b) より $\nabla_b S_d^b = -\{ (\nabla_b L^b) C_d + 2L^b \nabla_b C_d \} = a C_d$ なので、これを (A6) に代入し、

$$\nabla_b \check{S}_d^b = (a\kappa + \partial_\lambda \kappa) C_d \quad (\text{A8})$$

となる。ここに、 $\kappa \equiv \partial_\lambda \Lambda$ と定義した。従って、 a に対して $\partial_\lambda \ln \kappa = -a$ から Λ を決定すると、その様な Λ に対して、(A8) は

$$\nabla_b \check{S}_d^b = 0 \quad (\text{A9})$$

となる。

8. 参考文献

- [1] 今井 功：電磁気学を考える、1990 サイエンス社
- [2] Landau L.D. & E.M. Lifshitz: Fluid Mechanics 2nd Ed. 1987, Reed Educational and Professional Publishing Ltd.
- [3] Lamb, H.: Hydrodynamics 6th Ed. Cambridge Univ. Press.
- [4] Huggett, S. A. and Tod, K. P.: An introduction to Twistor theory 2nd Ed. 1994, Cambridge Univ. Press.
- [5] H. Sakuma and I. Ojima: On the synchronized electromagnetic and gravitational radiation, preprint.
- [6] 高崎金久：ツイスターの世界、2005 共立出版
- [7] Tonomura, A., Osakabe, N., Matsuda, T., Kawasaki, T., Endo, J., Yano, S. and Yamada, H.: Evidence for Aharonov-Bohm effect with magnetic field completely shielded from electron wave. Phys. Rev. Lett., 56, 792-795, 1986.

Non-Markovian evolution of quantum systems

Dariusz Chruściński^{*)} and Andrzej Kossakowski^{**)}

*Institute of Physics, Nicolaus Copernicus University
Grudziądzka 5/7, 87-100 Toruń, Poland*

We analyze a local approach to the non-Markovian evolution of open quantum systems. It turns out that any dynamical map representing evolution of such a system may be described either by non-local master equation with memory kernel or equivalently by equation which is local in time. Local approach turns out to be more suitable for practical problems which is illustrated by several examples.

§1. Introduction

The dynamics of open quantum systems attracts nowadays increasing attention.^{1)–3)} It is relevant not only for the better understanding of quantum theory but it is fundamental in various modern applications of quantum mechanics. Since the system-environment interaction causes dissipation, decay and decoherence it is clear that dynamic of open systems is fundamental in modern quantum technologies, such as quantum communication, cryptography and computation.⁴⁾

The usual approach to the dynamics of an open quantum system consists in applying the Markovian approximation, that leads to the following local master equation for the dynamical map $\Lambda(t)$

$$\frac{d}{dt}\Lambda(t) = \mathcal{L}\Lambda(t), \quad \Lambda(0) = \mathbb{1}, \quad (1.1)$$

where \mathcal{L} denotes the time-independent generator of the dynamical semigroup possessing the following well known representation^{3),5),6)}

$$\mathcal{L}\rho = -i[H, \rho] + \sum_{\alpha} \left(V_{\alpha}\rho V_{\alpha}^{\dagger} - \frac{1}{2}\{V_{\alpha}^{\dagger}V_{\alpha}, \rho\} \right). \quad (1.2)$$

The above structure of \mathcal{L} guaranties that dynamical map $\Lambda(t)$ is completely positive and trace preserving for all $t \geq 0$. It is clear that $\Lambda(t)$ generates the following evolution of the density matrix $\rho(t) = \Lambda(t)\rho_0$. The dynamical semigroup is characterized by the following property

$$\Lambda(t+s) = \Lambda(t)\Lambda(s), \quad (1.3)$$

for all $t, s \geq 0$. In general the external conditions which influence the dynamics of an open system may vary in time. The natural generalization of the Markovian master equation (1.1) involves time-dependent generator $\mathcal{L}(t)$ which has exactly the same representation as in (1.2) with time-dependent Hamiltonian $H(t)$ and time-

^{*)} E-mail: darch@fizyka.umk.pl

^{**)} E-mail: kossak@fizyka.umk.pl

dependent Lindblad operators $V_\alpha(t)$. Therefore one gets the following master equation for the dynamical map $\Lambda(t, t_0)$

$$\frac{d}{dt}\Lambda(t, t_0) = \mathcal{L}(t) \Lambda(t, t_0) , \quad \Lambda(t_0, t_0) = \mathbb{1} , \quad (1.4)$$

which leads to the following solution

$$\Lambda(t, t_0) = \text{T exp} \left(\int_{t_0}^t \mathcal{L}(\tau) d\tau \right) , \quad (1.5)$$

where T stands for the chronological operator. Clearly, $\Lambda(t, t_0)$ no longer depends upon ' $t - t_0$ ' but it still satisfies inhomogeneous composition law

$$\Lambda(t, s) \cdot \Lambda(s, t_0) = \Lambda(t, t_0) , \quad (1.6)$$

for $t \geq s \geq t_0$. We stress that (1.4) although time-dependent is perfectly Markovian. Note, however, that the solution (1.5) has only a formal meaning since the evaluation of T-product is in general not feasible.

In this paper we analyze a non-Markovian generalization of Markovian semigroup $\Lambda(t)$. Let us recall that $\Lambda(t)$ is divisible if

$$\Lambda(t) = V(t, s)\Lambda(s) , \quad (1.7)$$

where $V(t, s)$ is completely positive for all $t \geq s$. Throughout this paper we use the following definition of Markovianity

Definition 1 *A dynamical map $\Lambda(t)$ is Markovian if and only if it is divisible.*

Note that Markovian semigroup is trivially divisible with $V(t, s) = \Lambda(t - s)$. Our goal is to characterize legitimate dynamical maps which are not divisible (hence non-Markovian). The problem of non-Markovian evolution was recently analyzed by many authors.⁷⁾⁻²²⁾

The paper is organized as follows: in the next section we compare a general approach to non-Markovian evolution based on the non-local memory kernel and local in time generator and in section 3 we present simple examples to illustrate a general approach. Section 4 is devoted to general analysis of the non-Markovian dynamics of a single qubit. Final remarks are collected in section 5.

§2. Memory kernel vs. local generator

Consider now time-homogeneous situation $\Lambda(t, t_0) = \Lambda(t - t_0)$ and let us fix the initial time $t_0 = 0$. The standard approach to the dynamics of open system uses the Nakajima-Zwanzig projection operator technique²³⁾ (see also Refs.^{1),3)}) which shows that under fairly general conditions, the master equation for the dynamical map $\Lambda(t)$ defined via

$$\Lambda(t)\rho = \text{Tr}_E [e^{-itH_{SE}} \rho \otimes \omega e^{itH_{SE}}] \quad (2.1)$$

where H_{SE} denotes the system-environment Hamiltonian and ω is a fixed state of the environment, takes the form of the following non-local equation

$$\frac{d}{dt} \Lambda(t) = \int_0^t \mathcal{K}(t - u)\Lambda(u) du , \quad \Lambda(0) = \mathbb{1} , \quad (2.2)$$

in which quantum memory effects are taken into account through the introduction of the memory kernel $\mathcal{K}(t)$: this simply means that the rate of change of $\Lambda(t)$ at time t depends on its history (starting at $t = 0$). Unfortunately, we do not know condition for the memory kernel $\mathcal{K}(t)$ which guarantee that the corresponding dynamical map $\Lambda(t, t_0)$ is a legitimate quantum evolution, i.e. it is completely positive and trace preserving. Therefore, instead of non-local approach we propose to analyze much simpler approach which is based on the local in time Master Equation (this approach is usually called time-convolutionless (TCL)^{1,24}). Note, that each solution $\Lambda(t)$ of (2.2) satisfies the following local in time equation²²

$$\frac{d}{dt}\Lambda(t) = \mathcal{L}(t)\Lambda(t) , \quad \Lambda(0) = \mathbb{1} , \quad (2.3)$$

where the time-dependent generator $\mathcal{L}(t)$ is defined by the following logarithmic derivative of the dynamical map

$$\mathcal{L}(t) := \dot{\Lambda}(t) \cdot \Lambda^{-1}(t) . \quad (2.4)$$

We stress that equations (1.4) and (2.3) sharing the same structure have completely different interpretation: the former corresponds to the inhomogeneous case and describes 2-parameter family of maps $\Lambda(t, t_0)$ with arbitrary initial time t_0 , whereas the latter corresponds to the homogeneous case and describes 1-parameter family $\Lambda(t)$, where (using homogeneity) we fixed the initial time $t_0 = 0$. Note, that performing the Laplace transform of (2.2) one obtains

$$\tilde{\Lambda}(s) = \frac{1}{s - \tilde{\mathcal{K}}(s)} , \quad (2.5)$$

and hence one finds the non-local analog of (2.4)

$$\tilde{\mathcal{K}}(s) = [s\tilde{\Lambda}(s) - \mathbb{1}] \tilde{\Lambda}(s)^{-1} . \quad (2.6)$$

Hence, knowing $\Lambda(t)$ one may define a local generator via (2.4) and non-local one via (2.6). Note, that both formulae (2.4) and (2.6) are rather formal, since in general it would be difficult to compute the inverse of $\Lambda(t)$ or the inverse laplace transform of $\tilde{\mathcal{K}}(s)$. We stress, however, that we do not know how to construct $\mathcal{L}(t)$ and $\mathcal{K}(t)$ from ‘first principle’ such that the resulting evolution $\Lambda(t)$ defines a legitimate dynamical map.

The situation simplifies in the case of commutative dynamics, that is, $\Lambda(t)$ satisfying $\Lambda(t)\Lambda(u) = \Lambda(u)\Lambda(t)$ for all $t, u \geq 0$. In this case one has the following solution to the local equation

$$\Lambda(t) = \exp\left(\int_0^t \mathcal{L}(\tau)d\tau\right) , \quad (2.7)$$

that is, one may avoid to use chronological T-product. Now comes the following

Theorem 1 $\Lambda(t)$ defines a legitimate quantum dynamics if and only if $\int_0^t \mathcal{L}(\tau)d\tau$ defines a legitimate Lindblad generator for all $t \geq 0$.

In particular if the local generator has the following form

$$\mathcal{L}(t) = a_1(t)\mathcal{L}_1 + \dots + a_n(t)\mathcal{L}_n , \quad (2.8)$$

where $\mathcal{L}_1, \dots, \mathcal{L}_n$ are Lindblad generators, then $\Lambda(t)$ defines a legitimate quantum dynamics if and only if

$$\int_0^t a_k(\tau) d\tau \geq 0 , \quad k = 1, \dots, n . \quad (2.9)$$

Recall, that $\Lambda(t)$ is Markovian if and only if $a_k(t) \geq 0$ for all $t \geq 0$ and all $k = 1, \dots, n$. Hence condition (2.9) is a natural generalization of Markovianity. Hence, if (2.9) is satisfied but at least one $a_k(t)$ takes strictly negative value at some moment of time, then $\Lambda(t)$ necessarily describes non-Markovian quantum dynamics.

§3. Non-Markovian evolution: examples

To illustrate general approach let us consider few simple examples.

Example 1. Consider the dynamical map for a qudit (d -level quantum system) given by

$$\Lambda(t) = \left(1 - \int_0^t f(u) du \right) \mathbb{1} + \int_0^t f(u) du \mathcal{P} , \quad (3.1)$$

where $\mathcal{P} : \mathcal{B}(\mathbb{C}^d) \rightarrow \mathcal{B}(\mathbb{C}^d)$ denotes completely positive trace preserving projection. For example take a fixed qudit state ω and define \mathcal{P} by the following formula $\mathcal{P}\rho = \omega \text{Tr}\rho$. The real function ‘ f ’ satisfies:

$$0 \leq \int_0^t f(u) du \leq 1 ,$$

for any $t > 0$. Note that $f(u)$ needs not be positive. If $f(u) \geq 0$ ($u \geq 0$), then $\Lambda(t)$ defines quantum semi-Markov process and the function $f(u)$ may be interpreted as a waiting time distribution for this process.^{8),18)} Clearly, $\Lambda(t)$ being a convex combination of $\mathbb{1}$ and \mathcal{P} is completely positive trace preserving map and hence it defines legal quantum dynamics of a qudit. The corresponding memory kernel is well known^{8),18)} and it is given

$$\mathcal{K}(t) = \kappa(t)\mathcal{L}_0 , \quad (3.2)$$

where the function $\kappa(t)$ is defined in terms of its Laplace transform as follows

$$\tilde{\kappa}(s) = \frac{s\tilde{f}(s)}{1 - \tilde{f}(s)} , \quad (3.3)$$

and \mathcal{L}_0 is defined by $\mathcal{L}_0 = \mathcal{P} - \mathbb{1}$. Note, that \mathcal{L}_0 has exactly the structure of the Markovian generator (1.2) with $H = 0$, and the Lindblad operators V_α define Kraus representation of \mathcal{P} , that is $\mathcal{P}\rho = \sum_\alpha V_\alpha \rho V_\alpha^\dagger$. One easily finds for the corresponding generator

$$\mathcal{L}(t) = \alpha(t)\mathcal{L}_0 , \quad (3.4)$$

where

$$\alpha(t) = \frac{f(t)}{1 - \int_0^t f(u)du} . \quad (3.5)$$

Let us observe that

$$\int_0^t \alpha(u)du = -\ln \left(1 - \int_0^t f(u)du \right) \geq 0 ,$$

and hence this example gives rise to $\mathcal{L}(t) = \alpha(t)\mathcal{L}_0$, with Markovian \mathcal{L}_0 and $\alpha(t)$ satisfying $\int_0^t \alpha(u)du \geq 0$. We stress that $\alpha(t)$ needs not be positive. It is positive if and only if $f(t)$ corresponds to the waiting time distribution.^{8),18)} Note the striking similarity between formulae (3.3) and (3.5). It should be stressed that in this case one knows an explicit formula for time-local generator $\mathcal{L}(t)$. Note, however, that in general one is not able to invert the Laplace transform of $\tilde{\kappa}(s)$ from the formula (3.3) and hence the explicit formula for the memory kernel $\mathcal{K}(t)$ is not known.

Consider for example $f(t) = \varepsilon\gamma e^{-\gamma t}$, with $\gamma > 0$ and $\varepsilon \in (0, 1]$. One finds from (3.5) the following formula $\alpha(t) = \varepsilon\gamma[(1 - \varepsilon)e^{\gamma t} + \varepsilon]^{-1}$. Note, that for $\varepsilon = 1$ it reduces to $\alpha(t) = \gamma$, that is, it corresponds to the purely Markovian case. Hence, the parameter ‘ $1 - \varepsilon$ ’ measures the non-Markovianity of the dynamics.

Example 2. Suppose now that $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ and consider the dynamics of two qubits governed by (3.1) with $f(t) = \varepsilon\gamma e^{-\gamma t}$. The entanglement of the asymptotic state is governed by

$$C(\infty) = 2 \max\{c_1(\infty), c_2(\infty), 0\} ,$$

with

$$c_1(\infty) = (1 - \varepsilon)|\rho_{23}| - \sqrt{\rho_{11}\rho_{44}} ,$$

and

$$c_2(\infty) = (1 - \varepsilon)|\rho_{14}| - \sqrt{\rho_{22}\rho_{33}} .$$

It is clear that in the Markovian case ($\varepsilon = 1$) the asymptotic state is always separable ($C(\infty) = 0$). However, for sufficiently small ‘ ε ’ (i.e. sufficiently big non-Markovianity parameter ‘ $1 - \varepsilon$ ’) one may have $c_1(\infty) > 0$ or $c_2(\infty) > 0$, that is, the asymptotic state might be entangled. This example proves the crucial difference between Markovian and non-Markovian dynamics of composed systems. In particular controlling ‘ ε ’ we may avoid sudden death of entanglement.²⁵⁾

Example 3. Consider the pure decoherence model defined by the following Hamiltonian $H = H_R + H_S + H_{SR}$, where H_R is the reservoir Hamiltonian, $H_S = \sum_n \varepsilon_n P_n$ ($P_n = |n\rangle\langle n|$) the system Hamiltonian and

$$H_{SR} = \sum_n P_n \otimes B_n \quad (3.6)$$

the interaction part, $B_n = B_n^\dagger$ being reservoirs operators. The initial product state $\rho \otimes \omega_R$ evolves according to the unitary evolution $e^{-iHt}(\rho \otimes \omega_R)e^{iHt}$ and by partial

tracing with respect to the reservoir degrees of freedom one finds for the evolved system density matrix

$$\rho(t) = \text{Tr}_R[e^{-iHt}(\rho \otimes \omega_R)e^{iHt}] = \sum_{n,m} c_{mn}(t)P_m\rho P_n ,$$

where $c_{mn}(t) = \text{Tr}(e^{-iZ_m t}\omega_R e^{iZ_n t})$, $Z_n = \epsilon_n \mathbb{I}_R + H_R + B_n$ being reservoir operators. Note that the matrix $c_{mn}(t)$ is semi-positive definite and hence

$$\Lambda(t)\rho = \sum_{n,m} c_{mn}(t)P_m\rho P_n . \quad (3.7)$$

defines the Kraus representation of the completely positive map $\Lambda(t)$. The solution of the pure decoherence model can therefore be found without explicitly writing down the underlying master equation. Our method, however, enables one to find the corresponding generator $\mathcal{L}(t)$. It is given by the following formula

$$\mathcal{L}(t)\rho = \sum_{n,m} \alpha_{mn}(t)P_m\rho P_n , \quad (3.8)$$

where the functions $\alpha_{mn}(t)$ are defined by $\alpha_{mn} = \dot{c}_{mn}/c_{mn}$. It shows that the pure decoherence model may be defined by local in time master equation (2.3) with the non-Markovian generator (3.8). Note, that for $d = 2$ one finds

$$\mathcal{L}(t)\rho = \alpha_{01}(t)P_0\rho P_1 + \bar{\alpha}_{01}(t)P_1\rho P_0 . \quad (3.9)$$

Using

$$P_0 = \frac{1}{2}(\mathbb{I}_2 + \sigma_z) , \quad P_1 = \frac{1}{2}(\mathbb{I}_2 - \sigma_z) ,$$

one obtains the following formula for the local generator

$$\mathcal{L}(t)\rho = i\frac{\Omega(t)}{2}[\sigma_z, \rho] - \frac{\Gamma(t)}{2}(\sigma_z\rho\sigma_z - \rho) , \quad (3.10)$$

where

$$\Omega = \text{Im } c_{01} , \quad \Gamma = \text{Re } c_{01} .$$

which has the standard Lindblad form.

§4. Non-Markovian dynamics of a qubit

In this section we analyze a general dynamical map for a qubit

$$\Lambda(t)\rho = \sum_{m,n=0}^1 c_{mn}(t)P_m\rho P_n + d_1(t)\sigma_+\rho\sigma_- + d_0(t)\sigma_-\rho\sigma_+ , \quad (4.1)$$

where the positive matrix c_{mn} reads as follows

$$c_{mn}(t) = \begin{pmatrix} x_0(t) & \gamma(t) \\ \gamma^*(t) & x_1(t) \end{pmatrix} , \quad (4.2)$$

and

$$d_0(t) = 1 - x_1(t) , \quad (4.3)$$

$$d_1(t) = 1 - x_2(t) . \quad (4.4)$$

Note, that $\Lambda(t)$ is completely positive if and only if

$$x_0(t), x_1(t) \in [0, 1] . \quad (4.5)$$

and

$$x_0(t)x_1(t) \geq |\gamma(t)|^2 . \quad (4.6)$$

Finally, $\Lambda(0) = \text{id}$ iff $x_0(0) = x_1(0) = \gamma(0) = 1$. We use standard notation: if $|0\rangle, |1\rangle$ denotes an orthonormal basis in \mathbb{C}^2 , then $P_k = |k\rangle\langle k|$ and σ_{\pm} are raising and lowering operators in the qubit space, that is, $\sigma_+ = |1\rangle\langle 0|$ and $\sigma_- = |0\rangle\langle 1|$. The corresponding dynamics of the density operator $\rho(t)$ reads as follows

$$\begin{aligned} \rho_{00}(t) &= \rho_{00} x_0(t) + \rho_{11} [1 - x_1(t)] , \\ \rho_{11}(t) &= \rho_{00} [1 - x_0(t)] + \rho_{11} x_1(t) , \\ \rho_{01}(t) &= \rho_{01} \gamma(t) . \end{aligned} \quad (4.7)$$

One easily finds for the local generator

$$\mathcal{L}(t)\rho = -i\frac{\Omega(t)}{2} [\sigma_3, \rho] + \sum_{k=0}^1 y_k(t)\mathcal{L}_k\rho + \frac{\Gamma(t)}{2}\mathcal{L}_z\rho . \quad (4.8)$$

where the generators $\{\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_z\}$ read as follows

$$\begin{aligned} \mathcal{L}_0\rho &= \sigma_+\rho\sigma_- - \frac{1}{2}\{\sigma_-\sigma_+, \rho\} , \\ \mathcal{L}_1\rho &= \sigma_-\rho\sigma_+ - \frac{1}{2}\{\sigma_+\sigma_-, \rho\} , \\ \mathcal{L}_z\rho &= \sigma_z\rho\sigma_z - \rho , \end{aligned} \quad (4.9)$$

and

$$\begin{aligned} y_0 &= \frac{\dot{x}_0 x_1 + \dot{x}_1 (1 - x_0)}{1 - x_0 - x_1} , \\ y_1 &= \frac{\dot{x}_0 (1 - x_1) + \dot{x}_1 x_0}{1 - x_0 - x_1} , \\ \Gamma &= -\frac{y_0 + y_1}{2} - \text{Re} \frac{\dot{\gamma}}{\gamma} , \\ \Omega &= \text{Im} \frac{\dot{\gamma}}{\gamma} , \end{aligned} \quad (4.10)$$

where for simplicity we skipped time dependence.

Corollary 1 *Qubit dynamics (4.1) is Markovian if and only if $y_0(t), y_1(t), \Gamma(t) \geq 0$ for all $t \geq 0$.*

In particular taking $x_0 = x_1 = 1$ one recovers pure decoherence model with

$$\mathcal{L}(t)\rho = -i\frac{\Omega(t)}{2}[\sigma_z, \rho] + \frac{\Gamma(t)}{2}\mathcal{L}_z\rho, \quad (4.11)$$

where

$$\Gamma = -\operatorname{Re}\frac{\dot{\gamma}}{\gamma}, \quad \Omega = \operatorname{Im}\frac{\dot{\gamma}}{\gamma}.$$

§5. Conclusions

In conclusion, any non-Markovian quantum evolution may be described either by the non-local equation (2.2) or by a time-local equation (2.3). Local approach is more simple and well suited for practical purposes. The Markovian dynamics corresponds to divisible dynamics maps. Note, that divisibility may be checked on the level of the local approach – it corresponds to the time dependent Lindblad generator. We stress that the problem of necessary and sufficient condition for the local generator which do guarantee that the corresponding dynamical map gives rise to the legitimate quantum evolution is still open and it deserves further studies.

Acknowledgments

One of the authors (DC) thanks prof. Ojima for his kind hospitality during the conference in Kyoto and for interesting discussions about the history of Japan.

References

- 1) H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford Univ. Press, Oxford, 2007).
- 2) U. Weiss, *Quantum Dissipative Systems*, (World Scientific, Singapore, 2000).
- 3) R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications* (Springer, Berlin, 1987).
- 4) M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
- 5) G. Lindblad, *Comm. Math. Phys.* **48**, 119 (1976).
- 6) V. Gorini, A. Kossakowski, and E.C.G. Sudarshan, *J. Math. Phys.* **17**, 821 (1976).
- 7) J. Wilkie, *Phys. Rev. E* **62**, 8808 (2000); J. Wilkie and Yin Mei Wong, *J. Phys. A: Math. Theor.* **42**, 015006 (2009).
- 8) A. A. Budini, *Phys. Rev. A* **69**, 042107 (2004); *ibid.* **74**, 053815 (2006).
- 9) H.-P. Breuer, *Phys. Rev. A* **69**, 022115 (2004); *ibid.* **70**, 012106 (2004).
- 10) S. Daffer, K. Wódkiewicz, J.D. Cresser, and J.K. McIver, *Phys. Rev. A* **70**, 010304 (2004).
- 11) A. Shabani and D.A. Lidar, *Phys. Rev. A* **71**, 020101(R) (2005).
- 12) S. Maniscalco, *Phys. Rev. A* **72**, 024103 (2005).
- 13) S. Maniscalco and F. Petruccione, *Phys. Rev. A* **73**, 012111 (2006).
- 14) J. Piilo, K. Härkönen, S. Maniscalco, K.-A. Suominen, *Phys. Rev. Lett.* **100**, 180402 (2008); *Phys. Rev. A* **79**, 062112 (2009).
- 15) E. Andersson, J. D. Cresser and M. J. W. Hall, *J. Mod. Opt.* **54**, 1695 (2007).
- 16) A. Kossakowski and R. Rebolledo, *Open Syst. Inf. Dyn.* **14**, 265 (2007); *ibid.* **15**, 135 (2008).
- 17) A. Kossakowski and R. Rebolledo, *Open Syst. Inf. Dyn.* **16**, 259 (2009).
- 18) H.-P. Breuer and B. Vacchini, *Phys. Rev. Lett.* **101** (2008) 140402; *Phys. Rev. E* **79**, 041147 (2009).
- 19) M. Moodley and F. Petruccione, *Phys. Rev. A* **79**, 042103 (2009).
- 20) B. Vacchini and H.-P. Breuer, *Phys. Rev. A* **81**, 042103 (2010).

- 21) D. Chruściński, A. Kossakowski, and S. Pascazio, Phys. Rev. A **81**, 032101 (2010)
- 22) D. Chruściński and A. Kossakowski, Phys. Rev. Lett. **104**, 070406 (2010).
- 23) S. Nakajima, Prog. Theor. Phys. **20**, 948 (1958); R. Zwanzig, J. Chem. Phys. **33**, 1338 (1960).
- 24) H.-P. Breuer, B. Kappler and F. Petruccione, Phys. Rev. A **59**, 1633 (1999).
- 25) T. Yu and J. H. Eberly, Opt. Comm. **264**, 393 (2006); Q. Inf. Comp. **7**, 459 (2007); Phys. Rev. Lett. **97**, 140403 (2006); *ibid.* **93**, 140404 (2004).

Who has seen a free photon?

Izumi Ojima ^{*}
 Research Institute for Mathematical Sciences,
 Kyoto University
 and
 Hayato Saigo [†]
 Nagahama Institute of the Bio-Science and Technology

Abstract

While the notion of the position of photons is indispensable in the quantum optical situations, it has been known in mathematical physics that any position operator cannot be defined for a massless free particle with a non-zero finite spin. This dilemma is resolved by introducing the “effective mass” of a photon due to the interaction with matter. The validity of this interpretation is confirmed in reference to the picture of “polariton”, a basic notion in optical and solid physics.

1 A dilemma

Who has seen the wind?

Neither I nor you.

But when the leaves hang trembling

The wind is passing through.

This famous verse [15] reminds the authors of a fundamental dilemma about “localization of a photon”.

In the recent advanced quantum-optical technology, the notion of positions of photons has played indispensable roles in the experimental situations, as is exemplified by the most familiar Mach-Zender interferometry consisting of the beam splitters and of the photon detector to specify the position of detection of photons. Theoretically speaking, this is the contexts where the localization of a photon is to be described by means of some *position observables* associated with the quantum electromagnetic field. Here we encounter the following serious difficulty: since the paper of Newton and

^{*}E-mail: ojima@kurims.kyoto-u.ac.jp

[†]E-mail: h_saigoh@nagahama-i-bio.ac.jp

Wigner [11], it has been known in mathematical physics that any position operator cannot be defined for a massless free particle with a non-zero finite spin, in sharp contrast to the cases of massive particles which can be localized. This statement is clearly in contradiction to the above familiar situations where almost all physicists have used the notion of “position of a photon” as one of the basic ingredients of theory and application of quantum mechanics.

Then, who has seen a free photon?

2 Strategy

We propose the following resolution: **while a massless free photon is not localizable, a “real” photon can be made localizable by its dynamical interaction with matter such as media or devices (for its detection).** To be precise, this simple picture can be shown to have a sound basis as follows. First we see in Section 3 that **a photon can be localized only if photon-matter interaction provides a positive effective mass**, by our reinterpretation of Wightman’s theorem [16] following from the arguments by Newton and Wigner [11]. Then we show in Section 4 how the photon-matter interaction can be reduced in this context into the picture of “a free particle with positive effective mass”. **This interpretation becomes clearer if we refer to the picture of “polariton”** [6, 5], a basic notion in optical and solid-state physics.

3 Formulation

3.1 Newton-Wigner-Wightman analysis

In 1949, Newton and Wigner [11] raised the question of localizability of single free particles. They attempted to formulate the properties of the localized states on the basis of natural requirements of relativistic covariance.

Physical quantities available in this formulation admitting direct physical meaning are restricted inevitably to the generators of Poincaré group $\mathcal{P}_+^\uparrow = \mathbb{R}^4 \rtimes L_+^\uparrow$ (with L_+^\uparrow the orthochronous proper Lorentz group) which is locally isomorphic to the semi-direct product $\mathcal{H}_2(\mathbb{C}) \rtimes SL(2, \mathbb{C})$ of the hermitian (2×2) -matrices and of $SL(2, \mathbb{C})$, consisting of the energy-momentum vector P_μ and of the Lorentz generators $M_{\mu\nu}$ (composed of angular momenta M_{ij} and of Lorentz boosts M_{0i}). The problem is then to find conditions under which “position operators” can naturally be constructed from the Poincaré generators $(P_\mu, M_{\mu\nu})$. In [11], position operators have been shown to exist in massive cases in an essentially unique way for “elementary” systems in the sense of the irreducibility of the corresponding representations of \mathcal{P}_+^\uparrow so that

localizability of a state can be defined in terms of such position operators. In massless cases, however, no localized states are found to exist in the above sense. That was the beginning of the story.

Wightman [16] clarified the situation by recapturing the notion of “localization” in a general form as follows. First he has reformulated the usual approaches in terms of unbounded position operators into the form of general axioms (i)-(v) involving projection operators,

- (i) To each Borel subset Δ of \mathbb{R}^3 , there corresponds a projection operator $E(\Delta)$ in a Hilbert space \mathfrak{H} , whose expectation value gives the probability of finding the system in Δ ;
- (ii) $E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2)$;
- (iii) $E(\Delta_1 \cup \Delta_2) = E(\Delta_1) + E(\Delta_2)$, if $\Delta_1 \cap \Delta_2 = \phi$;
- (iv) $E(\mathbb{R}^3) = 1$;
- (v) $E(\mathcal{R}\Delta + \mathbf{a}) = U(\mathbf{a}, \mathcal{R})E(\Delta)U(\mathbf{a}, \mathcal{R})^{-1}$, where $\mathcal{R}\Delta + \mathbf{a}$ is the set obtained from Δ by applying a translation \mathbf{a} after a rotation \mathcal{R} , and $U(\mathbf{a}, \mathcal{R})$ is the corresponding unitary operator in \mathfrak{H} .

Note that the notion of localizability discussed above is concerned with *localization of states in space at a given time*. If we consider the axioms like (i)-(v) on the whole space-time, it would imply the validity of the CCR relations between 4-momenta p_μ and space-time coordinates x^ν , which would imply the Lebesgue spectrum covering the whole \mathbb{R}^4 for both observables \hat{p}_μ and \hat{x}^ν . Then, any such physical requirements as the spectrum condition cannot be imposed on the energy-momentum spectrum \hat{p}_μ , and hence, the notion of localizability in space-time does not make sense.

According to Mackey’s theory of induced representations, Wightman’s formulation can easily be seen as the condition for the set of operators $\{E(\Delta)\}$ to constitute a *system of imprimitivity* [8] under the action of the unitary representation $U(a, \mathcal{R})$ in \mathfrak{H} of the three-dimensional Euclidean group $SE(3) := \mathbb{R}^3 \rtimes SO(3)$. In a more algebraic form, the pair (E, U) can also be viewed as a *covariant representation*

$$E(\tau_{(\mathbf{a}, \mathcal{R})}(f)) = U(a, \mathcal{R})E(f)U(a, \mathcal{R})^{-1} \quad \text{for } f \in L^\infty(\mathbb{R}^3), (\mathbf{a}, \mathcal{R}) \in SE(3), \quad (1)$$

of an action $SE(3) \curvearrowright_\tau L^\infty(\mathbb{R}^3)$, $[\tau_{(\mathbf{a}, \mathcal{R})}(f)](\mathbf{x}) := f(\mathcal{R}^{-1}(\mathbf{x} - \mathbf{a}))$ on the algebra $L^\infty(\mathbb{R}^3)$ generated by the position operators in the representation $E : L^\infty(\mathbb{R}^3) \ni f \mapsto E(f) = \int f(\mathbf{x})dE(\mathbf{x}) \in B(\mathfrak{H})$, s.t. $E(\chi_\Delta) = E(\Delta)$.

Thus Wightman’s formulation of the Newton-Wigner localizability problem is just to examine whether the Hilbert space \mathfrak{H} of the representation (U, \mathfrak{H}) of $SE(3)$ can accommodate a representation E of the algebra $L^\infty(\mathbb{R}^3)$

consisting of position operators, covariant under the action of $SE(3)$ in the sense of (1) .

Applying the general theory of Mackey to the case of three-dimensional Euclidean group $SE(3)$, Wightman proved the fundamental result below as a purely kinematical consequence.

Theorem 1 ([16], **excerpt from theorem 6 and 7**) *A Lorentz or Galilei covariant massive system is always localizable. For the Lorentz case, the only localizable massless elementary system (i.e. irreducible representation) has spin zero. For the Galilei case, no massless elementary system is localizable.*

Corollary 2 *A free photon is not localizable.*

The essential mechanism of (non-)localizability in the sense of Newton-Wigner-Wightman is determined by the structure of little groups defined by Wigner as the stabilizer groups of standard four-momenta on each type of \mathcal{P}_+^\uparrow -orbits in p -space.

On an orbit $p^2 = m^2 > 0$ under \mathcal{P}_+^\uparrow , we can choose a standard momentum $p_0 := (m, \mathbf{0})$ which specifies a rest frame of a particle with mass $m \neq 0$. Then, the little group at p_0 is the group $SO(3)$ of spatial rotations, corresponding to the degrees of freedom remaining in the rest frame. As a consequence, “the space of all rest frames” along the orbit becomes $SO(1,3)/SO(3) \cong \mathbb{R}^3$. Here the coordinates of “rest frames” just play the role of the order parameters (or, “sector parameters”) on each \mathcal{P}_+^\uparrow -orbit as the space of “degenerate vacua” associated with a symmetry breaking of boost invariance, which should play the roles of position operators appearing in the imprimitivity system as follows. Denoting by Λ_p such a Lorentz boost as transforming p_0 into $p = \Lambda_p p_0 = (\sqrt{m^2 c^2 + \mathbf{p}^2}, \mathbf{p})$, we have for each boost $\Lambda \in SO(1,3)$ such a relation as $\Lambda = \Lambda_{\Lambda p} R(\Lambda) \Lambda_p^{-1}$ with the so-called Wigner rotation $R(\Lambda) \in SO(3)$ stabilizing p_0 . On the other hand, we have for $p_\Lambda = \Lambda p_0$ a relative velocity $\mathbf{u}_\Lambda \in \mathbb{R}^3 \cong SO(1,3)/SO(3)$ between the Lorentz frames $(1, \mathbf{0})$ and $\frac{p_\Lambda}{m} =: u_\Lambda = \left(\frac{1}{\sqrt{1 - \mathbf{u}_\Lambda^2/c^2}}, \frac{\mathbf{u}_\Lambda}{\sqrt{1 - \mathbf{u}_\Lambda^2/c^2}} \right)$, $u_\Lambda^2 = 1$.

Thus, the homeomorphism $\mathbb{R}^3 \cong SO(1,3)/SO(3)$ physically means the relation between a Lorentz boost $\Lambda = \Lambda(\mathbf{u})$ transforming p_0 into $p_\Lambda = \Lambda p_0$ and a relative velocity $\mathbf{u} = \mathbf{u}(\Lambda) \in \mathbb{R}^3 \cong SO(1,3)/SO(3)$ satisfying the relation $\frac{p_\Lambda}{m} = \left(\frac{1}{\sqrt{1 - \mathbf{u}^2/c^2}}, \frac{\mathbf{u}}{\sqrt{1 - \mathbf{u}^2/c^2}} \right)$, which describes a non-trivial action of $\mathbf{u} \in \mathbb{R}^3$ on p belonging to the \mathcal{P}_+^\uparrow -orbit $p^2 = m^2 > 0$ via the Wigner rotation $R(\Lambda) = \Lambda_{\Lambda p}^{-1} \Lambda \Lambda_p$. In the limit $c \rightarrow \infty$, this action yields the imprimitivity system used by Wightman.

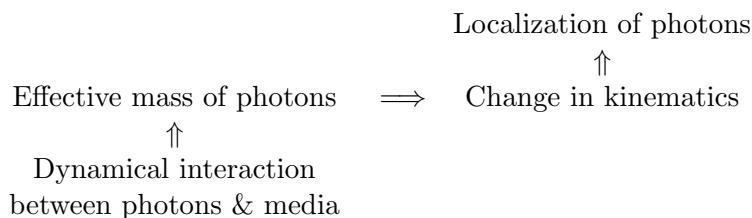
In sharp contrast, there is no rest frames for a massless particle and the little group becomes isomorphic to two-dimensional Euclidean group

$SE(2)$, whose rotational generator corresponds to the helicity. Since the other two translation generators corresponding to gauge transformations span *non-compact* directions in distinction from the massive cases with compact $SO(3)$, the allowed representation is only the trivial one which leaves the transverse modes invariant, and hence, the little group cannot provide position operators in the massless case.

Since Newton-Wigner-Wightman, many discussions around the photon localization problem have been developed. So far as we know, the arguments seem to be divided into two opposite viewpoints, one relying on purely dynamical bases [4] and another on pure kinematics [2], where it is almost impossible to find any meaningful agreements. Below we propose an alternative strategy based on the notion of “effective mass”, which can provide a reasonable reconciliation between these conflicting ideas because of its “kinematical” nature arising from some dynamical origin.

3.2 Wightman’s theorem as the “basis” for localization

Our scheme of the localization for photons can be summarized as follows, which is essentially in accordance with the basic formulation of “quadrality scheme” [13] underlying the Micro-Macro duality [12]:



Once a positive effective mass appears, Wightman’s theorem itself provides the “kinematical basis” for the localization of a photon. From our point of view, therefore, this theorem so far regarded as a no-go theorem against the localizability becomes actually an affirmative support for it, conveying such a strongly selective meaning that, whenever a photon is localized, it should carry a non-zero effective mass.

In the next section, we explain the meaning of our scheme from a physical point of view.

4 Resolution

4.1 How to define effective mass of a photon

Now we focus on a photon interacting with homogeneous medium, in the case of the monochromatic light with angular frequency ω as a classical light wave. For simplicity, we neglect here the effect of absorption, that is, the

imaginary part of refractive index. When a photon interacting with matter can be treated as a single particle, it is natural to identify its velocity \mathbf{v} with the “signal velocity” of light in medium. The relativistic total energy of the particle E should be related to $v := \sqrt{\mathbf{v} \cdot \mathbf{v}}$ by its mass m_{eff} :

$$E = \frac{m_{\text{eff}}c^2}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (2)$$

Since v is well known to be smaller than the light velocity c (theoretically or experimentally), m_{eff} is positive (when the particle picture above is valid). Then we may consider m_{eff} as the relativistic “effective (rest) mass of a photon”, and identify its momentum \mathbf{p} with

$$\mathbf{p} = \frac{m_{\text{eff}}\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (3)$$

Hence, as long as “an interacting photon” can be approximately treated as a single particle, it should be massive, according to which its “localization problem” is resolved. The validity of this picture will be confirmed in the next subsection.

While the concrete forms of energy/momentum are related to the Abraham-Minkowski controversy [1, 10, 3], our argument itself does not depend on any specific choice of energy/momentum formulae. The only essential point is that the interactoin can make a massive particle from a massless one. That is, while a free photon satisfies

$$E_{\text{free}}^2 - c^2p_{\text{free}}^2 = 0, \quad (4)$$

an interacting photon satisfies

$$E^2 - c^2p^2 = m_{\text{eff}}^2c^4 > 0. \quad (5)$$

To sum up, an “interacting photon” can gain a positive effective mass, while a “free photon” remains massless! This is the key we have sought for. However, the argument in this section is based on the assumption that “a photon dressed with interaction” can be viewed as a single particle. Then we proceed to consider the validity of our picture, especially the existence of particles whose effective mass is obtained by the interaction, analogous to Higgs mechanism: Such a universal model for photon localization exists. It is the notion of polariton, well known in optical and solid physics.

4.2 Polaritons as a universal model for photon localization

In optical and solid-state physics, the propagation of light in a medium is viewed as follows: By the interaction between light and matter, creation of

an “exciton (an excited state of polarization field above the Fermi surface)” and annihilation of a photon will be followed by annihilation of an exciton and creation of a photon, ..., and so on. This chain of processes itself is often considered as the motion of particles called polaritons (in this case “exciton-polaritons”), which constitute particles associated with the coupled wave of the polarization wave and electromagnetic wave.

Remark 3 *In spite of the similarity in its name, a polariton should not be confused with a polaron [7] which represents a fermion as a charged matter dressed by polarization field. In contrast, a polariton is a boson which represents a “dressed photon”.*

The notion of polariton has been introduced to develop the microscopic theory of electromagnetic interactions in materials ([6], [5]). An injected photons become polaritons by the interaction with matter. As exciton-phonon interaction is dissipative, the polariton picture gives a scenario of absorption. It has provided a better approximation than the scenarios without a polariton. Moreover, the group velocity of polaritons discussed below gives another confirmation of the presence of an effective mass.

As is well known, permittivity $\epsilon(\omega)$ is given by the following equality,

$$\epsilon(\omega) = n^2 = \frac{c^2 k^2}{\omega^2}, \quad (6)$$

and hence, we obtain the dispersion relation (a relation between frequency and wave number) of polariton once the formula of permittivity is given.

Remark 4 *In general, this dispersion relation implies branching, analogous to the Higgs mechanism. The signal pulse corresponding to each branches can also be detected in many experiments, for example, in [9] cited below.*

In the simple case, the permittivity is given by the transverse frequency ω_T of exciton’s (lattice vibration) as follows:

$$\epsilon(\omega) = \epsilon_\infty + \frac{\omega_T^2(\epsilon_{st} - \epsilon_\infty)}{\omega_T^2 - \omega^2}, \quad (7)$$

where ϵ_∞ denotes $\lim_{\omega \rightarrow \infty} \epsilon(\omega)$ and $\epsilon_{st} = \epsilon(0)$ (static permittivity). With a slight improvement through the wavenumber dependence of the exciton energy, the theoretical result of polariton group velocity $\frac{\partial \omega}{\partial \mathbf{k}} < c$ based on the above dispersion relation can explain satisfactorily experimental data of the passing time of light in materials (for example, [9]). This strongly supports the validity of the polariton picture.

From the above arguments, polaritons can be considered as a universal model of the “interacting photons in a medium” in the previous subsection 4.1. The positive mass of a polariton gives a solution to its “localization

problem”. Conversely, as the “consequence” of Wightman’s theorem, it follows that “all” physically accessible photons as particles which can be localized are more or less polaritons (or similar particles) because only the interaction can give a photon its effective mass, if it does not violate particle picture. In this way, the dilemma between Newton-Wigner-Wightman theorem and the position observable of photons is successfully resolved by combining useful mathematical methods and meaningful physical concepts, which were separated before causing a negative result.

5 Concluding remarks

The discussions in the present note has clarified the important roles played by interactions in making sense of the notion of localization. We can expect that it also shed some new light on the idea of the “emergence” of space-time proposed in [13]. In combination with a possible scenario for the mass generation, we can summarize the argument above in the following quadrality scheme:

$$\begin{array}{ccc}
 & & x : \begin{array}{l} \text{localization} \\ \text{of photons} \end{array} \\
 & & \uparrow \\
 m_{\text{eff}} : \text{Effective mass of photon} & \implies & v : \text{kinematics} \\
 \uparrow & & \\
 p : \begin{array}{l} \text{dynamical interaction} \\ \text{between photons \& media} \end{array} & &
 \end{array}$$

When all the above ingredients are established, the “mechanics of mass points” becomes meaningful. A photon, which is something quite dissimilar to a “mass point”, appears ubiquitously since the electromagnetic field works as a universal medium to mediate the interaction between charged particles which provides an idealized standard reference system.

Then, who has seen a free photon?

The answer to this original question can now be found in the following modified form of the verse at the beginning:

Who has seen a free photon?
 Neither I nor you.
 But when the matter reacts trembling
 the photons are passing through.

Acknowledgements

The authors are grateful to Prof. Shogo Tanimura, who has informed them of the importance of the dilemma. They thank Mr. Ryo Harada for many useful comments to refine drafts. They also thank Mr. Hiroshi Ando, Mr. Takahiro Hasebe, and Mr. Kazuya Okamura for their interests and comments.

References

- [1] Abraham, M., Zur Elektrodynamik bewegter Körper, Rend. Circ. Mat. Palermo **28**, 1 (1909).
- [2] Angelopoulos, E., Bayen, F. and Flato, M., On the localizability of massless particles, Physica Scripta **9**, 173 (1974).
- [3] Barnett, S.M., Resolution of the Abraham-Minkowski Dilemma, Phys. Rev. Lett. **104**, 070401 (2010).
- [4] Haag, R., Local Quantum Physics (second edition), Springer-Verlag (1996)
- [5] Hopfield, J.J., Theory of the contribution of excitons to the complex dielectric constant of crystals, Phys. Rev. **112**, 1555 (1958).
- [6] Fano, U., Atomic theory of electromagnetic interactions in dense materials, Phys. Rev. **103**, 1202 (1956)
- [7] Landau, L.D., Über die Bewegung der Elektronen in Kristallgitter, Phys. Z. Sowjetunion **3**, 644 (1933).
- [8] Mackey, G.W, Induced representations and quantum mechanics, Benjamin/Cummings (1968).
- [9] Masumoto, Y., Unuma, Y., Tanaka, Y. and Shionoya, S., Picosecond time of flight measurements of excitonic polariton in CuCl, J. Phys. Soc. Jpn. **47**, 1844 (1979).
- [10] Minkowski, H., Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern, Nach. Königl. Ges. Wiss. Göttingen **53** (1908); Math. Ann. **68** 472 (Reprinted 1910).
- [11] Newton, T.D. and Wigner, E.P. Localized states for elementary systems Rev. Mod. Phys. **21**, 400 (1949).
- [12] Ojima, I., Micro-Macro Duality in Quantum Physics, pp.143–161, in Proc. Intern. Conf. “Stochastic Analysis: Classical and Quantum”,

World Scientific, 2005; Micro-Macro duality and emergence of macroscopic levels, *Quantum Probability and White Noise Analysis*, **21**, 217 – 228 (2008)

- [13] Ojima, I., Meaning of Non-Extensive Entropies in Micro-Macro Duality, *J. Phys.: Conf. Ser.* 201 012017 (2010): Invited talk at RIMS International Workshop, “Mathematical Aspects of Generalized Entropies and their Applications”, July 2009.
- [14] Ojima, I. and Saigo, H., in preparation.
- [15] Christine G. Rossetti, “The Wind”.
- [16] Wightman, A.S., On the localizability of quantum mechanical systems, *Rev. Mod Phys.* **34**, 845 (1962).

Weak Value as Hilbert-Schmidt Bundle

Akio Hosoya^{1,*}

¹*Department of Physics, Tokyo Institute of Technology, Tokyo, Japan*

(Dated: May 31, 2011)

We present an elegant geometric picture of the weak values in the framework of the Hilbert-Schmidt bundle over the state space.

I. INTRODUCTION

In the conventional Copenhagen interpretation of quantum mechanics, the value of physical quantity emerges only after some measurements but not before. However, Ozawa [1] has pursued the possibility to assign a context dependent value for a physical quantity before measurement, where the context describes outcomes of the measurements to be performed.

Let a set of observables which act on the Hilbert space \mathcal{H} of finite dimensions ($\dim\mathcal{H} = N$) be \mathcal{N} . Let $\mathcal{V}(\mathcal{N})_{\max}$ be the set of all maximal abelian subalgebras of \mathcal{N} . We choose a maximal abelian subalgebra $V_{\max} \in \mathcal{V}(\mathcal{N})_{\max}$, noting that the choice is not unique. Then we can select a complete orthonormal system $\{|\omega\rangle\}_{\omega \in \Omega}$ as the set of all simultaneous eigenstates of all elements of V_{\max} , where $\Omega := \{\omega_1, \dots, \omega_N\}$. We introduce the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ adopting the standard notation of Kolmogorov's measure theory of probability [2] as follows. The sample space is $\tilde{\Omega} := \{\omega \in \Omega \mid \langle \omega | \psi \rangle \neq 0\}$, and the σ -field $\tilde{\mathcal{F}}$ over $\tilde{\Omega}$ is the power set of $\tilde{\Omega}$. \tilde{P} is a probability measure over $\tilde{\mathcal{F}}$.

In the recent work [3] we have demonstrated that the context dependent value $\lambda_\omega : \mathcal{N} \rightarrow \mathbb{C}$ of an observable $A \in \mathcal{N}$ coincides with the weak value

$$\lambda_\omega(A) = \frac{\langle \omega | A | \psi \rangle}{\langle \omega | \psi \rangle}, \text{ for } \omega \in \tilde{\Omega}, \quad (1)$$

advocated by Aharonov et al.[4, 5].

In the proof we assumed (i) the sum rule: $\lambda_\omega(A + B) = \lambda_\omega(A) + \lambda_\omega(B)$, $\forall A, B \in \mathcal{N}$, (ii) the product rule: $\lambda_\omega(TS) = \lambda_\omega(T)\lambda_\omega(S)$, $\forall T, S \in V_{\max}$ and (iii) initial condition: $\lambda_\omega(|\psi^\perp\rangle\langle\psi^\perp|) = 0$ for all the states $\{|\psi^\perp\rangle\}$ orthogonal to the prepared state $|\psi\rangle$. (iv) invariance: the expectation value

$$Ex(A) := \sum_{\omega \in \tilde{\Omega}} \tilde{P}(\omega) \lambda_\omega(A) \quad (2)$$

and the variance

$$Var(A) := \sum_{\omega \in \tilde{\Omega}} \tilde{P}(\omega) |\lambda_\omega(A)|^2 \quad (3)$$

should be independent of the choice of V_{\max} but depends only on the prepared state $|\psi\rangle$ and the observable A . The invariance has been already noticed in the papers [6, 7].

The bonus is Born's rule [8]. The probability measure is given by

$$\tilde{P}(\omega) = |\langle \omega | \psi \rangle|^2. \quad (4)$$

and we obtain the expectation value

$$Ex(A) = \langle \psi | A | \psi \rangle, \quad (5)$$

which is an axiom in the conventional quantum mechanics. and the variance

$$Var(A) = \langle \psi | A^2 | \psi \rangle. \quad (6)$$

*Electronic address: ahosoya@th.phys.titech.ac.jp

It is rather straightforward to generalize the previous work for pure states to mixed states. We just quote the result. Now the initial state is given by a density operator ρ and let us expand it in terms of its eigenstates $|\xi_i\rangle$ of the eigenvalue μ_i ,

$$\rho = \sum_i \mu_i |\xi_i\rangle \langle \xi_i|. \quad (7)$$

Let the probability measure for the initial state $|\xi_i\rangle$ to obtain the final state $|\omega\rangle$ be $\tilde{P}^i(\omega)$. The probability theory says that the expectation value has a form,

$$Ex(A) := \sum_{\omega \in \tilde{\Omega}, i} \tilde{P}^i(\omega) \lambda^i_\omega(A), \quad (8)$$

where $\lambda^i_\omega(A)$ is the value of the observable A in the context of $\{|\xi^i\rangle, |\omega\rangle\}$. The result is

$$\tilde{P}^i(\omega) = |\langle \omega | \xi^i \rangle|^2 \mu_i. \quad (9)$$

and

$$Ex(A) = \text{Tr}[\rho A]. \quad (10)$$

As is well known, the density operator can be alternatively expressed in terms of non-orthogonal basis $|\psi^m\rangle$ as

$$\rho = \sum_m p_m |\psi^m\rangle \langle \psi^m|, \quad p_m \in [0, 1], \quad \sum_m p_m = 1, \quad (11)$$

an interpretation of which is that the state $|\psi^m\rangle$ is prepared with the probability p_m .

There is an ambiguity of the choice [9] of $|\psi^m\rangle$ studied by Schrödinger long ago. Let us expand $|\psi^m\rangle$

$$|\psi^m\rangle = \sum_i C_i^m |\xi_i\rangle. \quad (12)$$

Here the coefficients are given by $C_i^m = \sqrt{\frac{\mu_i}{p_m}} u_i^m$, where u_i^m is a unitary matrix. The probability p_m is given by $p_m = \sum_i \mu_i |u_i^m|^2$. Therefore the state $|\psi^m\rangle$ is a non-linear function of the arbitrary unitary matrix u_i^m .

Then we have

$$Ex(A) := \sum_{\omega \in \tilde{\Omega}, m} \tilde{P}^m(\omega) \lambda^m_\omega(A), \quad (13)$$

where the probability measure $\tilde{P}^m(\omega)$ and the value of the observable A in the context of $\{|\psi^m\rangle, |\omega\rangle\}$, $\lambda^m_\omega(A)$ can be written as

$$\tilde{P}^m(\omega) = |\langle \omega | \psi^m \rangle|^2, \quad (14)$$

$$\lambda^m_\omega(A) = \frac{\text{Tr}[W^m A]}{\text{Tr}[W^m]}, \quad (15)$$

$$W^m = |\psi^m\rangle \langle \omega|. \quad (16)$$

It is important to observe that the projection $\pi : W^m \mapsto \rho$ defined by

$$\sum_m p_m W^m W^{m\dagger} = \rho \quad (17)$$

gives a natural fiber bundle picture of the space of $\{W^m\}$ with the parametrization of the fiber $u_i^m, v \in U(N)$.

II. GEOMETRY OF QUANTUM CONTEXT

We will see that the operator (16) has a beautiful geometric feature.

The weak value consists of an observable A and a prescribed context $\{|\omega\rangle, |\psi^m\rangle\}$. The context can be neatly organized as the W operator $W^m = |\psi^m\rangle \langle \omega|$ and the weak value is given by $\frac{\text{Tr}(AW^m)}{\text{Tr}(W^m)}$. The W operator can be

regarded as an element of the principal fiber bundle B over the state $\rho = \sum_m p_m |\psi^m\rangle\langle\psi^m|$ with the projection $\pi : W^m \rightarrow \sum_m p_m W^m W^{m\dagger}$ and the structure group $U(N)$. The group acts from the right, $W \rightarrow WU$, $U \in U(N)$. We call B as the context bundle. The Hilbert-Schmidt inner product is a natural metric in the context bundle so that the context bundle is a Hilbert-Schmidt bundle [10]. The induced metric in the base space of states by the projection turns out to be the standard Fubini-Study metric in the pure state case. The picture of principal fiber bundle unifies the two important aspects of quantum mechanics: probability and phase. It is straightforward to develop the standard tools of geometry; the horizontal lift or connection, which defines the parallel transport and then the geometric phase as the holonomy for a closed loop in the state space.

III. SUMMARY

We have shown that the contextual value of an observable is the weak value by demanding the consistency of quantum mechanics with Kolmogorov's measure theory of probability in conjunction with the consideration of the value of observables. The crucial requirement was that the expectation value of an observable should be independent of the way of measurements. This leads eventually to Born's rule in quantum mechanics. The contextual value has a geometrical structure of fiber bundle over the state space. The detailed analysis is in progress.

This work is done partially with Minoru Koga.

-
- [1] M. Ozawa, Quantum Reality and Measurement: A Quantum Logical Approach, to appear in Foundations of Physics, DOI:10.1007/s10701-010-9462-y and references therein.
 - [2] A.N. Kolmogorov, "Foundation of the Theory of Probability", second English translation edited by N. Morrison with an added bibliography, (A.D. Barcha-Reid University).
 - [3] A.Hosoya and M.Koga, "Weak value as Context Dependent Value of Observable and Born's Rule", ArXiv 1104.1873 (2011)
 - [4] Y. Aharonov, D. Z. Albert, and L. Vaidman, *Phys. Rev. Lett.* **60**, 1351 (1988).
 - [5] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes* (Wiley-VCH, Weinheim, 2005).
 - [6] Y.Shikano and A. Hosoya, *Phys. A:Math. Theor.* **43**, 025304 (2010).
 - [7] A. Hosoya and Y.Shikano, *Phys. A:Math. Theor.* **43**, 385307 (2010).
 - [8] M. Born, *Zeitschrift für Physik*, **37** 863 footnote of p865 (1926),(German)
 - [9] E.Schrödinger, *Proc.Camb.Phil.Soc.* **31**, 446 (1936).
 - [10] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, 2006).

Universality classes and the Anderson transition in two dimensions

Tomi Ohtsuki, Koji Kobayashi

Department of Physics, Sophia University,
Kioi-cho 7-1, Chiyoda-ku, Tokyo 102-8554, Japan

Keith Slevin

Department of Physics, Osaka University,
1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan

(Dated: May 6, 2011)

Abstract

The Anderson transition is a disorder driven quantum phase transition between metallic and insulating phases. In contrast to the common belief that two dimensional (2D) systems are always insulating and that the Anderson transition does not occur in 2D, in certain universality classes 2D systems can be metallic. We review the recent development of the theory of the Anderson transition in 2D. There are ten universality classes: three Wigner-Dyson classes, three chiral universality classes, and four Bogoliubov-de Gennes classes. We report results for critical exponents and distributions of conductance for the symplectic universality class. We emphasize that, on the one hand, the existence of a topological insulating phase does not alter the value of the critical exponent, while on the other, it strongly affects the form of the conductance distribution at the transition.

INTRODUCTION

In systems with a periodic potential, the wave functions of electrons are extended consistent with Bloch's theorem. In strongly disordered systems, however, due to destructive quantum interference, the envelope of the electron wave functions decays exponentially on a length scale called the localization length ξ ,

$$(\vec{r}) = a(\vec{r} - \vec{r}_0) \exp(-|\vec{r} - \vec{r}_0|/\xi). \quad (1)$$

Here, \vec{r}_0 is the localization center and a is a random function with a decay that is weaker than exponential. As a function of disorder, a localization-delocalization transition, called the Anderson transition [1], occurs. This transition is characterized by the divergence of the localization length ξ ,

$$\xi \sim \frac{1}{|x - x_c|^\nu}, \quad (2)$$

where x is a parameter such as Fermi energy E or the strength of the random potential W that is used to drive the transition, and ν is the critical exponent. (In the metallic phase, ξ is again finite but is there interpreted as a correlation length.) The value of the exponent ν is thought to be highly universal depending only on the universality class and the dimension (1D, 2D, 3D etc.) of the system.

The transition has been widely studied analytically [2], numerically [3], and experimentally [4], not only in semiconductors, but also in optical [5] and acoustic systems [6]. The recent development of experimental techniques that utilize Bose-Einstein condensation has shed new light on the Anderson transition [7]. In this report, we describe the classification into universality classes, and report recent numerical results for one of the universality classes, the symplectic universality class, as an example.

UNIVERSALITY CLASSES

Random Hamiltonian matrices are classified according to whether the system is invariant under the operations of time reversal T and spin rotation S . Systems with both time reversal and spin rotation symmetries comprise the orthogonal class, systems with time reversal symmetry but with broken spin rotation symmetry comprise the symplectic class, and systems with broken time reversal symmetry comprise the unitary class. (Once

T	S	Symmetry class
Yes	Yes	Orthogonal
Yes	No	Symplectic
No	not relevant	Unitary

TABLE I: Wigner-Dyson classes and symmetry

time reversal symmetry is broken, spin rotation symmetry is no longer relevant). These are called Wigner-Dyson classes [8, 9]. See Table I.

Recently, it has been found necessary to extend the Wigner-Dyson classification [10–12] to describe a wider variety of random systems such as disordered superconductors. The new classification is based on Lie algebra.

Let H be an $N \times N$ Hermitian matrix and let $X = iH$. Then, X is anti-Hermitian. Such matrices X are elements of the Lie algebra $\mathfrak{u}(N)$, and $\exp(X)$ elements of the Lie group $U(N)$. This is the unitary class in the Wigner-Dyson classification. In the absence of time reversal symmetry (or any other special symmetries) the Hamiltonian of a disordered system is in this class.

Any Hermitian matrix can be decomposed as $H = H_1 + iH_2$, where H_1 is a real symmetric matrix and H_2 a real antisymmetric matrix. The matrices H_2 are the elements of a Lie algebra that is a subalgebra of $\mathfrak{u}(N)$. The corresponding Lie group is $SO(N)$, which is a subgroup of $U(N)$. The tangent space to the symmetric space $U(N)/O(N)$ is the space of real symmetric matrices (up to a factor i). The Hamiltonians of systems with time reversal and spin rotation symmetry are of this form. This is the orthogonal class in the Wigner-Dyson classification.

When the electron spin degree is included in the description, the number of degrees of freedom is doubled and the Hamiltonian is a $2N \times 2N$ Hermitian matrix. We may decompose the Hamiltonian into 2×2 blocks c_{ij} containing matrix elements between up and down spin states and express each block in the form

$$c_{ij} = (a_{ij}^0 + ib_{ij}^0)\tau_0 + (a_{ij}^1 + ib_{ij}^1)\tau_1 + (a_{ij}^2 + ib_{ij}^2)\tau_2 + (a_{ij}^3 + ib_{ij}^3)\tau_3. \quad (3)$$

Here, $\tau_0 = 1_2$ the 2×2 identity matrix, $\tau_k = i\sigma_k$ ($k = 1, 2, 3$), where the σ_k are the Pauli matrices, and a_{ij}^k, b_{ij}^k ($k = 0, 1, 2, 3$) are real numbers. Since the Hamiltonian is Hermitian,

the coefficients $a_{i,j}$ and $b_{i,j}$ must satisfy the following conditions

$$a_{ij}^0 = a_{ji}^0, \quad a_{ij}^k = -a_{ji}^k \quad (k = 1, 2, 3), \quad b_{ij}^0 = -b_{ji}^0, \quad b_{ij}^k = b_{ji}^k \quad (k = 1, 2, 3). \quad (4)$$

Using (3) a general Hamiltonian may be decomposed into $H = H_1 + H_2$ where H_1 is a matrix with c_{ij} of the form

$$c_{ij} = a_{ij}^0 \tau_0 + a_{ij}^1 \tau_1 + a_{ij}^2 \tau_2 + a_{ij}^3 \tau_3, \quad (5)$$

and H_2 is the remainder, i.e. involving the b_{ij}^k ($k = 0, 1, 2, 3$). We may define $X = iH_2$. The matrices X satisfy

$$JX + X^T J = 0, \quad J_{ij} = i_j \tau_2, \quad (6)$$

and are the elements of a Lie algebra that is a subalgebra of $\mathfrak{u}(2N)$. The corresponding Lie group is $\text{Sp}(2N)$. The tangent space to the symmetric space $\text{U}(2N)/\text{Sp}(2N)$ is the space of matrices H_1 (up to a factor i). The Hamiltonians of systems with time reversal symmetry but where spin rotation symmetry is broken are of precisely this form. (Incidentally, this means that the Hamiltonians of such systems can be expressed as $N \times N$ matrices of quaternions, which can simplify both analytic and numerical calculations.) This is the symplectic class in the Wigner-Dyson classification.

In certain physical problems in disordered systems we encounter Hamiltonians of the form

$$H = \begin{pmatrix} 0_N & h \\ h^\dagger & 0_M \end{pmatrix}. \quad (7)$$

Here, 0_N denotes the $N \times N$ zero matrix, and h an $N \times M$ matrix. This describes the situation where the diagonal elements (potential energies) are vanishing and hopping is allowed only between different sublattices. Such a Hamiltonian satisfies

$$H = - \begin{pmatrix} 1_N & 0 \\ 0 & -1_M \end{pmatrix} H \begin{pmatrix} 1_N & 0 \\ 0 & -1_M \end{pmatrix}. \quad (8)$$

This property is called chiral symmetry [10, 11]. We may decompose a general $(N + M) \times (N + M)$ Hamiltonian into the form $H = H_1 + H_2$ where H_1 has the form (7) and H_2 is the remainder, which has the form

$$H_2 = \begin{pmatrix} h_N & 0 \\ 0 & h_M \end{pmatrix}. \quad (9)$$

Class	T	S	Symmetric space	Symbol
Wigner-Dyson	Yes	Yes	$U(N)/O(N)$	AI
Wigner-Dyson	Yes	No	$U(2N)/Sp(2N)$	AII
Wigner-Dyson	No	irrelevant	$U(N)$	A
chiral	Yes	Yes	$SO(N+M)/(SO(N) \times SO(M))$	BDI
chiral	Yes	No	$Sp(2N+2M)/(Sp(2N) \times Sp(2M))$	CII
chiral	No	irrelevant	$U(N+M)/(U(N) \times U(M))$	AIII
BdG	Yes	Yes	$Sp(2N)/U(N)$	CI
BdG	Yes	No	$SO(2N)/U(N)$	DIII
BdG	No	Yes	$Sp(2N)$	C
BdG	No	No	$SO(2N)$	D

TABLE II: Ten universality classes and corresponding Lie group (symmetric spaces). The last column is the symbol for the symmetry class.

We may then define $X = iH_2$. The matrices X are the elements of a Lie algebra with corresponding Lie group $U(N) \times U(M)$. The tangent space of the symmetric space $U(N+M)/(U(N) \times U(M))$ are precisely the Hamiltonians of the form (7) (up to a factor i).

In disordered superconductors, quasi-particles are described by a Bogoliubov-de Gennes (BdG) Hamiltonian of the form [12],

$$H = \begin{pmatrix} h & \Delta \\ -\Delta^* & -h^T \end{pmatrix}, \quad (10)$$

which satisfies

$$H = - \begin{pmatrix} 0_N & 1_N \\ 1_N & 0_N \end{pmatrix} H^T \begin{pmatrix} 0_N & 1_N \\ 1_N & 0_N \end{pmatrix} = -\tau_x H^T \tau_x, \quad (11)$$

where $1_N, 0_N$ are N -dimensional unit and zero matrices, respectively. By a unitary rotation of the basis we may define $H' = g^\dagger H g$, $g = (1_{2N} + i\tau_x)/\sqrt{2}$ that satisfies $H' = -H'^T$. We may then define matrices $X = iH'$. The matrices are elements of a Lie algebra $so(2N)$ with the corresponding Lie group $SO(2N)$.

By imposing time reversal and spin rotation symmetries, in addition to the chiral or BdG symmetries, we arrive at the ten universality classes listed in Table II.

Random Hamiltonians can be mapped to non-linear sigma models [2, 13]. Reflecting

the symmetries of the Hamiltonian, the non-linear sigma models are associated with different symmetric spaces. More details can be found in review articles such as [14–16].

ANDERSON TRANSITION IN TWO DIMENSIONS

The scaling theory of localization [17] predicts that all states are localized in 2D. The argument rests on two assumptions. First, that the conductance g (in units of e^2/h) of a system of size L obeys a single parameter scaling law

$$\beta(g) = \frac{d \log g(L)}{d \log L}. \quad (12)$$

And, second, that this β -function is monotonic. For a d -dimensional system, perturbation theory in $1/g$ gives $\beta(g) \approx d - 2 - a/g$ for large g . While for small g , where $g(L) \sim \exp(-bL/\xi)$, we expect $\log(g)$. (Here, a, b are numerical factors of the order of 1.) If we interpolate the beta function between large and small g , assuming it to be monotonic, we find that $\beta(g) < 0$ for all g , and hence that the conductance always vanishes in the limit of large L .

However, this argument applies only to the Wigner-Dyson orthogonal class and, in fact, this is the only symmetry class where in 2D the states are always localized and $g(L)$ scales to zero. All the remaining universality classes listed in Table II permit the existence of delocalized states in 2D.

Wigner-Dyson symplectic class

The Wigner-Dyson symplectic class is realized in systems with spin-orbit scattering, which conserves time reversal symmetry but breaks spin rotation symmetry. One of the model Hamiltonians used in the study of this class is the so called SU(2) model [18],

$$H = \sum_{i,\sigma} \epsilon_i c_{i\sigma}^\dagger c_{i\sigma} - \sum_{(i,j)\sigma,\sigma'} R(i,j)_{\sigma,\sigma'} c_{i\sigma}^\dagger c_{j\sigma'}, \quad (13)$$

with i, j the site indices, σ, σ' the spin indices, (i, j) pairs of nearest neighboring sites, and $R(i, j)$ a random SU(2) matrix uniformly distributed with respect to the group invariant measure. The on-site random potential ϵ_i is uniformly distributed in the range

$$-\frac{W}{2} < \epsilon_i < \frac{W}{2}. \quad (14)$$

In contrast to other models, where corrections to scaling need to be taken into account [19], the SU(2) model in 2D exhibits vanishingly small corrections to scaling. This property enables high precision finite size scaling analyses of the Lyapunov exponent [3, 20]. The critical exponent ν for SU(2) model has been found to be [18, 21],

$$\nu = 2.75 \pm 0.01. \quad (15)$$

At the transition, where the length scale ξ diverges, the correlation function of the local density of states $\langle (E, \vec{r}) (E, \vec{r}') \rangle$ has a power law decay

$$\langle (E, \vec{r}) (E, \vec{r}') \rangle \sim |\vec{r} - \vec{r}'|^{-2\eta}, \quad (16)$$

while in a quasi-1D system it has an exponential decay,

$$\langle (E, \vec{r}) (E, \vec{r}') \rangle_{\text{q1D}} \sim \exp(-2|\vec{r} - \vec{r}'|/\xi_{\text{q1D}}). \quad (17)$$

It follows from the assumption of conformal invariance at the critical point that these decays are related by [22, 23]

$$\frac{\xi_{\text{q1D}}}{L} = \frac{1}{\eta}, \quad (18)$$

with L the width of quasi-1D strip. In numerical simulations of the Anderson transition, the finite size scaling analysis of the Lyapunov exponent λ , which is estimated using the transfer matrix method [3, 20, 24], has proved to be the most powerful and precise approach. For the Lyapunov exponent it has been claimed [25, 26] that conformal invariance implies a different relation of the form

$$\lambda L = (\alpha_0 - 2), \quad (19)$$

with α_0 the position of the maximum of the multifractal spectrum $f(\alpha)$ of the distribution of the wave function (for a review, see for example [16]). There is strong numerical evidence in support of this relation [18, 27]. This suggests that the critical point in the symplectic class has conformal symmetry.

\mathbb{Z}_2 topological phase

In some systems in the Wigner-Dyson symplectic class, there appear edge states in the insulating phase that are stable against perturbations that preserve time reversal

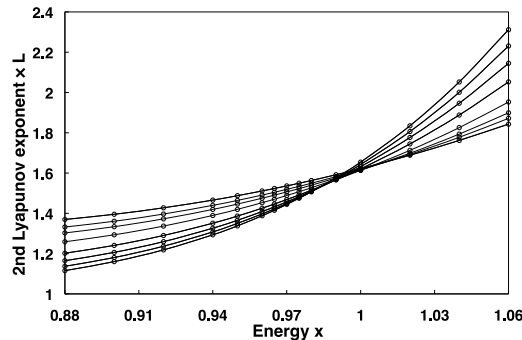


FIG. 1: The product of the second smallest positive Lyapunov exponent and the system size versus energy for the \mathbb{Z}_2 network model near the metal-topological insulator transition. The lines are a finite size scaling fit, with different lines corresponding to different system sizes. The critical exponent, $\nu = 2.77 \pm 0.06$ [30], is consistent with (15).

symmetry. Such edge states carry current, resulting in quantized conductance $2e^2/h$, which is observed experimentally [28]. If there are even numbers of edge states at one edge of a system, they are mixed and back scattered, resulting in vanishing conductance. On the other hand, in the case of an odd number of edge states, one edge state survives and the conductance is again $2e^2/h$. This insulating phase is called the \mathbb{Z}_2 topological insulator.

The critical exponent of the metal-topological insulator transition was conjectured to be the same as for the ordinary metal-insulator transition [29] and this has recently been confirmed by an analysis of the second smallest Lyapunov exponent [30] (Fig 1).

In disordered systems, the conductance is a statistical quantity and not the value but the distribution of conductance is important. (In this sense, the conductance that appears in the β -function should be regarded as a suitable average.) The distribution function, $P(g)$ should obey the scaling form [31],

$$P(g) = f(g, L/\xi). \quad (20)$$

At the Anderson transition $\xi \rightarrow \infty$, the distribution becomes system size independent. This size independent distribution is called the critical conductance distribution and is denoted by $P_c(g)$. It has been demonstrated that $P_c(g)$ is universal, i.e. model independent (see, for example, [32]). However, the form of this universal distribution has been found to be different depending on whether the adjacent insulating phase is a topological insulator or not [32] (Fig. 2).

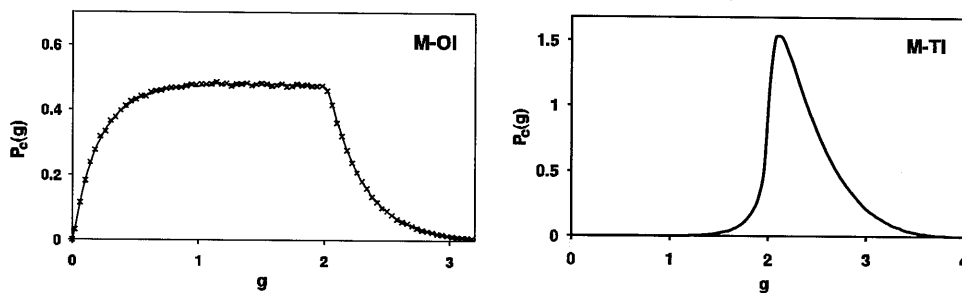


FIG. 2: Critical conductance distributions at the metal-ordinary insulator (M-OI) transition and at the metal- \mathbb{Z}_2 topological insulator (M-TI) transition. The distributions shown are for the \mathbb{Z}_2 network model (solid lines) and the $SU(2)$ model (\times). See [32] for details.

DISCUSSION

The properties of the Anderson transition in 2D systems in the Wigner-Dyson symplectic class (multifractal distribution of the wave function amplitude, conformal invariance of the critical theory, universality of the conductance distribution etc.) are conjectured to apply also to other universality classes. The most important case is the unitary class where the Anderson transition is the quantum Hall transition [33]. The quantum Hall insulator is a \mathbb{Z} topological insulator, where the current is carried by an integer number of edge states. Confirmation of conformal invariance, however, is still pending. The central difficulty is that corrections to scaling vanish only very gradually with increasing system size [34]. Further efforts, both numerical as well as analytic, will be needed to understand the quantum Hall transition quantitatively.

-
- [1] P. Anderson, *Phys. Rev.*, **109**, 1492 (1958).
 - [2] K. Efetov, *Supersymmetry in disorder and chaos* (Cambridge Univ. Press, 1997).
 - [3] B. Kramer, A. MacKinnon, *Rep. Prog. Phys.*, **56**, 1469 (1993).
 - [4] K. Itoh *et al.*, *J. Phys. Soc. Jpn.*, **73**, 173 (2004).
 - [5] S. John, *Phys. Today*, **44**, 32 (1991).
 - [6] S. Faez, A. Strybulevych, J.H. Page, Ad Lagendijk, B.A. van Tiggelen, *Phys. Rev. Lett.*, **103**, 155703 (2009).
 - [7] J. Chabe *et al.*, *Phys. Rev. Lett.*, **101**, 255702 (2008).

- [8] E.P. Wigner, Proc. Cambridge Philos. Soc., **47**, 790 (1951).
- [9] F.J. Dyson, J. Math. Phys., **3**, 140,157,166 (1962).
- [10] R. Gade, F. Wegner, Nuclear Physics B, **360**, 213 (1991).
- [11] R. Gade, Nuclear Physics B, **398**, 499 (1993).
- [12] A. Altland, M.R. Zirnbauer, Phys. Rev. B, **55**, 1142 (1997).
- [13] S. Hikami, Phys. Rev. B, **24**, 2671 (1981).
- [14] M.R. Zirnbauer, arXiv:math-ph/9808012.
- [15] B. Kramer, T. Ohtsuki, S. Kettemann, Phys. Rep., **417**, 211 (2005).
- [16] F. Evers, A.D. Mirlin, Rev. Mod. Phys., **80**, 1355 (2008).
- [17] E. Abrahams *et al.*, Phys. Rev. Lett., **42**, 673 (1979).
- [18] Y. Asada, K. Slevin, T. Ohtsuki, Phys. Rev. Lett., **89**, 256601 (2002).
- [19] K. Slevin, T. Ohtsuki, Phys. Rev. Lett., **82**, 382 (1999).
- [20] A. MacKinnon, B. Kramer, Z. Phys., **B53**, 1 (1983).
- [21] Y. Asada, K. Slevin, T. Ohtsuki, Phys. Rev. B, **70**, 035115 (2004).
- [22] J.L. Cardy, J. Phys. A, **17**, L385 (1984).
- [23] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, 1996).
- [24] J-L. Pichard, G. Sarma, J. Phys. C, **14**, L127 (1981).
- [25] A. Dohmen, P. Freche, M. Janssen, Phys. Rev. Lett., **76**, 4207 (1996).
- [26] M. Janssen, Phys. Rep., **295**, 1 (1998).
- [27] H. Obuse, R. Subramaniam, A. Furusaki, I. A. Gruzberg, A.W.W. Ludwig, Phys. Rev. Lett., **98**, 156802 (2007).
- [28] M. König *et al.*, Science, **318**, 766 (2007).
- [29] H. Obuse, A. Furusaki, S. Ryu, C. Mudry, Phys. Rev. B, **76**, 075301 (2007).
- [30] K. Kobayashi, T. Ohtsuki, K. Slevin, unpublished.
- [31] B. Shapiro, Phys. Rev. Lett., **65**, 1510 (1990).
- [32] K. Kobayashi, T. Ohtsuki, H. Obuse, K. Slevin, Phys. Rev. B, **82**, 165301 (2010).
- [33] M. P. K. von Klitzing, G. Dorda, Phys. Rev. Lett., **45**, 494 (1980).
- [34] K. Slevin, T. Ohtsuki, Phys. Rev. B, **80**, 041304(R) (2009).

植物の生き方と情報ネットワーク

東京理科大学 理工学部 応用生物科学科 朽津 和幸

E-mail: kuchitsu@rs.tus.ac.jp

植物は、脳のような中央統御型情報処理システムを進化させて来なかったが、環境変化を巧みに察知して、自らの体を移動させることなく体を再構築して適応するという独自の情報処理システムを進化させて来た。最近、植物の情報の処理・伝達を担う分子やその生成・制御の素過程が急速に明らかにされつつある。私たちは、植物が多種多様なストレスを感知して自らの体を再構築し適応する過程において、情報伝達素子としての Ca^{2+} と活性酸素種(ROS)に着目し、濃度の時間的・空間的制御機構や、情報を他の化学信号に変換し機能を発現する機構の解明を進めている。こうした素過程の知見に基づいて、植物の情報処理ネットワークの全体像を構築する研究を推進することが今後の大きな課題であり、環境・食糧・エネルギー問題等への貢献を視野に入れた新世代の植物バイオテクノロジーを展開する上でも重要な基礎となろう。

1. はじめに

生命の歴史の中で生物は多様化し、動植物は共通の祖先からそれぞれ独自の戦略を採用して高度な進化を遂げて来た。これまで動物、植物、微生物は、それぞれの専門家によって独立に研究される傾向にあった。しかし近年、ヒトのゲノムと並行してシロイヌナズナ、イネ、ミヤコグサ、ブドウ、ポプラ等多くの高等植物を始め、非常に多くの生物種のゲノムが解読されるなど、ゲノム解析が急激に進展した結果、ゲノム情報を多様な生物間の架け橋として利用することにより、研究成果を統合し、遺伝子機能の研究に新しい光を当てることが出来る時代が到来した。すなわち、生命科学は、情報科学の技術を活用して、細分化されたさまざまな分野が新たに統合される時代を迎えつつある。

生物の生命活動を支える上で、生体分子の代謝ネットワークと同時に、情報伝達ネットワークが重要な役割を果たす。実際、これまで機能面から分類されていた種々の生命現象の多くが、本質的に情報伝達の問題に帰着されることが明らかとなって来ている。

植物は、動物とは異なる適応戦略を持って進化して来た。実際、遺伝子の全塩基配列が決定されたシロイヌナズナやイネ等の植物のゲノム上に存在する情報伝達関連の遺伝子を動物と比較すると、それぞれが異なる遺伝子群を多様化させていることがわかる。多様な環境変化を察知する能力を発達させて来た植物が獲得した独自の情報伝達機構を研究することは、基礎、応用の両面から今後ますます重要性を増すと考えられる。

2. 植物の生体防御と情報の処理・伝達 ～植物に免疫・神経はあるか？

植物は、動物のように移動して不利な環境から逃げる代わりに、病原菌の感染を認識し、生体防御応答を誘導するなど、悪環境や外敵から自分を守る巧妙な仕組みを進化の過程で獲得して来た。近年「植物免疫」という概念が提唱され、動

植物の自然免疫システムの一部が、同一の起源から独立に進化した可能性も議論されている。病原体の増殖と拡散を阻止する感染部位局所的な防御応答としては、活性酸素種(reactive oxygen species; ROS)の生成、細胞壁中の糖鎖の架橋反応を介した細胞壁の物理的強化、二次代謝系の活性化による種々の抗菌性低分子物質の合成、病原体を攻撃する種々のタンパク質の合成、病原体を道連れにして封じ込めるための局所的なプログラムされた細胞死などが含まれる。

植物は病原体由来の種々の物質を認識し、一連の防御応答を誘導する。こうしたシグナル物質の中には、病原体の表面に存在する、または分泌されるタンパク質またはペプチド、病原体の表面から遊離されるオリゴ糖や糖脂質などがあり、細菌の鞭毛タンパク質や糖脂質など、哺乳動物の Toll 様受容体で認識されるような pathogen-associated molecular pattern (PAMP; 病原体関連分子パターン)と共通の物質も多い。また植物が誘導する抵抗性反応は、病原体に限らず、広く微生物一般により誘導されることから、PAMP の概念をさらに広げて、microbe-associated molecular pattern (MAMP)と呼ぶことも提唱されている。

MAMP の多くは、動植物自身が持たない物質であり、非自己の認識に基づいて、一般的な生体防御応答を誘導する機構が動植物に共通して存在すると考えられる。自然界に存在する無数の微生物のうち、特定の植物に感染できるものは、ごくわずかで、大部分のものは植物の非特異的・基礎的な抵抗性反応によって撃退される。逆に、病原体は、植物の基盤的抵抗性反応を打破する機構を備えることによって初めて感染が可能になる。

植物は、細胞外(細胞壁中)に、微生物の細胞表層因子(細胞壁多糖等)の加水分解酵素を分泌している。病原菌の感染に伴い、こうした加水分解酵素が微生物の細胞表層に対して働き、MAMP としてオリゴ糖を遊離させる。それらが植物の細胞膜上の MAMP に対する受容体タンパク質によって認識される。一方 MAMP の下流で発現が誘導されるタンパク質の中には、chitinase や β -glucanase 等の微生物の細胞表層因子の加水分解酵素が含まれることから、植物は正のフィードバック機構により感染シグナルを増幅し、自然免疫応答を活性化すると考えられる。

私たちは、植物培養細胞が病原菌由来の感染シグナル分子を認識し、自律的な細胞死を含む防御応答を誘導するモデル実験系を構築し、植物が誘導する免疫・生体防御系の制御機構を解析して来た。感染シグナルの認識に伴い、膜電位の脱分極[1]や、特徴的なパターンを示す Ca^{2+} の動員[2]と過酸化水素(H_2O_2)などの ROS の生成[3]が誘導され、 Ca^{2+} と ROS が生体情報の担い手として、防御応答の制御に重要な役割を果たすことが明らかとなった。細胞膜の電位変化(一過性の脱分極)の波形は、神経細胞の興奮に伴う活動電位と類似しているが、植物のゲノム中には Na^+ チャンネルは存在せず、電気化学的信号の発生機構は大きく異なると考えられる。植物細胞における膜電位の脱分極においては、 Ca^{2+} チャンネルと陰イオンチャンネルが重要な役割を果たすと考えられる。

3. 植物の情報伝達ネットワークの物質的基盤

3.1. 情報を担う素子としてのカルシウムイオン

一般に生物は、受容体分子で外界の情報を認識し、一連の情報伝達カスケードあるいはネットワークを介して、核における遺伝子発現その他の細胞機能の調節を制御する。この過程では、タンパク質同士の相互作用と並んで、二次メッセンジャーと総称される低分子物質が情報の担い手として重要な働きを担っている。種々の植物ホルモンに対する応答、乾燥、浸透圧、温度、機械刺激などの環境ストレス、病原体の感染や微生物の共生に伴う植物微生物相互作用など、多種多様な刺激の伝達に、 Ca^{2+} が二次メッセンジャーとして関与することが示唆されている。

さまざまな刺激からのシグナルが Ca^{2+} 濃度変化に変換されるとすれば、それぞれの情報がどのように識別されるのかは、非常に大きな課題である。最近、細胞質の Ca^{2+} 濃度の時間的・空間的な動態を生きたまま解析する手法が工夫され、その時間的・空間的なパターンの違いを見分けて情報が識別されていることが明らかになりつつある。

一般に植物の細胞外の Ca^{2+} 濃度は $10^{-3} \sim 10^{-2}$ M 程度と高く、また細胞内の主たる貯蔵器官である液胞にも、細胞外と同程度の濃度の Ca^{2+} が存在している。反面、細胞質の Ca^{2+} 濃度は、動物細胞同様 10^{-7} M 程度と、細胞外の1万分の1程度に保たれている。その結果、細胞外や細胞内小器官(細胞内 Ca^{2+} 貯蔵庫)からわずかな量の Ca^{2+} が細胞質に流入することで、細胞質の Ca^{2+} 濃度が大きく上昇し、 Ca^{2+} が二次メッセンジャーとして情報を担うことを可能にしている。この濃度(電気化学ポテンシャル)勾配の維持には、主として細胞膜や液胞膜等に存在する輸送分子を介した能動輸送機構が関与している。

3.2 カルシウムイオン濃度変化の時空間パターンの形成 (coding)

情報伝達の過程ではしばしば Ca^{2+} 濃度の急激な変化が誘導される。生体膜を貫通する Ca^{2+} チャネルタンパク質が開口すると、電気化学ポテンシャル勾配に従って Ca^{2+} の flux が起きる。細胞内局在部位や、制御機構の異なる複数の Ca^{2+} チャネルが、 Ca^{2+} 濃度変化の時空間パターンを形成し、生体情報が Ca^{2+} シグナルに変換(coding)される。

Ca^{2+} シグナル伝達の研究においては、主に化学物質を用いた薬理学とパッチクランプ法に代表される電気生理学とを組み合わせた研究が先行し、種々のイオンチャネルの電気生理学的性質が解析されて来た。しかし、 Ca^{2+} シグナルが細胞内でどのように仕分けされてさまざまな応答反応を引き起こすのか、細胞質の Ca^{2+} 濃度の時間的・空間的なパターンがどのような分子機構により形成されるかの理解は遅れている。その主な原因は、特に植物において Ca^{2+} 動員に関与する Ca^{2+} チャネルの分子の実体がなかなか明らかにされなかったため、分子遺伝学的解析が遅れている点にある。

近年、シグナル伝達系に関与する Ca^{2+} チャネルの遺伝子がようやく同定され始め、 Ca^{2+} 動員制御の分子機構の研究が本格的に開始されつつある。ゲノム解析の進展や、豊富な変異系統から得られる情報、そしてバイオイメージング技術の発展により、それぞれの組織、細胞における複数の情報伝達機構の存在や、植物独自

の Ca^{2+} 動員制御機構が徐々に明らかになりつつある。

3.3 カルシウムシグナルの decoding と情報伝達分子スイッチシステム

Ca^{2+} シグナル伝達系の下流では、種々の Ca^{2+} センサータンパク質が Ca^{2+} と結合し、情報ネットワークの素子として情報の decoding に関与する。植物のゲノム中には、哺乳動物以上に多様な Ca^{2+} 結合領域(EF-hand)を持つ多様な Ca^{2+} 結合タンパク質が存在する。calmodulin (CaM)は、ヒトには基本的に1種類しか存在しないが、植物には少なくともアミノ酸配列の異なる3種が存在し、機能分担し、生体防御系においても重要な役割を果たす[4-6]。

生体の情報伝達系では、タンパク質の特定のアミノ酸がリン酸化あるいは脱リン酸化されることによりタンパク質の機能がON/OFF制御されることが多い。これは、タンパク質リン酸化を介した情報伝達分子スイッチシステムであり、スイッチをOFFからONにする際には、タンパク質リン酸化酵素(protein kinase)が関与する。植物のゲノム中には、プロテインキナーゼドメインのC末端側に、CaMと相同な Ca^{2+} 結合領域(EF-hand)を持つ、CaM domain (Ca^{2+} -dependent) protein kinase (CDPK)が大きなファミリーを形成して存在しており、生体防御応答にも関与することが明らかになりつつある。

最近私たちは、RNA干渉(RNAi)法によりプロテインキナーゼ OsCIPK14/15の発現を抑制したイネの培養細胞では、カビ由来のタンパク質 TvX により誘導される過敏感細胞死、防御遺伝子の発現誘導などの応答が顕著に抑制されることを見出した。逆に OsCIPK14/15 過剰発現株では、TvX 誘導性の過敏感細胞死が亢進された。こうした結果は、イネの感染防御応答やプログラム細胞死の制御に、 Ca^{2+} 制御型 OsCIPK14/15 が関与する可能性を示唆する[7]。OsCIPK14/15は、FISLモチーフを介して Ca^{2+} センサータンパク質 OsCBL4 と強く相互作用する。つまり CBL-CIPK系は、CDPK系と並んで、 Ca^{2+} シグナルをリン酸化シグナルに変換する役割を担う。CBL (calcineurin B like protein)は、動物の免疫系で重要な役割を果たす calcineurin と呼ばれるタンパク質の Ca^{2+} センサーサブユニットとよく似ており、共通の祖先タンパク質から進化した ortholog (相同タンパク質) と考えられる。動植物の免疫系に類似の因子が機能していることは興味深い。

3.4 活性酸素種(ROS)～毒? 薬? 情報?

酸素呼吸に伴い必然的に生成される ROS は、多くの生体物質を酸化し、生体にさまざまな害を与える。実際、さまざまな生活習慣病を始めとしてヒトの多くの疾患が ROS の生成と密接に関連することが明らかにされつつある。ROS の生成は、酸素呼吸を行う生物にとって一つの宿命とも言える。さらに植物は、葉緑体において光合成を行う際、水を光分解して酸素を発生する。一方葉緑体内には、 CO_2 固定反応のエネルギー源とするため、NADPH などの還元力の高い物質が蓄積される。そのため、何らかのストレスが加わるなどして光合成反応のバランスが崩れると、葉緑体内で大量の ROS が生成される。植物の生存にとって、高い ROS 分解能力を持つことが必須となる。植物がビタミン C に代表される抗酸化物質を大量に蓄積しているのはこのためである。

一方で動植物は、ROS を積極的に生成する。例えば動植物の自然免疫系(病原体

に対する生体防御応答)の初期過程で、細胞表層で ROS が生成される。この反応は、oxidative burst と呼ばれ、生成された ROS は、病原菌を直接攻撃するだけでなく、植物では糖タンパク質や多糖類の架橋反応を司る細胞壁中の peroxidase の基質として細胞壁の強化に寄与し、二次的な感染に備えている。植物の細胞壁は、分子架橋の生成と切断の制御を通して、その堅さ、柔らかさを巧みに調節している。細胞伸張の際には、細胞壁が緩むのに対して、完成防御応答の過程では、細胞壁を堅くして、病原体の細胞内への進入を防ぐ。

同時に ROS は、redox (酸化還元)制御系を介して、下流の防御応答や細胞死を制御すると考えられる。さらに植物の環境ストレス応答や発生、分化、形態形成などの過程においても、ROS は積極的に生成され、シグナル物質として機能することが明らかにされつつある(図 1)。

病原菌認識後に誘導される ROS 生成は、タンパク質リン酸化反応や Ca^{2+} 流入の下流で誘導され、植物の防御応答に重要な役割を果している。細胞膜上の Nox を介した O_2 の発生は、哺乳動物の好中球や食細胞における ROS の生成と一見類似している。実際、植物には、哺乳動物の食細胞に存在する NADPH 酸化酵素(NADPH oxidase2; Nox2)に類似の構造を持つ遺伝子 *Respiratory burst oxidase homolog (Rboh)* が多数存在する。*Rboh* 遺伝子にコードされる Rboh タンパク質は、NOX2 と同様に、6 箇所の膜貫通領域と、触媒ドメインを形成する NADPH, FAD 結合モチーフなどを持つが、その一方で Nox2 とは異なり、N 末端に Ca^{2+} 結合モチーフ (EF-hand) 2 個を含む長い細胞質領域を持つ(図 2)。また、哺乳動物の Nox2 や Nox1 の活性制御に必要な活性調節因子は、低分子量 G タンパク質 Rac を除いて植物に存在しない。

植物の Nox には、恒常発現型と誘導型の双方が存在する。感染シグナル認識後の ROS 生成は、しばしば二相性のパターンを示すが、初期の ROS 生成は、Nox の活性化、後期の大量の ROS 生成は、誘導型の Nox の発現誘導を介した反応と考えられる。また、細胞膜上の NADPH oxidase 以外に、細胞壁中に存在する peroxidase

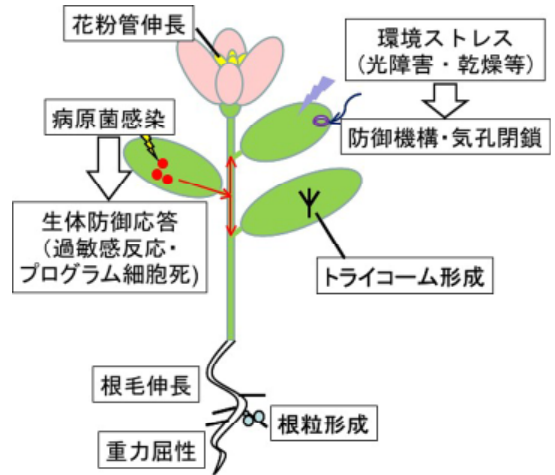


図 1 植物における活性酸素種 (ROS) の積極的生成の生理的意義

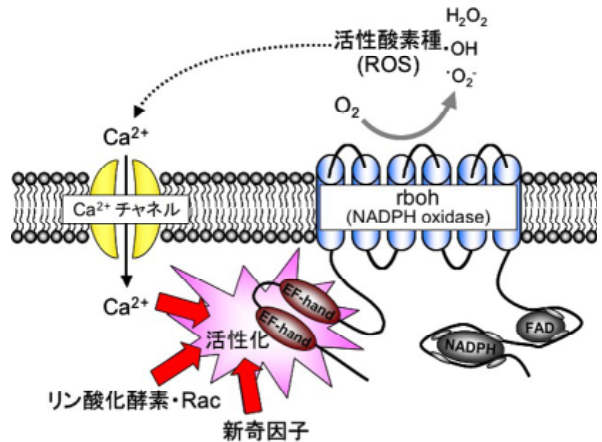


図 2 植物の活性酸素種(ROS)生成酵素 rboh を介した ROS と Ca^{2+} との相互依存的な正のフィードバック制御

などの酵素を介した ROS 生成機構も提唱されている。

感染シグナルの認識後、しばしば局所的な細胞死が誘導される。この細胞死は、病原体の感染による受動的な死ではなく、抵抗性発現のために遺伝的にプログラムされた積極的、自律的な細胞死である。この自発的な細胞死は、病原体を道連れにしてその増殖を抑制すると共に、感染シグナルを周囲の細胞に伝達する役割を担っていると考えられる。ROS 生成は、この細胞死の引き金として機能すると考えられているが、詳細な制御機構は不明であり、ROS を介した情報伝達系に関する今後の研究が待たれる。

3.5 カルシウムイオンと活性酸素の相互依存的な正のフィードバック制御機構

植物の Nox である Rboh の活性化機構は、未解明の点が多い。そこで私たちは、導入した遺伝子由来の ROS 生産活性を測定できる HEK293T 細胞発現系を確立し、Rboh の活性制御機構を解析した。AtRbohD を一過的に発現させた HEK293T 細胞に、ionomycin (生体膜に作用して Ca^{2+} を透過させる試薬; Ca^{2+} ionophore) を添加したところ、迅速な ROS 生成が観察された。この ROS 生成は、ionomycin の量に依存して増加した。このことは、Rboh が Ca^{2+} によって活性化されることを意味する。逆の視点から見ると、ROS 生成酵素である Rboh は、情報伝達系において Ca^{2+} センサータンパク質として機能すると考えられる。

そこで、AtRbohD の EF-hand モチーフ内のアミノ酸に Ca^{2+} 結合能を低下させるような部位特異的変異体 (AtRbohD-E277Q, AtRbohF-E288Q) を作製し、ROS 生産活性に対する影響を調べた。その結果、双方の EF-hand 変異体で ROS 生成が著しく減少した。同時に、AtRbohD の EF-hand 領域を含む 85 アミノ酸残基からなる組換えタンパク質を調整し、蛍光・円偏光二色性スペクトル法により Ca^{2+} 結合能と二次構造変化を解析したところ、 Ca^{2+} 濃度依存的な立体構造変化が観察された。E277Q 変異体では Ca^{2+} 依存的な立体構造変化が抑制されるなど、AtRbohD の EF-hand 変異体による ROS 生産活性と立体構造変化との間に相関関係が見られた。これらのことは、Rboh が、EF-hand 領域への Ca^{2+} の結合とそれに伴う分子内構造変化により活性化されることを示唆する。

病原菌の感染を認識して植物が ROS 生成を誘導する際、タンパク質リン酸化反応が関与することが薬理的な実験から示されていた[3]。AtRbohD を発現させた細胞に、タンパク質脱リン酸化酵素阻害剤 calyculin A (CA) を処理したところ、持続的な ROS 生成が観察された。この ROS 生成は、細胞外の Ca^{2+} を除いたときにも観察された。加えて AtRbohD の EF-hand の Ca^{2+} 結合能を低下させた変異体タンパク質 (ARbohD-E277Q) を発現させた細胞でも観察された。これらのことは、タンパク質リン酸化が、 Ca^{2+} とは独立に AtRbohD を活性化していることを示唆している。また、AtRbohD を発現した細胞に CA を前処理した後、ionomycin を添加したところ、ionomycin 単独で誘導される ROS 生成に比べて、より強い ROS 生成が観察された。この CA による増強は、AtRbohD-E277Q では観察されなかった。これらのことは、AtRbohD の ROS 生成活性において、 Ca^{2+} とタンパク質リン酸化が相乗的に働いていることを示している[8]。

さらに、植物の感染防御応答や環境ストレス応答で重要な働きを担う AtRbohD,

AtRbohF の活性制御機構を詳細に解析したところ、 Ca^{2+} の結合による活性化にはタンパク質リン酸化が必須だが、タンパク質リン酸化による活性化は EF-hand 領域への Ca^{2+} の結合を必要としないことが明らかとなった[9]。

一方、根の養分吸収に関与する根毛が形成される際、Nox の一種で AtRbohC が先端成長部位に局在し機能することが明らかとなった。AtRbohC の欠損突然変異体では根毛の伸長が著しく抑制される。突然変異体に同様に異種発現系を用いた解析を進めたところ、AtRbohD, F と同様に、 Ca^{2+} とタンパク質リン酸化により相乗的に活性化されることが明らかとなった。一方で ROS は、生体防御、根毛伸長の双方の系で、 Ca^{2+} チャンネルを活性化し、 Ca^{2+} の流入を促す。以上の結果から私たちは、 Ca^{2+} と ROS が双方向の正のフィードバック制御を受け、生体情報の増幅に関与するとのモデルを提唱した(図 2)[9, 10]。こうした機構は、感染防御応答、乾燥や低温などの環境ストレスに対する応答、根毛の先端成長の制御など多様な局面で情報の処理・伝達を司っていると考えられる。

こうした情報伝達素子が生体においてどのような機能を担うかを解明するためには、特定の遺伝子の機能を破壊する逆遺伝学的手法が有効である。現在私たちは、生体における ROS を介した情報ネットワークの生理的意義の全体像を捉えるべく、モデル植物シロイヌナズナにおいて ROS シグナルの発生に関与すると考えられる 10 種の Rboh 遺伝子の突然変異体の網羅的な解析を進めている。

4. むすび

植物においても、情報の処理・伝達を担う分子(情報伝達系における素子)や、その生成・制御の素過程が急速に明らかにされつつある。植物は、脳のような中央統御型情報処理システムを進化させて来なかったが、環境変化を巧みに察知して、自らの体を移動させることなく自分の体を再構築して適応するという独自の情報処理システムを進化させて来た。素過程の知見に基づいて、植物の情報処理ネットワークの全体像を構築する研究を推進することが今後の大きな課題であり、環境・食糧問題の解決を目指して、植物の耐病性、環境ストレス耐性を高め生産力を向上させる新世代の植物バイオテクノロジーを展開する上でも重要な基礎となろう。

植物を理解し、植物の持つ潜在能力を生かし、人間が植物とうまく付き合っていくことは、今後の人類にとって必要不可欠な課題である。最近、植物の免疫力を高めることにより、殺菌剤殺虫剤等の化学農薬の使用を軽減する試みが注目を集めている。品種改良や遺伝子組換え技術などによる耐病性の遺伝的増強と、植物の免疫力を高める生理活性物質の探索・利用の双方のアプローチは、次世代の農業を大きく変える可能性を秘めている。本研究の過程で私たちの研究グループも最近、植物の免疫力を高める新規の物質を発見し、作用機構の解明と実用化の検証を進めている。こうした研究を進める上での根幹は、植物が独自に進化させて来た、動物とは全く異なる免疫システムの情報ネットワークを分子レベルで理解することにある。植物の免疫応答における情報伝達機構を研究することは、今後ますます重要性を増すと考えられ、研究の発展が期待される。

また、研究の過程で、数理物理学、情報科学分野の研究者と対話を進めて来た。植物の情報ネットワークを量子生命情報の視点から捉える研究はこれまでほとんど行われて来なかったが、今後学際的研究を進めることで予想外の展開が期待で

きるかも知れない。2010年11月に京都大学基礎物理学研究所で開かれた研究会「量子科学における双対性とスケール」や2010年3月、2011年3月に東京理科大学野田キャンパスで開かれた量子生命情報国際会議(IC-QBIC)で依頼された講演の後の反響の大きさは、これまでになかった新たな学際研究の萌芽を予感させた。こうした意外性に富む学問の新展開を期待して、学際的な対話を継続することの重要性を痛感している。

References

- [1] K. Kuchitsu *et al.*, *N*-Acetylchitoooligosaccharides, biotic elicitor for phytoalexin production, induce transient membrane depolarization in suspension-cultured rice cells. *Protoplasma* **174**, 79-81 (1993).
- [2] Y. Kadota *et al.*, Cryptogein-induced initial events in tobacco BY-2 cells: pharmacological characterization of molecular relationship among cytosolic Ca^{2+} transients, anion efflux and production of reactive oxygen species. *Plant Cell Physiol.* **45**, 160-170 (2004).
- [3] K. Kuchitsu *et al.*, EPR evidence for generation of hydroxyl radical triggered by *N*-acetylchitoooligosaccharide elicitor and a protein phosphatase inhibitor in suspension-cultured rice cells. *Protoplasma* **188**, 138-142 (1995).
- [4] E. Karita *et al.*, Three types of tobacco calmodulins characteristically activate plant NAD kinase at different Ca^{2+} concentration and pHs. *Plant Cell Physiol.* **45**, 1371-1379 (2004).
- [5] S. Katou *et al.*, Catalytic Activation of Plant MAPK Phosphatase NtMKP1 by Its Physiological Substrate SIPK, but Not by Calmodulins. *J. Biol. Chem.* **280**, 39569-39581 (2005).
- [6] T. Takabatake *et al.*, Pathogen-induced calmodulin isoforms in basal resistance against bacterial and fungal pathogens in tobacco. *Plant Cell Physiol.* **48**, 414-423 (2007).
- [7] T. Kurusu *et al.*, Regulation of microbe-associated molecular pattern-induced hypersensitive cell death, phytoalexin production and defense gene expression by calcineurin B-like protein-interacting protein kinases, OsCIPK14/15, in rice cultured cells. *Plant Physiol.* **153**, 678-692 (2010).
- [8] Y. Ogasawara *et al.*, Synergistic Activation of Arabidopsis NADPH Oxidase AtrbohD by Ca^{2+} and Phosphorylation. *J. Biol. Chem.* **283**, 8885-8892 (2008).
- [9] S. Kimura *et al.*, Protein phosphorylation is a prerequisite for the Ca^{2+} -dependent activation of Arabidopsis NADPH oxidases and may function as a trigger for the positive feedback regulation of Ca^{2+} and reactive oxygen species. *Biochim. Biophys. Acta -Molecular Cell Research*, in press (2011).
- [10] S. Takeda *et al.*, Local Positive Feedback Regulation Determines Cell Shape in Root Hair Cells. *Science* **319**, 1241-1244 (2008).

量子力学における測定・相関・文脈性

名古屋大学 大学院情報科学研究科 小澤 正直

量子測定は、測定対象と測定装置の間に相関を確立することであるとよく言われるが、その正確な意味について十分な理解が得られているとは言えない。本稿では、この問題を解明するために、エンタングルメントでは表現できない、二つの物理量が異なる時刻でもつ値の相関に関する理論を展開し、量子測定が実現する対象と装置の間の相関の意味を解明する。この成果に基づいて、非可換物理量の同時測定に関する理論を展開し、その可能性と特徴付けを与える。また、これらの成果を通して、測定が達成されると言うことが量子力学的文脈性にどのように関わるかを考察する。

1 はじめに

量子測定理論の伝統的な理解によると、量子測定は、測定対象と測定装置の間に相関を確立することであるとよく言われる。しかし、その正確な意味について十分な理解が得られているとは言いがたい。つまり、伝統的な見方によると、量子測定では対象と装置の間の相互作用によって、対象と装置の合成系の状態が、対象に属する被測定量の値と装置に含まれるメータ物理量の値の間に一対一対応が成立する状態に変化するとされ、その相関こそがその相互作用が測定を意味する理由だとされる。確かに、測定後の状態において、メータ物理量を測定すれば、メータ物理量の値だけでなく、それと対応した被測定量の値を知ることができる。しかし、そのこと自体は、その相互作用が被測定量の測定を実現していることの原因にはならない。なぜなら、測定によって知りたいのは、測定後の被測定量の値ではなく、測定直前の被測定量の値だからである。

それでは、伝統的な理解は全く的外れだったのかと言えば、そうではない。von Neumann によって定式化された伝統的な測定理論では、測定の前で被測定量の値が変化しない測定だけが扱われているからである。そのような場合には、測定後のメータの値、測定後の被測定量の値、測定前の被測定量の値の3者の間に一対一対応が成立し、結局、測定後のメータの値から測定前の被測定量の値を知ることができる。つまり、伝統的な理論では、測定前の被測定量の値と測定後の被測定量の値が一致していることが暗黙の前提となり、そのもとで、測定後の状態における測定対象と測定装置の間の相関を考えれば十分だと言う論理構造になっている。

しかし、あらゆる可能な測定を考える現代の進歩した量子測定理論では、そのような議論はもはや通用しない。現代の一般的な測定理論では、測定前の被測定量の値と測定後の被測定量の値が一致するという前提は成立せず、もはや、測定後の状態における測定対象と測定装置の間の相

関を考えれば十分だと言うことにはならない。それにも関わらず、量子測定は、測定対象と測定装置の間に相関を確立することであるという言い方は正しいであろう。なぜなら、測定後のメータの値から測定前の被測定量の値が知られるためには、測定相互作用によってそれらの間に何らかの相関が構成される必要があるからである。つまり、ある物理量の測定において、観測者が実際に読み取るのが、メータ物理量の値であったとしても、観測者が報告するのは、あくまでも被測定量の値が何々であったということであり、そのことは、測定後のメータの値と測定前の被測定量の値に相関がなければ、正当化されないと考えられる。

しかし、そのような相関を一般的に扱う理論は、伝統的な量子力学の中には存在しない。それでは、一般的な測定に対して、量子測定は、測定対象と測定装置の間に相関を確立することであるという命題は、厳密に述べればどういう意味を有するのであろうか。

本稿では、まず、この問題を解明して、ある物理量の値を測定するという言葉の意味を明らかにする。このことから直ちに導かれるのは、同時測定に関する新しい理論である。従来、同時測定の可能性と可換性は同等な概念と考えられてきたが、それは、両者を状態に依存しない形で定式化した場合である。つまり、二つの物理量が任意の状態でも同時測定可能であるということとそれらが作用素として可換であることは、同等である。しかし、ある装置がある物理量を正しく測定するかどうかは、一般に、どの状態で測定を行なうかに依存する。例えば、ガンマ線顕微鏡で電子の位置が（少なくともある望ましい誤差で）測定できるのは、電子の波動関数がガンマ線顕微鏡の視野に局在しているときに限り、視野の外にある電子の位置が測定できるわけではない。従って、同時測定可能性に限らず、測定可能性一般について、理論は状態依存的でなければならない。物理量の相関という概念は、それらの相関が状態から定まると言う意味で、まさしく、状態依存的であり、物理量の測定の概念が状態依存的に確立されるならば、それによって、状態依存的な同時測定の理論の構築が自然になされる。このことから明らかになるのは、本来的に同時に値を付与できないはずの至る所非可換な物理量でさえも同時測定が可能な場合があるということである。このことは、従来からも指摘されてきたことであるが、理論的に一般的な定式化のもとでその可能性が明らかにされる。

すると、問題になるのは、その場合に物理量の値と測定値の関係をどう考えるべきかという問題である。そのために、測定が達成されると言うことが量子力学的文脈性にどのように関わるかを明らかにする必要がある。ここで、量子力学的文脈性とは、Kochen-Specker の定理 [1] などで明らかにされた物理量の値という概念の使用の制約に関わる概念である。これまで、測定前の被測定量の値という言葉で、暗黙のうちに、被測定量の値という概念を取り扱ってきたが、物理量の値という概念は、一般に文脈依存的である。つまり、すべての物理量に対してそれらの値を付与することはできないと考えられている。本稿では、物理量の集まりに対する一つの付値が有効な範囲、または、その範囲の物理量に関する観測命題の集まりを文脈と呼んでいる。つまり、いくつかの物理量の値に言及する場合には、その言及がどの文脈においてなされているのかを明らかにする必要がある。従って、「物理量の測定とは、測定前の被測定量の値と測定後のメータの値の一致である」と述べるならば、それらの値への言及を正当化する文脈を明らかにする必要がある。

これらの問題に関する量子論理的アプローチを明らかにして、一つの解釈を導く。本稿は、これらの論考によって、物理量の測定とは、測定前の被測定量の値と測定後のメータの値の一致であるという見方の正当性と有効性を明らかにすることを目的としている。

2 von Neumann の測定過程

本稿では、有限次元 Hilbert 空間 \mathcal{H} で記述される量子系 \mathbf{S} を考察する。 \mathcal{H} 上の自己共役作用素を **物理量** と呼び、 \mathcal{H} の単位ベクトルを **状態** と呼ぶ。 von Neumann [2] は、 Hilbert 空間 \mathcal{H} の縮退のない物理量 $A = \sum_n a_n |\phi_n\rangle\langle\phi_n|$ の測定のプロセスを考察して、測定は次のような対象と装置の合成系のユニタリ変換で表されることを示した。

$$U : |\phi_n\rangle \otimes |\xi\rangle \mapsto |\phi_n\rangle \otimes |\xi_n\rangle. \quad (1)$$

ここで、 $|\xi\rangle$ は対象のある状態ベクトル、 $\{|\xi_n\rangle\}$ は装置の Hilbert 空間 \mathcal{K} のある正規直交基底であり、この場合、 $M = \sum_n a_n |\xi_n\rangle\langle\xi_n|$ をメータ物理量と呼ぶ。

対象の初期状態が $\psi = \sum_n c_n |\phi_n\rangle$ だとすると、 U の線形性から

$$U(|\psi\rangle \otimes |\xi\rangle) = \sum_n c_n |\phi_n\rangle \otimes |\xi_n\rangle \quad (2)$$

が得られる。

この変換が測定と見なされる理由について、 Wigner [3] は次のように述べている。

「測定で得られた状態 (2) において、対象の状態と装置の状態の間に統計的な相関が存在する。つまり、対象プラス装置の合成系において、本来の被測定量 A と装置のメータ物理量 M を同時に測定すると、常に同じ測定値が得られる。その結果、それらのうち一方の測定は不要である。つまり、対象の状態は装置の測定によって確認される。このことは、状態ベクトル (2) が、 $m \neq n$ に対して $|\phi_m\rangle \otimes |\xi_n\rangle$ のような形の項を一切含まないと言う特別な形をしている結果である。」

上の説明は、状態 (2) において、被測定量 A とメータ物理量 M の間に統計的相関が存在することを正しく指摘している。しかし、このことが上の変換が正しい測定を与える理由なのであろうか？上で指摘されている相関は、測定相互作用後の被測定量 A とメータ物理量 M の間の相関であるから、確かに、測定相互作用後のメータ物理量 M を見れば、測定相互作用後の被測定量 A の値がわかるという風になっている。しかし、本来、測定で知りたいのは、測定相互作用前の被測定量 A の値であるから、上の説明は、十分な説明とは言えないであろう。例えば、測定相互作用前の A の確率分布と測定相互作用後の M の確率分布が一致することは、正しい測定の重要な必要条件であるが、上の説明から導かれるのは、ともに測定相互作用後の A と M の確率分布が一致することであって、このことが導かれるわけではない。

3 確率分布再現性

任意の入力状態に対して、測定相互作用前の A の確率分布と測定相互作用後の M の確率分布が一致するという条件を、**確率分布再現性**と呼ぶことにしよう。

1970年代に、Helstrom [4] によって確率作用素値測度 (POVM) の概念が導入され、すべての測定装置は固有の POVM を持つことが知られるようになった。離散的な POVM は、 $\sum_x \Pi(x) = 1$ を満たす正作用素の族 $\{\Pi(x)\}$ で表すことができ、状態 $|\psi\rangle$ で POVM $\{\Pi(x)\}$ を持つ測定装置で測定を行なうと、測定値の確率分布は、

$$\Pr\{\mathbf{x} = x \mid |\psi\rangle\} = \langle \psi | \Pi(x) | \psi \rangle \quad (3)$$

で与えられる。従って、確率分布再現性が成立することは、測定装置の POVM が

$$\Pi(a_n) = |\phi_n\rangle\langle\phi_n| \quad (4)$$

を満たすことと同等である。

そこで、再び von Neumann の測定過程に戻ると、von Neumann の測定過程では、(1) 式が要請されたが、確率分布再現性に必要な条件は、実は、それより弱い条件

$$U : |\phi_n\rangle \otimes |\xi\rangle \mapsto |\phi'_n\rangle \otimes |\xi_n\rangle \quad (5)$$

で表されることがわかる。ここで、 $\{|\phi'_n\rangle\}$ は任意の状態の族であって、 $|\phi'_n\rangle = |\phi_n\rangle$ が成り立つ必要はない。実際、対象の初期状態が $|\psi\rangle = \sum_n c_n |\phi_n\rangle$ だとすると、 U の線形性から

$$U(|\psi\rangle \otimes |\xi\rangle) = \sum_n c_n |\phi'_n\rangle \otimes |\xi_n\rangle \quad (6)$$

が得られるが、この状態で観測者がメータ物理量 M を測定すれば、測定値 \mathbf{x} の確率分布は

$$\begin{aligned} \Pr\{\mathbf{x} = a_n \mid |\psi\rangle\} &= \|I \otimes E^M(a_n) U(|\psi\rangle \otimes |\xi\rangle)\|^2 \\ &= \|c_n |\phi'_n\rangle \otimes |\xi_n\rangle\|^2 = |c_n|^2 \end{aligned}$$

となって、確率分布再現性が示される。ここで、一般に $E^X(x)$ は部分空間 $\{|\psi\rangle \mid X|\psi\rangle = x|\psi\rangle\}$ への射影を表す。この部分空間は、 a が固有値ならば、 a に対する A の固有空間であり、そうでないならば、0次元部分空間 $\{0\}$ である。

さて、 $\{|\phi'_n\rangle\}$ は任意の状態の族でよかったのであるから、特に、ある状態 $|\phi\rangle$ が存在して、 $|\phi'_n\rangle = |\phi\rangle$ となる場合は、

$$U(|\psi\rangle \otimes |\xi\rangle) = |\phi\rangle \otimes \sum_n c_n |\xi_n\rangle \quad (7)$$

となって、測定直後の合成系の状態において、Wigner が主張したような対象と装置の間の相関は存在しない。

4 測定と相関

それでは、「量子測定は、測定対象と測定装置の間に相関を確立することである」というのは、正しくないのだろうか。「相関」を「エンタングルメント」と言い替えて、「量子測定は、測定対象と

測定装置の間にエンタングルメントを確立することである」という主張が正しくないことは、これまでの議論で明らかである。つまり、(7)式の場合には、測定相互作用はエンタングルメントを構成しない。しかし、確率分布再現性に必要な条件である(5)式を見ると、確かに、 $|\phi_n\rangle$ と $|\xi_n\rangle$ の間の一対一対応が存在し、測定は、対象と装置の間に「異なるタイプの量子相関」を確立していると考えられる。ここで、異なるタイプというのは、次のような意味がある。つまり、エンタングルメントは、スペースライクな事象の間の相関であり、Wignerが指摘したのは、von Neumann型の測定には測定直後の時刻に対象と装置という系の間にエンタングルメントと呼ばれるスペースライクな相関があるということであった。しかし、(5)式で示唆されていることは、測定直前の対象の状態 $|\phi_n\rangle$ と測定直後の装置の状態 $|\xi_n\rangle$ が示すタイムライクな相関が存在するということである。つまり、Wignerが指摘した相関は、測定直後のメータの値を読めば、測定直後の被測定量の値が知られると言うタイプの相関であったのに対して、測定が本来、実現すべき相関とは、測定直後のメータの値を読めば、測定直前の被測定量の値が知られるというタイプの相関である。問題は、対象の初期状態が任意の状態のときにもそのような相関が存在するかどうかということであり、とりわけ、(7)式の様なケースでも確かに、測定は対象と装置の間に相関を確立しているかどうかを明らかにすることである。もちろん、測定が正しいならば、測定直後のメータの値を読めば、測定直前の被測定量の値が知られるというタイプの相関は、成り立っている筈である。

ところが、測定直前の被測定量と測定直後のメータ物理量が、一方の値を見れば、他方の値がわかるという完全相関を持つかどうかを議論するのは、難しい問題を含んでいる。と言うのは、測定直前の被測定量と測定直後のメータ物理量は、一般に交換可能ではないからである。

交換可能な2つの物理量 X, Y は、任意のベクトル状態 $|\psi\rangle$ において結合確率分布

$$\Pr\{X = x, Y = y \mid |\psi\rangle\} = \langle \psi | E^X(x) E^Y(y) | \psi \rangle \quad (8)$$

を持ち、それらが完全相関して確率1で値が同一であるとは、 $\Pr\{X = Y \mid |\psi\rangle\} = 1$ が成り立つことである。ただし、事象 $X = Y$ の確率は、

$$\Pr\{X = Y \mid |\psi\rangle\} = \sum_x \Pr\{X = x, Y = x \mid |\psi\rangle\} \quad (9)$$

で与えられる。

ここで、測定の相互作用が時刻0から時刻 Δt まで行なわれたとすれば、対象と装置の合成系のユニタリ変換 U は合成系の時刻0から時刻 Δt までの時間発展を表す。そこで、 $A(0) = A \otimes I$, $A(\Delta t) = U^\dagger A(0) U$, $M(0) = I \otimes M$, $M(\Delta t) = U^\dagger M(0) U$ 等と表わすことにする。Wignerが指摘したことを言い替えると、2つの物理量 A, B が状態 $U(|\psi\rangle \otimes |\xi\rangle)$ で完全相関して確率1で値が同一であるということになる。これをHeisenberg描像で述べれば、合成系の初期状態 $|\psi\rangle \otimes |\xi\rangle$ において、 $A(\Delta t)$ と $M(\Delta t)$ が完全に相関して確率1で値が同一であるということである。しかし、測定が本来、実現すべき相関とは、 $A(0)$ と $M(\Delta t)$ が完全に相関して確率1で値が同一であるということであろう。ここで、 $A(0)$ と $M(\Delta t)$ は可換とは限らないから、上の議論を当てはめるわけにはいかない。このように、測定直前の被測定量と測定直後のメータ物理量が、一方の値を見れば

ば、他方の値がわかるという完全相関を持つかどうかを議論するのは、難しい問題を含んでいる。測定直前の被測定量と測定直後のメータ物理量は、一般に交換可能ではないからである。

以下では、必ずしも可換とは限らない物理量 X, Y について、それらが完全に相関して確率1で値が同一であるということを経学的にどのように定式化すべきかという問題を考察しよう。

5 状態依存的可換性

従来の定式化では、2つの物理量が同時測定可能であることとそれらが作用素として可換であることが同等であるとされてきた。しかし、この命題は、状態に依存しない定式化に基づいている。同時測定可能性を状態依存的に取り扱えば、ある特殊な状態においては、非可換な物理量が同時測定可能であるといういくつかの主張が知られている。本稿では、状態に依存する同時測定可能性について考察し、ある与えられた状態で二つの物理量が同時測定可能であるための条件を決定する。そのためには、まず、物理量の状態に依存する可換性の概念を確立することから始めよう。

二つの物理量 A と B が状態 $|\psi\rangle$ で可換である ($A \leftrightarrow_{|\psi\rangle} B$) とは、任意の $a, b \in \mathbf{R}$ に対して、

$$[E^A(a), E^B(b)]|\psi\rangle = 0 \quad (10)$$

が成り立つことをいう。これは、 A, B で生成される C^* 代数の状態 $|\psi\rangle$ で生成される巡回表現が可換であることと同値である。もちろん、 A, B が作用素の積に関して可換であることと、任意の状態 $|\psi\rangle$ で可換であることは同値である。 n 次元確率分布とは、非負実数の族 $\{p(x_1, \dots, x_n) \mid (x_1, \dots, x_n) \in \mathbf{R}^n\}$ で $\sum_{x_1, \dots, x_n} p(x_1, \dots, x_n) = 1$ を満たすものである。ここで、 \mathbf{R} は実数直線を表す。2つの物理量 A, B の状態 $|\psi\rangle$ における結合確率分布 (JPD) とは、 a, b についての2次元確率分布 $\Pr\{A = a, B = b \mid |\psi\rangle\}$ で次の関係

$$\langle \psi | f(A, B) | \psi \rangle = \sum_{a, b} f(a, b) \Pr\{A = a, B = b \mid |\psi\rangle\} \quad (11)$$

が A と B の任意の多項式 $f(A, B)$ に対して成り立つものである。2つの射影作用素 E と F の下限とは、 E と F それぞれの値域の共通部分の上への射影 $E \wedge F$ のことである。従って、射影作用素 $E^A(a) \wedge E^B(b)$ は、部分空間 $\{|\psi\rangle \in \mathcal{H} \mid A|\psi\rangle = a|\psi\rangle \text{ かつ } B|\psi\rangle = b|\psi\rangle\}$ の上への射影作用素である。結合確率分布は、本質的に状態依存的な概念であり、われわれの理論で中心的な役割を果たす。Gudder [5] は、JPD の存在に関する次の特徴付けを得た。

定理 1 (Gudder 1968) 任意の物理量 A と B 及び任意の状態 $|\psi\rangle$ に対して、次の条件は同値である。

- (i) A, B の状態 $|\psi\rangle$ に対する JPD が存在する。
- (ii) A, B が状態 $|\psi\rangle$ で可換である。
- (iii) 状態 $|\psi\rangle$ は A と B の共通の固有状態の重ね合わせである。
- (iv) 実数の族 $p(a, b) = \langle \psi | E^A(a) \wedge E^B(b) | \psi \rangle$ は、 a, b に対する2次元確率分布である。
- (v) $\sum_{a, b} \langle \psi | E^A(a) \wedge E^B(b) | \psi \rangle = 1$ が成立する。

物理量 A と B が、どの状態においても可換でないとき、それらは、**至る所で非可換**であると言う。これは、それらが共通の固有状態を持たないことと同値である。異なる方向のスピン成分は、至る所で非可換である。

6 結合擬確率分布

結合確率分布が常に存在するとは限らないので、物理量の任意の対に対して存在し、可換な物理量の対に対しては、結合確率分布に帰着する数学的概念があると便利である。ここでは、結合確率の、弱と強の、2つの拡張概念を取り上げる [6]。

物理量 A と B の状態 $|\psi\rangle$ における**強結合擬確率分布**は、任意の $a, b \in \mathbf{R}$ に対して、

$$\Pr_S\{A = a, B = b \mid |\psi\rangle\} = \langle \psi | E^A(a) \wedge E^B(b) | \psi \rangle \quad (12)$$

によって定義される。物理量 A と B の状態 $|\psi\rangle$ における**弱結合擬確率分布**は、任意の $a, b \in \mathbf{R}$ に対して、

$$\Pr_W\{A = a, B = b \mid |\psi\rangle\} = \langle \psi | E^B(a) E^A(b) | \psi \rangle \quad (13)$$

によって定義される。強結合擬確率分布の概念は、Birkhoff と von Neumann [7] の仕事に由来する。彼らは、“ $A = a$ かつ $B = b$ ” という命題に射影作用素 $E^A(a) \wedge E^B(b)$ を対応させた。弱結合擬確率分布は Kirkwood [8] によって導入された。上の2つの概念は、 A と B が状態 $|\psi\rangle$ で交換可能ならば、結合確率分布に帰着する。

定理 2 任意の物理量 A, B 及び任意の状態 $|\psi\rangle$ に対して、物理量 A, B が $|\psi\rangle$ で可換であるための必要十分条件は、強及び弱結合擬確率分布が一致すること、すなわち、

$$\Pr_S\{A = a, B = b \mid |\psi\rangle\} = \Pr_W\{A = a, B = b \mid |\psi\rangle\} \quad (14)$$

が任意の $a, b \in \mathbf{R}$ に対して成立することである。

定理 1 (iv) より、物理量 A, B の $|\psi\rangle$ における強擬確率分布が2次元確率分布であるための必要十分条件は、 A と B が $|\psi\rangle$ で可換であることである。しかしながら、12節で議論するように、 A, B の $|\psi\rangle$ における弱結合擬確率分布は、 A 及び B が $|\psi\rangle$ で可換でないときにも2次元確率分布になることがある。

7 条件付き擬確率分布

確率論では、ある確率変数 A の確率変数 B に関する条件付き確率分布は、変数 A と変数 B の結合確率分布の変数 B の周辺分布に対する比として定義され、条件付き期待値は、条件付き確率分布に対する期待値として定義される。それらからの類推として、量子物理量の条件付き確率分布と条件

付き期待値が次のように定義される。もし、物理量 A, B が状態 $|\psi\rangle$ で可換で、 $\Pr\{B = b \mid |\psi\rangle\} > 0$ ならば、 $B = b$ の下での A の**条件付き確率分布**が

$$\Pr\{A = a \mid B = b \mid |\psi\rangle\} = \frac{\Pr\{A = a, B = b \mid |\psi\rangle\}}{\Pr\{B = b \mid |\psi\rangle\}} \quad (15)$$

によって自然に定義される。このとき、 $B = b$ の下での A の**条件付き期待値**が

$$\text{Ex}\{A \mid B = b \mid |\psi\rangle\} = \sum_a a \Pr\{A = a \mid B = b \mid |\psi\rangle\} \quad (16)$$

で定義される。さて、 $\Pr\{B = b \mid |\psi\rangle\} > 0$ であれば、 $A \leftrightarrow_{|\psi\rangle} B$ を仮定することなく、強、及び、弱条件付き擬確率分布が次のように定義される。 $B = b$ の下での A の**強条件付き擬確率分布**が

$$\Pr_S\{A = a \mid B = b \mid |\psi\rangle\} = \frac{\Pr_S\{A = a, B = b \mid |\psi\rangle\}}{\Pr\{B = b \mid |\psi\rangle\}} \quad (17)$$

によって定義され、**強条件付き擬期待値**が

$$\text{Ex}_S\{A \mid B = b \mid |\psi\rangle\} = \sum_a a \Pr_S\{A = a \mid B = b \mid |\psi\rangle\} \quad (18)$$

によって定義される。一方、 A, B の弱結合擬確率分布は、 A と B に関して対象ではないので、事前選択と事後選択に対応する2つの異なる条件付き擬確率分布が存在する。それらは、一方から他方へは容易に類推されるので、以下では、事後選択に対応するものだけを導入する。事後選択 $B = b$ の下での A の**弱条件付き擬確率分布**が

$$\Pr_W\{A = a \mid B = b \mid |\psi\rangle\} = \frac{\Pr_W\{A = a, B = b \mid |\psi\rangle\}}{\Pr\{B = b \mid |\psi\rangle\}} \quad (19)$$

によって定義される。また、事後選択 $B = b$ の下での A の**弱条件付き擬期待値**が

$$\text{Ex}_W\{A \mid B = b \mid |\psi\rangle\} = \sum_a a \Pr_W\{A = a \mid B = b \mid |\psi\rangle\} \quad (20)$$

によって定義される。明らかに、

$$\text{Ex}_W\{A \mid B = b \mid |\psi\rangle\} = \frac{\langle \psi \mid E^B(b) A \mid \psi \rangle}{\langle \psi \mid E^B(b) \mid \psi \rangle} \quad (21)$$

であり、 $A \leftrightarrow_{|\psi\rangle} B$ ならば、定理 12 から

$$\text{Ex}\{A \mid B = b \mid |\psi\rangle\} = \text{Ex}_S\{A \mid B = b \mid |\psi\rangle\} = \text{Ex}_W\{A \mid B = b \mid |\psi\rangle\} \quad (22)$$

が成り立つ。

8 弱値

Aharonov-Albert-Vaidman [9] によって導入された弱値の概念は、以下のように数学的に定式化することができる。物理量 A と2つの状態 $|\psi_i\rangle, |\psi_f\rangle$ について、事前選択状態 $|\psi_i\rangle$ 及び事後選択状態 $|\psi_f\rangle$ に関する A の弱値が

$$\langle \psi_f \mid A \mid \psi_i \rangle_w = \frac{\langle \psi_f \mid A \mid \psi_i \rangle}{\langle \psi_f \mid \psi_i \rangle} \quad (23)$$

によって定義される. Steinberg [10, 11] によって指摘されたように, 弱値の形式的定義は, 確率論的概念であるところの事後選択に関する弱条件付き擬期待値に一致する. 後者は, Kirkwood [8] の結合擬確率分布から自然に定義された.

定理 3 (Steinberg 1995) 物理量 $B = \sum_b b|B = b\rangle\langle B = b|$ に対して, 事前選択状態 $|\psi_i\rangle = |\psi\rangle$ と事後選択状態 $|\psi_f\rangle = |B = b\rangle$ に関する物理量 A の弱値は, 事後選択 $B = b$ の下での A の弱条件付き擬期待値に一致する. 即ち,

$$\langle B = b | A | \psi \rangle_w = \text{Ex}_W \{ A | B = b || \psi \} \quad (24)$$

が成り立つ.

9 状態依存的等値性

2つの物理量が同一の値を持つのはどういう場合かという問は, 量子力学において基本的な問題であるが, 長いこと満足のいく解決が得られていなかった. 筆者の最近の仕事 [12, 13] において, この問題の有力な解決が弱結合擬確率分布に基づいて与えられた.

2つの物理量 A, B が状態 $|\psi\rangle$ において, **量子同一性相関** または **量子完全相関** を持つ ($A =_{|\psi\rangle} B$) とは, $a \neq b$ ならば

$$\text{Pr}_W \{ A = a, B = b || \psi \} = 0 \quad (25)$$

が成り立つことを言う. 物理量 A, B と状態 $|\psi\rangle$ に対して, $A, B, |\psi\rangle$ で生成される巡回部分空間 $C(A, B, |\psi\rangle)$ を

$$C(A, B, |\psi\rangle) = \{ f(A, B) |\psi\rangle \mid f : \text{多項式} \}. \quad (26)$$

で定義する. この時, $A, |\psi\rangle$ で生成される巡回部分空間 $C(A, |\psi\rangle)$ は, $C(A, |\psi\rangle) = C(A, I, |\psi\rangle)$ によって定義される. 次の定理は, 二つの物理量の量子同一性相関を特徴付ける [12, 13].

定理 4 任意の物理量 A, B 及び状態 $|\psi\rangle$ について, 次の条件は同値である.

- (i) $A =_{|\psi\rangle} B$.
- (ii) $A \leftrightarrow_{|\psi\rangle} B$ かつ $\sum_a \text{Pr}\{A = a, B = a || \psi\} = 1$.
- (iii) $|\psi\rangle$ は, A と B の共通の固有値に対する共通の固有状態の重ね合わせである.
- (iv) $A = B$ が $C(A, |\psi\rangle)$ 上で成り立つ.
- (v) 任意の x に対して $E^A(x) = E^B(x)$ が $C(A, |\psi\rangle)$ 上で成り立つ.
- (vi) $\text{Pr}\{A = x || \phi\} = \text{Pr}\{B = x || \phi\}$ が任意の x と任意の $|\phi\rangle \in C(A, |\psi\rangle)$ について成り立つ.
- (vii) $\|A|\phi\rangle - B|\phi\rangle\| = 0$ が任意の $|\phi\rangle \in C(A, |\psi\rangle)$ について成り立つ
- (viii) $\sum_a \text{Pr}_S\{A = a, B = a || \psi\} = 1$.
- (ix) $\langle B = b | E^A(a) | \psi \rangle_w = \delta_{a,b}$ が $\text{Pr}\{B = b || \psi\} > 0$ ならば成り立つ.
- (x) $\text{Pr}_S\{A = a, B = b || \psi\} = \text{Pr}_W\{A = a, B = b || \psi\} = \delta_{a,b} \text{Pr}\{A = a || \psi\}$.

量子同一性相関の重要な性質は推移律である [12, 13]. 可換性はこの性質を持たない.

定理 5 関係 $=_{|\psi\rangle}$ は物理量の間と同値関係である。とりわけ、もし $A =_{|\psi\rangle} B$ かつ $B =_{|\psi\rangle} C$ ならば、 $A =_{|\psi\rangle} C$ が成り立つ。

10 測定過程

\mathcal{H} に対する**測定過程**とは、4つ組 $(\mathcal{K}, |\xi\rangle, U, M)$ のことで、Hilbert空間 \mathcal{K} , \mathcal{K} の状態 $|\xi\rangle$, $\mathcal{H} \otimes \mathcal{K}$ 上のユニタリ作用素 U , 及び、 \mathcal{K} 上の自己共役作用素 M からなる [14]. 測定過程 $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M)$ は、 \mathcal{H} で記述される被測定系 \mathbf{S} と \mathcal{K} で記述される**プローブ**と呼ばれる系 \mathbf{P} の間の**測定相互作用**と呼ばれる相互作用によって行なわれる測定を記述する。ここで、 \mathbf{P} は測定の直前に状態 $|\xi\rangle$ に準備される。ユニタリ作用素 U は、測定相互作用の間（時刻 0 から時刻 Δt までとする）の合成系 $\mathbf{S} + \mathbf{P}$ の時間発展を表す。 \mathcal{H} 上の物理量 A, B と \mathcal{K} 上の物理量 M に対して、Heisenberg 描像で $A(0) = A \otimes I$, $B(0) = B \otimes I$, $M(0) = I \otimes M$, $A(\Delta t) = U^\dagger(A \otimes I)U$, $B(\Delta t) = U^\dagger(B \otimes I)U$, $M(\Delta t) = U^\dagger(I \otimes M)U$ と表す。 n 次元 POVM とは、 $(x_1, \dots, x_n) \in \mathbf{R}^n$ に対する \mathcal{H} 上の正作用素 $\Pi(x_1, \dots, x_n)$ の族で $\sum_{x_1, \dots, x_n} \Pi(x_1, \dots, x_n) = I$ を満たすものである。測定過程 \mathbf{M} に対して、

$$\Pi(x) = \text{Tr}_{\mathcal{K}}[E^{M(\Delta t)}(x)(I \otimes |\xi\rangle\langle\xi|)] \quad (27)$$

によって1次元 POVM が定義され、 \mathbf{M} の POVM と呼ばれる。測定過程 \mathbf{M} の測定値は、**メータ物理量**と呼ばれる物理量 M の測定相互作用直後の値を測定することによって得られる。従って、任意の入力状態 $|\psi\rangle$ における**出力確率分布** $\Pr\{\mathbf{x} = x \mid |\psi\rangle\}$, 即ち、この測定の出力変数 \mathbf{x} の確率分布が測定過程 \mathbf{M} の POVM によって

$$\Pr\{\mathbf{x} = x \mid |\psi\rangle\} = \Pr\{M(\Delta t) = x \mid |\psi\rangle|\xi\rangle\} = \langle\psi|\Pi(x)|\psi\rangle \quad (28)$$

と与えられる。

\mathcal{H} 上の物理量 A と状態 $|\psi\rangle$ に対して、測定過程 $(\mathcal{K}, |\xi\rangle, U, M)$ が $|\psi\rangle$ で A を**測定**するとは、 $A(0)$ と $M(\Delta t)$ が量子同一性相関を持つこと、即ち、

$$M(\Delta t) =_{|\psi\rangle|\xi\rangle} A(0) \quad (29)$$

が成り立つことを言う [12, 13]. 次の定理から、このことは、測定の POM だけから定まる事がわかる [12, 13].

定理 6 測定過程 $(\mathcal{K}, |\xi\rangle, U, M)$ が POVM $\{\Pi(x)\}$ を持つとき、次の条件は互いに同値である。

- (i) $(\mathcal{K}, |\xi\rangle, U, M)$ が $|\psi\rangle$ で A を測定する。
- (ii) $a \neq b$ ならば、 $\langle\psi|\Pi(a)E^A(b)|\psi\rangle = 0$ が成り立つ。
- (iii) $\Pi(a)|\phi\rangle = E^A(a)|\phi\rangle$ が任意の $|\phi\rangle \in \mathcal{C}(A, \psi)$ について成り立つ。

この定理から、特に、測定の POVM が被測定量の単位の分解 E^A に等しい、つまり、測定装置が確率分布再現性を持つならば、任意の初期状態 $|\psi\rangle$ において物理量 A が測定されることが示さ

れる。これによって、POVMが E^A に等しいならば、確率分布だけでなく、被測定量の値も再現されることが導かれる。ところが、物理量の値という概念には、量子力学的文脈性の問題が関わってくる。この問題については、後の節で議論する予定である。

ここで、von Neumann型の測定において、 $A(0)$ と $M(\Delta t)$ が量子同一性相関を持つかを調べることにする。測定の初期状態は、 $|\psi\rangle \otimes |\xi\rangle = \sum_n c_n |\phi_n\rangle \otimes |\xi\rangle$ であり、

$$A(0)|\phi_n\rangle \otimes |\xi\rangle = A|\phi_n\rangle \otimes |\xi\rangle = a_n|\phi_n\rangle \otimes |\xi\rangle$$

であるから、 $|\psi\rangle \otimes |\xi\rangle$ は $A(0)$ の固有値 a_n に属する固有状態 $|\phi_n\rangle \otimes |\xi\rangle$ の重ね合わせである。ここで、合成系のユニタリ変換 U が(5)式を満たすと仮定すると、

$$\begin{aligned} M(\Delta t)|\phi_n\rangle \otimes |\xi\rangle &= U^\dagger(1 \otimes M)U(|\phi_n\rangle \otimes |\xi\rangle) \\ &= U^\dagger(1 \otimes M)(|\phi'_n\rangle \otimes |\xi_n\rangle) \\ &= a_n U^\dagger(|\phi'_n\rangle \otimes |\xi_n\rangle) \\ &= a_n |\phi_n\rangle \otimes |\xi\rangle \end{aligned}$$

であるから、 $|\phi_n\rangle \otimes |\xi\rangle$ は $M(\Delta t)$ の固有値 a_n に属する固有状態でもあり、定理4(iii)から $A(0)$ と $M(\Delta t)$ が量子同一性相関を持つことが結論される。ここで、われわれは条件(1)より明らかに弱い条件(5)しか仮定しなかった。このことは、von Neumannの測定過程だけではなく、(7)式の様まったくエンタングルメントを構成しない測定においても $A(0)$ と $M(\Delta t)$ が量子同一性相関を持つことが結論される。

ここで、Heisenberg描像で $A(0)$ および $M(\Delta t)$ は

$$A(0) = A \otimes |\xi\rangle\langle\xi| + A \otimes (I - |\xi\rangle\langle\xi|), \quad (30)$$

$$M(\Delta t) = A \otimes |\xi\rangle\langle\xi| + \sum_{n \neq m} a_m U^\dagger |\phi_n \otimes \xi_m\rangle\langle\phi_n \otimes \xi_m| U \quad (31)$$

と表される。 $\mathcal{C}(A(0), |\psi\rangle \otimes |\xi\rangle) \subseteq \mathcal{H} \otimes \{\xi\}^{\perp\perp}$ という関係が成り立つので、 $\mathcal{C}(A(0), |\psi\rangle \otimes |\xi\rangle)$ 上で $A(0) = M(\Delta t)$ が成り立ち、定理4(iv)から同じ結論を導くことができる。

11 普遍的な不確定性原理

2 変数測定過程とは、測定過程 $(\mathcal{K}, |\xi\rangle, U, M)$ と実関数の組 f, g からなる6つ組 $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M, f, g)$ のことを言う。これは、測定過程 $(\mathcal{K}, |\xi\rangle, U, M)$ の測定値が \mathbf{u} であるとき、 $(\mathbf{x}, \mathbf{y}) = (f(\mathbf{u}), g(\mathbf{u}))$ を測定値とする測定過程を表す。測定過程 $(\mathcal{K}, |\xi\rangle, U, M)$ のPOVMを Π とすると、2変数測定過程 $\mathbf{M} = (\mathcal{K}, |\xi\rangle, U, M, f, g)$ のPOVM $\Pi(x, y)$ は、

$$\Pi(x, y) = \sum_{u: f(x)=u, g(y)=u} \Pi(u) \quad (32)$$

によって定義される。2変数測定過程 $(\mathcal{K}, |\xi\rangle, U, M, f, g)$ の結合出力確率分布は、出力変数 \mathbf{x}, \mathbf{y} の結合確率分布であり、

$$\Pr\{\mathbf{x} = x, \mathbf{y} = y \mid |\psi\rangle\} = \Pr\{f(M(\Delta t)) = x, g(M(\Delta t)) = y \mid |\psi\rangle|\xi\rangle\}. \quad (33)$$

で定義される。このとき、

$$\Pr\{\mathbf{x} = x, \mathbf{y} = y \mid |\psi\rangle\} = \langle \psi | \Pi(x, y) | \psi \rangle \quad (34)$$

が成り立つ。

物理量 A, B と状態 $|\psi\rangle$ に対して、状態 $|\psi\rangle$ における A 測定と B 測定に関する M の平方根平均 2 乗 (rms) 誤差 $\epsilon(A)$ 及び $\epsilon(B)$ が、

$$\epsilon(A) = \|N(A)|\psi\rangle|\xi\rangle\|, \quad (35)$$

$$\epsilon(B) = \|N(B)|\psi\rangle|\xi\rangle\|, \quad (36)$$

によって定義される。ここで、誤差作用素 $N(A), N(B)$ は、 $N(A) = f(M(\Delta t)) - A(0)$ 及び $N(B) = g(M(\Delta t)) - B(0)$ で与えられる。この時、これらの誤差は、次の普遍的に真な関係式 [15, 16]

$$\epsilon(A)\epsilon(B) + \epsilon(A)\sigma(B) + \sigma(A)\epsilon(B) \geq \frac{1}{2} |\langle \psi | [A, B] | \psi \rangle| \quad (37)$$

を満たす。ここで、 $\sigma(A), \sigma(B)$ は A, B の $|\psi\rangle$ における標準偏差を表す。もし、平均誤差 $\langle \xi | \langle \psi | N(A) | \psi \rangle | \xi \rangle$ 及び $\langle \xi | \langle \psi | N(B) | \psi \rangle | \xi \rangle$ がプローブの初期状態 $|\xi\rangle$ から独立ならば、次の関係

$$\epsilon(A)\epsilon(B) \geq \frac{1}{2} |\langle \psi | [A, B] | \psi \rangle| \quad (38)$$

が得られる [15, 16]。これは、Heisenberg のガンマ線顕微鏡の思考実験 [17] で示唆された関係である。この関係は明らかに普遍的妥当性を持たない。なぜなら、 $\epsilon(A) = 0$ ならば $\epsilon(B) \sim \infty$ が $\langle \psi | [A, B] | \psi \rangle \neq 0$ を満たす状態で成り立たなければならないが、定義から明らかのように、一般に $\epsilon(B) \leq \|B\| + \|g(M)\| < \infty$ が成り立つので、これは正しくない。しかし、新しい関係 (37) は、新しい制約 $\sigma(A)\epsilon(B) \geq \frac{1}{2} |\langle \psi | [A, B] | \psi \rangle|$ を導く [15, 16]。Hall [18] は、(38) が成立しない場合における事前情報の役割を指摘した。

12 同時測定可能性

物理量 A, B と状態 $|\psi\rangle$ に対して、 A, B の $|\psi\rangle$ における同時測定とは、2変数測定過程 $(\mathcal{K}, |\xi\rangle, U, M, f, g)$ で

$$f(M(\Delta t)) =_{|\psi\rangle|\xi\rangle} A(0), \quad (39)$$

$$g(M(\Delta t)) =_{|\psi\rangle|\xi\rangle} B(0) \quad (40)$$

を満たすものである。この条件は、測定過程 $(\mathcal{K}, |\xi\rangle, U, f(M))$ が $|\psi\rangle$ で A を測定し、測定過程 $(\mathcal{K}, |\xi\rangle, U, g(M))$ が $|\psi\rangle$ で B を測定することと同値である。2つの物理量 A, B が状態 $|\psi\rangle$ で**同時測定可能**であるとは、 A, B の $|\psi\rangle$ における同時測定が存在することを言う。

定理4 (vii) から、 A, B の任意の同時測定は、 $\epsilon(A) = \epsilon(B) = 0$ を満たすので、 $\langle \psi|[A, B]|\psi\rangle = 0$ が成り立つ。更に、 A, B の任意の同時測定は、容易に $E^A(a), E^B(b)$ の同時測定に変形できるので、 $\langle \psi|[E^A(a), E^B(b)]|\psi\rangle = 0$ が任意の a, b に対して成り立つ。しかし、これは A と B が $|\psi\rangle$ で可換であることを意味しない。なぜなら、可換性は $\langle \psi|[E^A(a), E^B(b)]^2|\psi\rangle = 0$ と同値だからである。従って、至る所非可換な A と B の同時測定は、 $\langle \psi|[E^A(a), E^B(b)]|\psi\rangle = 0$ が任意の a, b に対して成り立ち、 $\langle \psi|[E^A(a), E^B(b)]^2|\psi\rangle \neq 0$ がある a, b に対して成り立つ場合に限り可能である。

次の定理が示すように、同時測定可能性と可換性は、状態依存的定式化においては同値な概念ではない [19].

定理 7 (i) 2つの物理量 A, B が状態 $|\psi\rangle$ で可換であるための必要十分条件は、2次元 POVM Π が存在して

$$\sum_y \Pi(x, y) = E^A(x) \quad \text{が} \quad C(A, B, |\psi\rangle) \quad \text{で,} \quad (41)$$

$$\sum_x \Pi(x, y) = E^B(y) \quad \text{が} \quad C(A, B, |\psi\rangle) \quad \text{で} \quad (42)$$

が成り立つことである。

(ii) 2つの物理量 A, B が状態 $|\psi\rangle$ で同時測定可能であるための必要十分条件は、2次元 POVM Π が存在して

$$\sum_y \Pi(x, y) = E^A(x) \quad \text{が} \quad C(A, |\psi\rangle) \quad \text{で,} \quad (43)$$

$$\sum_x \Pi(x, y) = E^B(y) \quad \text{が} \quad C(B, |\psi\rangle) \quad \text{で} \quad (44)$$

が成り立つことである。

(iii) 2つの物理量は、状態 $|\psi\rangle$ で可換ならば、同時測定可能である。

(iv) 2つの物理量が任意の状態 $|\psi\rangle$ で同時測定可能であるための必要十分はそれらが \mathcal{H} 上で可換であることである。

次の定理は、2つの至る所非可換な物理量が同時測定可能な場合があることを示している [19].

定理 8 (i) 任意の Hilbert 空間において、任意の物理量の対は一方の物理量の固有状態において同時測定可能である。

(ii) 4次元以上の任意の Hilbert 空間において、至る所非可換な物理量の対でいずれの物理量の固有状態でもない状態で同時測定可能なものがある。

$\dim(\mathcal{H}) = 2$ の場合は、次の特徴付けが成り立つ [19].

定理 9 $\dim(\mathcal{H}) = 2$ であるとき, 任意の物理量 A, B と状態 $|\psi\rangle$ に対して, 次の条件は同値である.

- (i) A と B は $|\psi\rangle$ で同時測定可能である.
- (ii) $\Pr_W\{A = a, B = b \mid |\psi\rangle\} \geq 0$ が任意の a, b に対して成り立つ.
- (iii) A と B が $|\psi\rangle$ で可換であるか, または, $|\psi\rangle$ は A または B の固有状態である.

上の条件 (i) と (ii) が任意の次元においても同値であるかどうかは, 未解決である.

従来の理論では, 2つの可換な物理量 A, B の状態 ψ における同時測定の結合出力確率分布がそれらの結合確率分布で与えられる. しかし, もし A, B が可換でないならば, この関係はもはや意味を持たない. 次の定理により, 同時測定の結合出力確率分布は, 常に, それらの弱結合擬確率分布として与えられる [6].

定理 10 物理量 A, B の状態 $|\psi\rangle$ における任意の同時測定過程 $(\mathcal{K}, |\xi\rangle, U, M, f, g)$ の結合出力確率分布は, 測定過程には依存せず, 常に, それらの弱結合擬確率分布に一致する. 即ち,

$$\Pr\{\mathbf{x} = x, \mathbf{y} = y \mid |\psi\rangle\} = \Pr_W\{A = x, B = y \mid |\psi\rangle\} = \Pr_W\{B = y, A = x \mid |\psi\rangle\} \quad (45)$$

が成り立つ. また, $\mathbf{y} = y$ の下での $\mathbf{x} = x$ の条件付き確率は, 事前状態 $|\psi\rangle$ と事後選択 $B = y$ に関する $E^A(x)$ の弱値, 及び, 事前選択 $B = y$ と事後状態 $|\psi\rangle$ に関する $E^A(x)$ の弱値に一致する. 即ち,

$$\Pr\{\mathbf{x} = x \mid \mathbf{y} = y \mid |\psi\rangle\} = \langle B = y \mid E^A(x) \mid \psi \rangle_w = \langle \psi \mid E^A(x) \mid B = y \rangle_w \quad (46)$$

が成り立つ.

13 観測命題の論理

$\mathcal{O}(\mathcal{H})$ で \mathcal{H} 上の物理量の集合を表す. \mathcal{H} に関する**観測命題**を以下の規則で定める.

(R1) 任意の $X \in \mathcal{O}(\mathcal{H})$ と $x \in \mathbf{R}$ に対して, 式 $X = x$ は観測命題である. これを**原子観測命題**と呼ぶ.

(R2) ϕ_1 及び ϕ_2 が観測命題ならば, $\neg\phi_1$ 及び $\phi_1 \wedge \phi_2$ も観測命題である.

従って, すべての観測命題は, 原子観測命題から有限個の結合子 \neg 及び \wedge によって構成されたものである. 構成を明らかにするために, 括弧は自由に使用することにする. 例えば, $X_1, X_2, X_3 \in \mathcal{O}(\mathcal{H})$ かつ $x_1, x_2, x_3 \in \mathbf{R}$ ならば, $(X_1 = x_1 \wedge X_2 = x_2) \vee X_3 = x_3$ は観測命題である. 結合子 \vee を次の定義で導入する.

(D1) $\phi_1 \vee \phi_2 := \neg(\neg\phi_1 \wedge \neg\phi_2)$.

観測命題 $(\phi_1 \wedge \phi_2) \wedge \phi_3$ を $\phi_1 \wedge \phi_2 \wedge \phi_3$ と表す. 同様の既約をより長い観測命題や結合子 \vee についても仮定する.

\mathcal{H} 上の射影作用素からなる束を \mathcal{H} 上の**量子論理**と呼び, $\mathcal{Q}(\mathcal{H})$ で表す.

任意の観測命題 ϕ に対して, $\mathcal{Q}(\mathcal{H})$ 値の真理値 $[[\phi]] \in \mathcal{Q}(\mathcal{H})$ を以下の規則で定める.

$$(T1) \llbracket X = x \rrbracket = E^X(x).$$

$$(T2) \llbracket \neg \phi \rrbracket = \llbracket \phi \rrbracket^\perp.$$

$$(T3) \llbracket \phi_1 \wedge \phi_2 \rrbracket = \llbracket \phi_1 \rrbracket \wedge \llbracket \phi_2 \rrbracket.$$

(D1), (T2), (T3) から次を得る.

$$(T4) \llbracket \phi_1 \vee \phi_2 \rrbracket = \llbracket \phi_1 \rrbracket \vee \llbracket \phi_2 \rrbracket.$$

$\llbracket \phi \rrbracket |\psi\rangle = |\psi\rangle$ が成立するとき、観測命題 ϕ は状態 $|\psi\rangle$ で**真である**、または、**成立する**と言う。

14 文脈的整合性

観測命題 ϕ_1, \dots, ϕ_n に対して、観測命題 $\bigvee_j \phi_j$ を $\bigvee_j \phi_j = \phi_1 \vee \dots \vee \phi_n$ と定義する。 $\text{Sp}(X)$ で X の固有値の集合を表す。

任意の物理量 X_1, \dots, X_n に対して、観測命題 $\text{com}(X_1, \dots, X_n)$ を

$$\text{com}(X_1, \dots, X_n) := \bigvee_{x_1 \in \text{Sp}(X_1), \dots, x_n \in \text{Sp}(X_n)} X_1 = x_1 \wedge \dots \wedge X_n = x_n \quad (47)$$

によって定義する。物理量 X_1, \dots, X_n が状態 $|\psi\rangle$ で**同時決定的**であるとは、観測命題 $\text{com}(X_1, \dots, X_n)$ が $|\psi\rangle$ で成立することである。この時、次の定理が成り立つ。

定理 11 物理量 X_1, \dots, X_n が状態 $|\psi\rangle$ で同時決定的であるための必要十分条件は、 $|\psi\rangle$ が X_1, \dots, X_n の共通固有状態の重ね合わせであることである。

n 次元確率分布 $\mu(x_1, \dots, x_n)$ が状態 $|\psi\rangle$ における $X_1, \dots, X_n \in \mathcal{O}(\mathcal{H})$ の**結合確率分布**であるとは、

$$\mu(x_1, \dots, x_n) = \Pr\{X_1 = x_1 \wedge \dots \wedge X_n = x_n \mid |\psi\rangle\} \quad (48)$$

を満たすことである。結合確率分布の存在は、物理量の同時決定性と同等である [5]。

定理 12 物理量 X_1, \dots, X_n が状態 $|\psi\rangle$ において同時決定的であるための必要十分条件は、 X_1, \dots, X_n の $|\psi\rangle$ における結合確率分布が存在することである。この時、物理量 X_1, \dots, X_n の任意の多項式 $f(X_1, \dots, X_n)$ に対して、

$$\langle \psi | f(X_1, \dots, X_n) | \psi \rangle = \sum_{(x_1, \dots, x_n) \in \mathbf{R}^n} f(x_1, \dots, x_n) \Pr\{X_1 = x_1 \wedge \dots \wedge X_n = x_n \mid |\psi\rangle\} \quad (49)$$

が成り立つ。

物理量 X_1, \dots, X_n を含む観測命題 $\phi(X_1, \dots, X_n)$ が状態 $|\psi\rangle$ で**文脈的に整合的である**とは、状態 $|\psi\rangle$ で X_1, \dots, X_n が同時決定的であることを言う。量子集合論の移行原理 [20] から、以下の定理が成り立つ。

定理 13 $\phi(X_1, \dots, X_n)$ が X_1, \dots, X_n を含む観測命題で, 結合子 \wedge, \neg に関して古典論理のトートロジーであれば,

$$\llbracket \text{com}(X_1, \dots, X_n) \rrbracket \leq \llbracket \phi(X_1, \dots, X_n) \rrbracket \quad (50)$$

が成り立つ. 特に, $\phi(X_1, \dots, X_n)$ が $|\psi\rangle$ において文脈的に整合的ならば, $\phi(X_1, \dots, X_n)$ は $|\psi\rangle$ で真である.

任意の物理量 X, Y に対して, 観測命題 $X = Y$ を

$$X = Y := \bigvee_{x \in \text{Sp}(X)} X = x \wedge Y = x \quad (51)$$

によって定義する. 観測命題 $X = Y$ が $|\psi\rangle$ で成立することと, $X =_{|\psi\rangle} Y$ が成立することは同値であり [19], このとき物理量 X と Y は状態 $|\psi\rangle$ において同一の値を持つという. 量子集合論から以下の定理が得られる [20].

定理 14 任意の物理量 X_1, \dots, X_n に対して,

$$\llbracket X_1 = X_2 \wedge \dots \wedge X_{n-1} = X_n \rrbracket \leq \llbracket \text{com}(X_1, \dots, X_n) \rrbracket \quad (52)$$

が成り立つ. 特に, 任意の状態 $|\psi\rangle$ に対して, $X_1 =_{|\psi\rangle} X_2, \dots, X_{n-1} =_{|\psi\rangle} X_n$ ならば, X_1, \dots, X_n は $|\psi\rangle$ において同時決定的である.

15 測定と文脈

われわれの理論では, $A(0)$ と $B(0)$ のメータ $M(\Delta t)$ による同時測定は, 次の2つの条件で定義されていた.

(i) $A(0) = f(M(\Delta t))$ が状態 $|\psi\rangle \otimes |\xi\rangle$ で成り立つ.

(ii) $B(0) = g(M(\Delta t))$ が状態 $|\psi\rangle \otimes |\xi\rangle$ で成り立つ.

(i) から $A(0)$ と $f(M(\Delta t))$ は $|\psi\rangle \otimes |\xi\rangle$ において同時決定的である. (ii) から同じことが $B(0)$ と $g(M(\Delta t))$ について成り立つ. しかし, このことは $A(0)$ と $B(0)$ が $|\psi\rangle \otimes |\xi\rangle$ で同時決定的であることを導かない. 従って, $A(0)$ と $B(0)$ の同時測定は, その2つの測定値が同時に実在を表していることを保証しない. これは, 同時測定を定義する2つの条件が, 一般に文脈的に整合的でないことによる.

条件 (i) より, 観測命題 $A(0) = f(M(\Delta t))$ が $|\psi\rangle \otimes |\xi\rangle$ で文脈的に整合的であり, 条件 (ii) より, 観測命題 $B(0) = g(M(\Delta t))$ についても同様である. 観測命題「 $A(0) = f(M(\Delta t))$ かつ $B(0) = g(M(\Delta t))$ 」は状態 $|\psi\rangle \otimes |\xi\rangle$ で真である. しかし, $A(0)$ と $B(0)$ が $|\psi\rangle \otimes |\xi\rangle$ で同時決定的でない, つまり, 可換でないならば, これは文脈的に整合的ではない. 従って, もし, 至る所で非可換な物理量 $A(0)$ と $B(0)$ の同時測定を行ない, $A(0) = a$ かつ $B(0) = b$ という測定値を得たとすれば, $A(0) = a$ という事実を $A(0)$ を含む一つの文脈で, あたかも古典論理に従う実在の性

質として未来の予測などに使用し、また、同様に、 $B(0) = b$ という事実を使用することができるが、しかし、「 $A(0) = a$ かつ $B(0) = b$ 」という統一された事実を實在の性質として使用することはできない。このような、文脈を越えた古典論理の使用は定理13によって、明らかに制限されている。たとえば、物理量の対の同一性を、

$$\langle A, B \rangle = \langle C, D \rangle := \bigvee_{x \in \text{Sp}(A), y \in \text{Sp}(B)} A = x \wedge B = y \wedge C = x \wedge D = y \quad (53)$$

で定義することができる。古典論理に従えば、

$$(A = C \wedge B = D) \Rightarrow \langle A, B \rangle = \langle C, D \rangle \quad (54)$$

はトートロジーであるが、量子論理での真理値は常に真とは限らない。定理13によれば、

$$\llbracket \text{com}(A, B, C, D) \rrbracket \leq \llbracket (A = C \wedge B = D) \Rightarrow \langle A, B \rangle = \langle C, D \rangle \rrbracket \quad (55)$$

となる。ただし、量子論理での \Rightarrow は、通常、 $\phi_1 \Rightarrow \phi_2 := (\neg \phi_1) \vee (\phi_1 \wedge \phi_2)$ と定義される。

量子論理は、量子力学の2つの側面に光を当てている。量子状態は、観測命題のなす束の上の確率測度として定義される。従って、観測命題に対する射影作用素値の真理値及び確率の付値は、Bornの統計公式と整合的であり、文脈に依存せず、多数回の同種の測定の結果の統計に関するわれわれの経験を説明する。この側面は、Ballentine [21]によって定式化され、しばしば、量子力学の統計的解釈と呼ばれる。一方、観測命題に対する2値の付値は文脈的であるが、単一の測定の結果についての我々の経験を、文脈によって定まる観測命題のある部分系に古典論理の法則を適用することによって、説明する。この側面は、古典物理学によって培われた物理的實在に関するわれわれの直観を量子力学的対象にどのように適用すべきかを説明している。例えば、保存量概念はこの意味で解釈できる。この立場は、Bohrの相補性原理 [22]によって、典型的に強調されてきた。

参考文献

- [1] S. Kochen, E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. 17 (1967) 59–87.
- [2] J. von Neumann, Mathematische Grundlagen der Quantenmechanik, Springer, Berlin, 1932.
- [3] E. P. Wigner, The problem of measurement, Am. J. Phys. 31 (1963) 6–15.
- [4] C. W. Helstrom, Quantum Detection and Estimation Theory, Academic, New York, 1976.
- [5] S. Gudder, Joint distributions of observables, J. Math. Mech. 18 (1968) 325–335.

- [6] M. Ozawa, Universal uncertainty principle, simultaneous measurability, and weak values, in: Proceedings of the 10th International Conference on Quantum Communication, Measurement and Computing, AIP Conference Proceedings, American Institute of Physics, Melville, 2011, e-print arXiv:1106.5083 [quant-ph].
- [7] G. Birkhoff, J. von Neumann, The logic of quantum mechanics, *Ann. Math.* 37 (1936) 823–845.
- [8] J. G. Kirkwood, Quantum statistics of almost classical assemblies, *Phys. Rev.* 44 (1933) 31–37.
- [9] Y. Aharonov, D. Z. Albert, L. Vaidman, How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100, *Phys. Rev. Lett.* 60 (1988) 1351–1354.
- [10] A. M. Steinberg, Conditional probabilities in quantum theory and the tunneling-time controversy, *Phys. Rev. A* 52 (1995) 32–42.
- [11] A. M. Steinberg, How much time does a tunneling particle spend in the barrier region?, *Phys. Rev. Lett.* 74 (1995) 2405–2409.
- [12] M. Ozawa, Perfect correlations between noncommuting observables, *Phys. Lett. A* 335 (2005) 11–19.
- [13] M. Ozawa, Quantum perfect correlations, *Ann. Phys. (N.Y.)* 321 (2006) 744–769.
- [14] M. Ozawa, Quantum measuring processes of continuous observables, *J. Math. Phys.* 25 (1984) 79–87.
- [15] M. Ozawa, Universally valid reformulation of the Heisenberg uncertainty principle on noise and disturbance in measurement, *Phys. Rev. A* 67 (2003) 042105 (1–6).
- [16] M. Ozawa, Uncertainty principle for quantum instruments and computing, *Int. J. Quant. Inf.* 1 (2003) 569–588.
- [17] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Z. Phys.* 43 (1927) 172–198.
- [18] M. J. W. Hall, Prior information: How to circumvent the standard joint-measurement uncertainty relation, *Phys. Rev. A* 69 (2004) 052113.
- [19] M. Ozawa, Quantum reality and measurement: A quantum logical approach, *Found. Phys.* 41 (2011) 592–607.
- [20] M. Ozawa, Transfer principle in quantum set theory, *J. Symbolic Logic* 72 (2007) 625–648.

- [21] L. E. Ballentine, The statistical interpretation of quantum mechanics, *Rev. Mod. Phys.* 42 (1970) 358–381.
- [22] N. Bohr, The quantum postulate and the recent development of atomic theory, *Nature* (London) 121 (1928) 580–590.

Space Time Noise

空間・時間・ノイズ

飛田 武幸 (Takeyuki Hida, 名古屋大学・名城大学 名誉教授)
Si Si (愛知 県立大学)

概要

First we explain the idea of using noise to discuss the analysis of random functions, and we explain the role of a noise. We then come to introducing a new noise depending on the space parameter and discuss its application.

1 Introduction

The main purpose of this report is to introduce a new noise, As an application, we can discuss decomposition of a compound Poisson process into infinitesimal Poisson type random variables which are mutually independent. Those independent variables form a system of *idealized elemental random variables*, often called simply a *noise*.

We can roughly classify various noises according to the parameters on which noises depend.

i) Depending on time $t \in R^1$:

White noise (Gaussian) $\dot{B}(t)$

Poisson noise $\dot{P}(t)$

ii) Depending on state variable $u \geq 0$:

New noise $P'(u)$.

This noise is exactly what we are going to introduce.

Before we come to the main subject, some of the motivations of our work shall be presented.

1) From classical theory.

We refer to the literature [18], where the *elementare Gruppe* acting on parameter sets, time and space, is considered and discusses the matter depending on them. We may, with this in mind, change the matter with the *noise* and discuss how the noise is parametrized by those parameters.

Another Weyl, what we should like to mention is P. Lévy's original work on the decomposition of Lévy process. See [8] that was published before his famous book [11].

He takes an additive process, actually the Lévy process $L(t)$. It is understood them as a continuous analogue of a sum of independent random variables. The sum is equivalent to the sequence of independent random variables. See Reduction, that we shall propose in the next section.

We can see there the so-called Lévy-Itô decomposition of a Lévy process $L(t)$. To establish the decomposition Lévy introduces a variable $N(t, u)$, which describes jump u at instant t , intuitively speaking. It is the key notion of forming the decomposition. It should be noted that the relationship between the time t and the space u is illustrated, although implicitly (from our viewpoint).

2) From quantum dynamics: no need to be mentioned in this report.

3) From the present viewpoint.

We have been trying to discover significant relationship between space and time, when we discuss a representation of them as systems of independent random variables, i.e. within random world.

Again, we remind the Lévy process $L(t)$, as a continuous analogue of a sum of independent identically distributed random variables, so that it should be taken to be a typical subject when we think of the reduction (see next section).

The approach to randomness by D. Mumford [14] has encouraged our study; in particular, his idea to appeal to the decomposition of a Lévy process.

2 Preliminaries

2.1 Reduction, Synthesis and Analysis

Our approach to stochastic analysis starts with a system of idealized elemental random variables (i.i.d., i.e.r.v.) which are independent, atomic, identically distributed, infinitesimal random variables, parametrized by a point of an ordered set. To fix the idea we assume the ordered set is either R^1 or its subinterval. In such a case, we often call it *noise*.

Given a random complex phenomenon. We wish to find a noise such that the phenomenon in question is expressed as a function of the noise so that it is ready to be analysed. This is the step of **reduction** and **synthesis**.

Then, follows the **analysis** of the functions, actually nonlinear functionals. The setup of the analysis involves the choice of the class of functionals. In fact, we propose to take a class of **generalized**

functionals of a noise and come to introduce **operators** acting on them. Thus, we are ready to carry on the analysis.

2.2 Systems of i.e.r.v.'s; Noises

We have a quick review of the typical noise. It is the (Gaussian) white noise and is realized by the time derivative $\dot{B}(t)$ of a Brownian motion $B(t)$.

The probability distribution μ of the system $\{\dot{B}(t)\}$ is given on the measurable space (E^*, \mathcal{B}) , where E^* is the dual space of a nuclear space E dense in $L^2(\mathbb{R}^1)$, and where the sigma-field \mathcal{B} is generated by cylinder subsets of E^* .

There is another noise called the Poisson noise, obtained by taking the time derivative $\dot{P}(t)$ of a Poisson process $P(t)$.

Although i.e.r.v.'s are idealized variable, they can rigorously be discussed in mathematics. They form a base of a Hilbert space and form a total set in the space of generalized linear functionals. Naturally, approximation by ordinary random variables would be requested. We may use the technique from discrete to continuous.

2.3 Birth of noises

One may ask what kind of noise can exist naturally. Some details on this topic, we have reported in the monograph [16] and the paper [5], so we do not go into details, but just state the results.

1) Gaussian noise

To fix the idea, we take the unit interval $I = [0, 1]$ to be the parameter set. Follow the idea of reduction and use the technique that from discrete to continuous. And we take a partition $\{\Delta_k^n, 1 \leq k \leq 2^n\}$ of I with $|\Delta_k^n| = 2^{-n}$. Let $X_k^n, 1 \leq k \leq 2^n$ be a family of independent random variables subject to the same distribution with finite variance σ^2 , and assume that each X_k^n is associated with Δ_k^n .

With this setup, we are ready to appeal to the Central Limit Theorem to have a Gaussian distribution. For any partial sum of X_k^n 's associated with consecutive subintervals Δ_k^n 's, we also have a Gaussian distribution. The collection of those Gaussian distributions forms a consistent family. It can be proved that the family determines a Brownian motion. Its time derivative can be associated to each point of I so that a white noise is determined.

Since atomic property is required we have finally come to random variables indexed by $u \in [0, 1]$.

2) Poisson noise

Interval I and the partitions are the same as above. The X_k^n is now atomic:

$$P(X_k^n = 1) = p_n, P(X_k^n = 0) = 1 - p_n.$$

Since those random variables are independent for any fixed n . Hence their sum is subject to the binomial distribution. Assume that n is large and p_n is small. Then, the law of small probability, we have a Poisson distribution. If $2^n p_n$ is kept to be constant, say λ , then the intensity of the Poisson distribution is λ . Further we see, as in 1), the consistent family of Poisson processes determined a Poisson process with intensity λ , denoted by $P(t, \lambda)$.

Important remark is that different intensities determine different type Poisson distributions.

3) For each λ we may define a Poisson process with jump $u = u(\lambda)$. Or we may consider a function $\lambda = \lambda(u)$.

Such an observation has suggested us to consider the characteristic functional that appears in the next section.

3 A noise depending on space parameter

3.1 Noise $P'(u)$

The idea is to choose a parameter different from time t , say space variable u .

Let us start out with a functional $C^P(\xi)$, where the variable ξ runs through a certain nuclear space E , say isomorphic to the Schwartz space \mathcal{S} ;

$$C^P(\xi) = \exp\left[\int (e^{iu\xi(u)} - 1)dn(u)\right], \quad (3.1)$$

where $dn(u)$ is a measure on $(0, \infty)$, which is specified later.

To fix the idea, we assume that the measure $dn(u)$ is equivalent to the Lebesgue measure, i.e. it is of the form $dn(u) = \lambda(u)du$ with $\lambda(u)$ positive a.e. and integrable.

Theorem 1 *Under these assumptions, the functional $C^P(\xi)$ is a characteristic functional. That is*

- i) $C^P(\xi)$ is continuous in ξ ,
- ii) $C^P(0) = 1$.
- iii) Positive definiteness is shown by noting the fact that $\exp[(e^{izu} - 1)\lambda]$ is a characteristic function.

Proof is easy, so is omitted. See e.g. [16].

Applying the Bochner-Minlos theorem, we can see that there exists a probability measure ν^P on E^* such that

$$C^P(\xi) = \int_{E^*} e^{i\langle x, \xi \rangle} d\nu^P(x). \quad (3.2)$$

We introduce a notation $P'(u, \lambda(u))$ or write it simply $P'(u)$. We understand that ν^P -almost all $x \in E^*$ is a sample function of $P'(u)$, more precisely, a sample function of $P'(u, \lambda(u))$.

Theorem 2 *The $\{P'(u), u \in (0, \infty)\}$ defines a system of i.e.r.v., in fact the system is a noise.*

Proof. Suppose ξ_1 and ξ_2 have disjoint support; let them be denoted by A_1 and A_2 . Then, the integral in the expression (2.1) can be expressed in the form

$$\int_{A_1} (e^{iu\xi(u)} - 1)dn(u) + \int_{A_2} (e^{iu\xi(u)} - 1)dn(u).$$

This proves the equality

$$C^P(\xi_1 + \xi_2) = C^P(\xi_1)C^P(\xi_2).$$

This means that $P'(u)$ has independent value at every point u . Also follows the atomic property.

3.2 Observations

We shall see intuitive meaning of $P'(u)$.

Start with the characteristic function of Poisson random variable $P(\lambda)$ with intensity λ :

$$E[e^{izP(\lambda)}] = e^{\lambda(e^{iz} - 1)}.$$

Consider $uP(\lambda)$ to have

$$E[e^{izuP(\lambda)}] = e^{\lambda(e^{izu} - 1)}.$$

Let $\lambda = \lambda(u)$. Take independent variables $u_j P(\lambda(u_j)), j = 1, 2, \dots$ and have a sum $\sum u_j P(\lambda(u_j))$. Then we have

$$E[e^{i \sum z_j u_j P(\lambda(u_j))}] = e^{\sum (e^{z_j u_j} - 1)\lambda(u_j)}$$

Now, set $\lambda(u_j) = \lambda'(u_j)du_j$ to make $P(\lambda(u))$ to be infinitesimal, and let $z_j = \xi(u_j)$. Then, we have

$$E[e^{i \sum \xi(u_j)u_j P(\lambda'(u_j)du_j)}] = e^{\sum (e^{i\xi(u_j)u_j} - 1)\lambda'(u_j)du_j}.$$

Set $P(\lambda'(u)du) = P'(\lambda'(u))du$. Then, we have the limit

$$E(e^{i \int \xi(u)u P'(\lambda'(u))du})$$

which is equal to the $C^P(\xi)$ from which we have started, where $\lambda(u)$ is replaced by $\lambda'(u)$.

Now we have the intuitive meaning of the symbol $P'(u)$.

Remark 3.1 *In the case where the parameter u runs through the negative interval $(0, -\infty)$, new noise can also be discussed in the similar manner, It is, however, noted that the single point mass at $u = 0$ is omitted.*

4 Representation of generalized functionals of $P'(u)$'s

4.1 Space of linear functionals

We now come to the discussion on the linear functionals of $P'(u)$. (For $P'(u)$ the symbol P_{du} was used in [2]) although there are somewhat formal expressions.

Anyhow linear functionals are ordinary random variables and they form a space $\mathcal{H}_1(P)$.

Compute

$$\frac{1}{i} \frac{d}{dt} C^P(\xi + t\eta)|_{t=0} = \int e^{iu\xi(u)} u\eta(u)\lambda(u)du \cdot C^P(\xi).$$

Up to the common factor $C^P(\xi)$, we are given a representation (similar to the U -functional in the Gaussian case) of $\mathcal{H}_1(P)$ -functionals expressed in the form

$$\int (e^{iu\xi(u)} - 1)u\eta(u)\lambda(u)du,$$

by subtracting off the constant (expectation) from

$$\int e^{iu\xi(u)} u\eta(u)\lambda(u)du.$$

This is linear in η . We are, therefore, given a linear space

$$\mathcal{F}_1 = \text{span} \left\{ \int (e^{iu\xi(u)} - 1)u\eta(u)\lambda(u)du, \eta \in E \right\}, \quad (4.1)$$

which is isomorphic to $\mathcal{H}_1(P)$.

Note that $\lambda(u)$ is chosen so as the above integral converges and continuous in η , in addition to the condition that $\lambda(u) \neq 0$ for $u \in (0, \infty)$.

We know that $\langle x, \xi \rangle, \xi \in E$, is viewed as a sample of a random variable $\langle P', \xi \rangle$, the characteristic function of which is given by $\varphi^{P'}(z) = C^P(z\xi)$. Hence its mean value is $\int u\xi(u)\lambda(u)du$ and variance is $\int u^2\xi(u)^2\lambda(u)du$.

The bijection

$$\xi \iff \langle P', \xi \rangle, \xi \in E$$

leads us to have

$$f \iff \langle P', f \rangle, \quad u f \in L^2((0, \infty), d\lambda),$$

where $d\lambda = \lambda(u)du$.

By using the representation, we can rigorously define the random measure $p(du)$ (which was briefly mentioned before) such that for any Borel subset $B \subset (0, \infty)$ with finite $d\lambda$ measure

$$p(B) = \langle P', I_B \rangle$$

is defined, where I_B is the indicator function of the Borel set B .

Stochastic integrals based on $p(du)$ are defined in the usual manner, and the collections of the stochastic integrals forms a Hilbert space which is in agreement with $\mathcal{H}_1(P)$.

Also, as in the Gaussian case, we can define the extended space $\mathcal{H}_1(P)^{(-1)}$, where each $P'(u)$ is a member of the *total* system in $\mathcal{H}_1(P)^{(-1)}$. Again, we note that $P'(u)$ corresponds to the kernel which is the δ -function, except $u = 0$.

There the $P'(u)$ has rigorous, and the definite identity.

4.2 Linear processes and their canonical representations

Define

$$X(u) = \int_0^u F(u, v) P'(v) dv, \quad (4.2)$$

which is called a *linear process* cf. P. Lévy [13]. His theory is concerned with $X(t)$ depending on the time parameter t , not on u , but we can discuss in the same mind.

Like Gaussian case, we can discuss *canonical representations*. The idea is the same as in the Gaussian case, but the definition is slightly different. The representation 4.2 is canonical if

$$\mathcal{B}_u(X) = \mathcal{B}_u(P').$$

It is our hope that the canonical property can be discussed in the sample function level. For one thing, the jump finding problem can be discussed in the $\mathcal{H}_1^{(-1)}$ -level (see [7]). For another thing, $P'(u)$ noise depends on the space variable u , so that the value of u should be concerned directly.

4.3 Nonlinear case

We can also consider multi-linear case in a similar manner, and the spaces \mathcal{F}_n with $n \geq 2$ can be defined similarly, and of course, there is required to have functionals renormalized.

We will not enter the general theory of nonlinear functionals of the $P'(u)$, however we should like to emphasize the significance of quadratic forms of the $P'(u)$.

参考文献

- [1] I.M. Gel'fand and N.Ya. Vilenkin, Generalized functions. vol.4, Academic Press. Russian original, Moscow 1961.
- [2] T. Hida, Stationary stochastic processes. Princeton Univ. Press, 1970.
- [3] T. Hida. Analysis of Brownian functionals. Carleton Math. Notes. no.13 , 1975. Carleton Univ.
- [4] 飛田武幸、Dual pairs の系から見たホワイトノイズ超汎関数空間の特徴。RIMS 共同研究 非可換解析とマイクロ・マクロ双対性、2007. 146-160.
- [5] T. Hida, A noise of new type and its generalized functionals. Proc. 13th Workshop on Noncommutative Harmonic Analysis. Bedlewo, 2010. to appear.
- [6] T. Hida and Si Si, Lectures on white noise functionals. World Sci. Pub. Co. 2008.
- [7] T. Hida, Si Si and Win Win Htay, A Noise of new type and its application. to appear.
- [8] P. Lévy, Sur les intégrales dont les éléments sont des variables aléatoires indépendentes. Ann. d. R. Scuola Normale Superiore di Pisa. ser II, 3 (1934) 337-366.
- [9] P. Lévy, Notice sur les Travaux Scientifiques. Herman & Cie, 1935.
- [10] P. Lévy, Sur les exponentielles de polynômes et sur l'arithmétique des produits de lois de Poisson. Ann. Ecole Normale Sup. 54 (1937), 231-292. (in particular Chap. 2).
- [11] P. Lévy, Théorie de l'addition des variables aléatoires. Gauthier-Villars, 1938. 2ème éd. 1954.
- [12] P. Lévy, Problèmes concrets d'analyse fonctionnelle. Gauthier-Villars, 1951.
- [13] P. Lévy, Fonctions aléatoires à corrélation linéaire. Illinois J. Math. 1 (1957), 217-258.
- [14] D. Mumford and B. Gidas, Stochastic models for generic images. preprint , (2000).
- [15] 小嶋 泉 編、量子科学における双対性とスケール。 RIMS 講究録 no.1705 (2010).

- [16] Si Si, Introduction to Hida distributions. World Sci. Pub. Co. to appear 2011.
- [17] Si Si A.H. Tsoi and Win Win Htay, Jump finding of a stable process. Quantum Information. vol. V, World Sci. Pub. Co. (2000), 193-201.
- [18] H. Weyl, Raum · Zeit · Materie, 1919, 5th ed.1923, Springer: Japanese Transl, by T.Uchiyama.

同種粒子系の量子もつれについて

近畿大学 量子コンピュータ研究センター 市川 翼¹
 東京大学大学院 理学系研究科 佐々木 寿彦
 高エネルギー加速器研究機構 理論センター 筒井 泉

1 はじめに

同種粒子性にまつわる混乱というのは、量子論の発見以来、度々繰り返されてきたのではないかと思われる。そもそも量子論発見の端緒となった黒体輻射自体、同種粒子性にまつわる問題であり、プランクの輻射公式の提案からボーズによるそのミクロな立場からの基礎付け迄にほぼ四半世紀を費している。ハンプリー・ブラウン・ツイス効果にまつわる研究史を紐解けば、やはり論争があったことが分かるだろう [1, 2, 3].

量子もつれ [4] も例外ではない。量子情報科学の進展に伴って近年盛んに研究がなされ、理論的整備もなされつつあるものの、それは異種粒子系に限った状況であり、同種粒子系に目を移せば百家争鳴である [5, 6, 7, 8, 9, 10]. しかも奇妙なことに、量子情報実験で使われる物理系はほぼ全て光子系や電子系などの同種粒子系であるにもかかわらず、異種粒子系での量子もつれの理論と整合的なのである。以上の二点からすれば、i) 諸家の提案を包摂し、ii) 異種粒子系での量子もつれの理論を内包する理論が、同種粒子系の量子もつれには望ましい。

本報告では我々が [11, 12] で提案した理論の概要を述べたい。これが i) と ii) を満足するものであることも以下で説明してゆく。

2 同種粒子系特有の問題

この節で、同種粒子系の量子もつれがはらむ問題点を浮き彫りにしたい。そのためには異種粒子系の量子もつれの定義が何に依拠しているかを概観し、同種粒子系のもつ諸性質と相容れるか否かを検討する必要があるだろう。簡単のため、以下では粒子の統計性の如何によらず純粋状態のみを扱い、有限自由度 2 体系のみを考察する。多体系への拡張は [11, 12] を参照のこと。

異種粒子系の量子もつれ さて、2 体異種粒子系を考えよう。i 番目の粒子の状態空間は d_i 次元複素ヒルベルト空間 $\mathcal{H}_i = \mathbb{C}^{d_i}$ で表され、全体系の状態空間はそのテンソル積空間 $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ で表される。このとき 1 粒子状態の直積状態を可分な状態 (separable state) という。つまり可分

¹E-mail: tsubasa@alice.math.kindai.ac.jp

な状態 $|\Psi\rangle \in \mathcal{H}$ については適当な状態 $|\psi_i\rangle_i \in \mathcal{H}_i$ が存在し、 $|\Psi\rangle = |\psi_1\rangle_1 \otimes |\psi_2\rangle_2$ と書くことができる。一方、可分ではない状態を量子もつれ状態 (entangled state) という。つまり 1 粒子状態の直積として書けない状態が量子もつれ状態である。

以上の定義の物理的帰結を吟味してみよう。そのために局所物理量を導入し、その相関を考える。 \mathcal{H}_i 上に作用するエルミート演算子を O_i とおく。局所物理量は \mathcal{H}_i 上の恒等演算子 $\mathbb{1}_i$ を用いて

$$\hat{O}_1 = O_1 \otimes \mathbb{1}_2, \quad \hat{O}_2 = \mathbb{1}_1 \otimes O_2 \quad (1)$$

と定義し、それらの相関を \hat{O} を $\hat{O} = \hat{O}_1 \hat{O}_2$ で定める。明らかに

$$\hat{O} = \hat{O}_1 \hat{O}_2 = \hat{O}_2 \hat{O}_1 = O_1 \otimes O_2 \quad (2)$$

であり。局所物理量の可分な状態に関する相関は

$$\langle \Psi | \hat{O} | \Psi \rangle = \langle \Psi | \hat{O}_1 | \Psi \rangle \langle \Psi | \hat{O}_2 | \Psi \rangle = {}_1\langle \psi_1 | O_1 | \psi_1 \rangle_1 \cdot {}_2\langle \psi_2 | O_2 | \psi_2 \rangle_2 \quad (3)$$

と部分系の期待値の積で与えられ、一方の部分系の情報は他方の情報には依存しないことが分かる。対して量子もつれ状態で \hat{O} の期待値をとった場合、式 (3) は成立せず、個々の部分系の独立性が担保できない。これが量子もつれ状態が持つ非局所性である。

同種粒子系の標準的記述 ここまでの議論で分かるのは、量子もつれ状態にせよ局所物理量にせよ、異種粒子系の状態空間の自然なテンソル積構造を用いて定義されていたということである。一方同種粒子系の場合、同種性から $d_1 = d_2 =: d$ であり、ボゾンについては対称化演算子 S を、フェルミオンについては反対称化演算子 A を用いることで、状態空間は

$$\mathcal{H}_\chi := \{ |\Psi\rangle \mid \mathcal{X}|\Psi\rangle = |\Psi\rangle, |\Psi\rangle \in \mathcal{H} \} \subset \mathcal{H}, \quad \mathcal{X} = A, S \quad (4)$$

で与えられる。 S 及び A は具体的には

$$S = (\mathbb{1}_1 \otimes \mathbb{1}_2 + \pi) / 2, \quad A = (\mathbb{1}_1 \otimes \mathbb{1}_2 - \pi) / 2 \quad (5)$$

であり、 π はエルミートで、1 粒子状態空間 \mathcal{H}_i のラベルへの互換 $(1, 2) \in \mathfrak{S}_2$ に対応する演算子。 \mathcal{H}_χ は \mathcal{H} の部分空間ではあるが、(反)対称化を受けているため、テンソル積構造の有無は非自明な問題である。よって例えフェルミオン系が直積として書きようがなかりと、異種粒子系での定義を用いて量子もつれの有無を判断することは意味がないことになる。このテンソル積構造の有無の非自明性が同種粒子系の量子もつれに関する核心であるといってもよい。

3 テンソル積構造とその物理的意味

前節では、同種粒子系の量子もつれが孕む問題点が状態空間のテンソル積構造の有無に帰着されることを指摘した。この問題への対策として、本報告では \mathcal{H}_χ (の部分空間) にテンソル積構造を入れる方法を具体的に与えることにする。この手法では、入るテンソル積構造に任意性が残るが、これについても自然な物理的解釈が可能である。

3.1 直交構造にもとづいた量子もつれ

本報告で手法を解説する準備として、いくつか集合や空間を定義する。まず 1 粒子状態空間 $\mathcal{H}_i \cong \mathbb{C}^d$ の直交部分空間の集合 V を

$$V = \{V_1, V_2\}, \quad V_1 \perp V_2, \quad V_i \subset \mathbb{C}^d \quad (6)$$

で定義する。この V を直交構造と呼ぶことにする。 \mathcal{H}_j の部分空間としての V_i を $V_i(\mathcal{H}_j)$ と書き、

$$\mathcal{H}(V) := V_1(\mathcal{H}_1) \otimes V_2(\mathcal{H}_2) \subset \mathcal{H} \quad (7)$$

なる空間と、その(反)対称化された空間

$$\mathcal{H}_\chi(V) := \{|\Psi\rangle \mid |\Psi\rangle = \chi|\Phi\rangle, \quad |\Phi\rangle \in \mathcal{H}(V)\} \subset \mathcal{H}_\chi \quad (8)$$

を用意する。 $\mathcal{H}_\chi(V)$ を用いて同種粒子系の状態空間 \mathcal{H}_χ は

$$\mathcal{H}_\chi = \mathcal{H}_\chi(V) \oplus \mathcal{H}_\chi(V)^\perp \quad (9)$$

と直和分解でき、これに応じて状態 $|\Psi\rangle \in \mathcal{H}_\chi$ も

$$|\Psi\rangle = |\Psi(V)\rangle + |\Psi(V)^\perp\rangle, \quad |\Psi(V)\rangle \in \mathcal{H}_\chi(V), \quad |\Psi(V)^\perp\rangle \in \mathcal{H}_\chi(V)^\perp \quad (10)$$

と分解できることになる。

以上の準備のもと、同種粒子系の状態 $|\Psi\rangle \in \mathcal{H}_\chi$ が量子もつれ状態であることの定義を与えよう。

定義 1 $|\Psi\rangle \in \mathcal{H}_\chi$ と V から式 (10) によって導かれる $|\Psi(V)\rangle \in \mathcal{H}_\chi(V)$ が

$$|\Psi(V)\rangle = \sqrt{2}\chi (|\psi_1\rangle_1 \otimes |\psi_2\rangle_2), \quad |\psi_i\rangle_i \in V_i(\mathcal{H}_i) \quad (11)$$

と書けないとき、 $|\Psi\rangle$ は V について量子もつれ状態である。

この定義が前節の問題意識のもとで適切なものであるかについては吟味が必要だろう。 $|\Psi(V)^\perp\rangle$ を考慮しなくてよい理由は後述するとして、ここで取り扱っている空間は本質的には $\mathcal{H}_\chi(V)$ であるから、 $\mathcal{H}_\chi(V)$ にテンソル積構造があり、なおかつ式 (11) を満たす $|\Psi(V)\rangle$ がそのテンソル積構造のもとで直積とみなせれば、異種粒子系の量子もつれと完全な対応が成り立ち、適切な定義といえる。以下で実際そうになっていることを示す。

3.2 同型対応

さて、 $\mathcal{H}_\chi(V)$ にテンソル積構造があること段階を踏んで示そう。まず、

$$f_\chi (|\psi_1\rangle_1 \otimes |\psi_2\rangle_2) = \sqrt{2}\chi (|\psi_1\rangle_1 \otimes |\psi_2\rangle_2) \quad (12)$$

で定義される線型写像

$$f_\chi : \mathcal{H}(V) \mapsto \mathcal{H}_\chi(V), \quad (13)$$

を導入する。この写像について以下の主張が成立する。

補題 1 任意の $|\psi_i\rangle_i \in V_i(\mathcal{H}_i)$ に対し,

$$\|f_{\mathcal{X}}(|\psi_1\rangle_1 \otimes |\psi_2\rangle_2)\| = \| |\psi_1\rangle_1 \otimes |\psi_2\rangle_2 \| . \quad (14)$$

証明. 式 (5) と式 (12) より

$$f_{\mathcal{X}}(|\psi_1\rangle_1 \otimes |\psi_2\rangle_2) = (|\psi_1\rangle_1 \otimes |\psi_2\rangle_2 \pm |\psi_2\rangle_1 \otimes |\psi_1\rangle_2) / \sqrt{2}. \quad (15)$$

$|\psi_i\rangle_j \in \mathcal{H}_j(V_i)$ であるから, V_1 と V_2 の直交性により ${}_i\langle\psi_1|\psi_2\rangle_i = 0$. よって $V_i(\mathcal{H}_1) \cong V_i(\mathcal{H}_2)$ を用いれば

$$\|f_{\mathcal{X}}(|\psi_1\rangle_1 \otimes |\psi_2\rangle_2)\|^2 = (\| |\psi_1\rangle_1 \otimes |\psi_2\rangle_2 \|^2 + \| |\psi_2\rangle_1 \otimes |\psi_1\rangle_2 \|^2) / 2 = \| |\psi_1\rangle_1 \otimes |\psi_2\rangle_2 \|^2 \quad (16)$$

がいえ, 題意が示せた. ■

補題 1 からただちに以下がいえる:

系 1 $f_{\mathcal{X}}$ は全単射.

証明. $\mathcal{H}(V)$ の基底系として, $V_i(\mathcal{H}_j)$ の正規直交基底系 $\{|e_i(\alpha)\rangle_j\}_{\alpha=1}^{\dim V_i}$ から構成した $\{|e_1(\alpha)\rangle_1 \otimes |e_2(\beta)\rangle_2\}_{\alpha,\beta}$ をとることができる. この基底系で $\mathcal{H}(V)$ の要素を展開し補題 1 を用いれば, $f_{\mathcal{X}}$ がノルム保存であることがいえる. よって $f_{\mathcal{X}}$ は単射. 全射性は $f_{\mathcal{X}}$ の定義より明らか. ■

系 1 より $\mathcal{H}_{\mathcal{X}}(V)$ は $\mathcal{H}(V)$ と同型. $\mathcal{H}(V)$ は明らかにテンソル積空間なので,

定理 1 $\mathcal{H}_{\mathcal{X}}(V)$ はテンソル積構造を持つ.

さらに式 (12) より, 式 (11) の右辺の状態がこのテンソル積構造のもとでの直積状態であることは言を俟たないだろう. 定義 1 はテンソル積構造に依拠しており, この意味で異種粒子系における量子もつれの定義と同格のものとなっているのである.

次に局所物理量の相関についても同種粒子系での対応物を構成しよう. O_i として $V_i(\mathcal{H}_i)$ 上の物理量を採用して式 (2) で \hat{O} を定義し, それを同型写像 $f_{\mathcal{X}}$ で

$$\tilde{O} = f_{\mathcal{X}} \hat{O} f_{\mathcal{X}}^\dagger = 2\mathcal{X} (O_1 \otimes O_2) \mathcal{X} \quad (17)$$

と写し, 同種粒子系での相関演算子とする. チルダ付きの演算子はこの写像でハット付きの演算子から移したものと約束する. このとき直和分解 (10) より

$$\tilde{O}|\Psi(V)^\perp\rangle = 0 \quad (18)$$

が成り立つ. これは $|\Psi(V)^\perp\rangle$ が相関を測定する実験には寄与しないことを示している. 量子もつれの有無は相関の有無に他ならないので, 量子もつれを判定する際には $|\Psi(V)^\perp\rangle$ を考慮する必要はない. これが定義 1 で $|\Psi(V)^\perp\rangle$ を考慮しない理由である. また, 実験においては得られた総データを基準に規格化がなされる. 先の考察から, $|\Psi\rangle$ において実験にかかる成分は $|\Psi(V)\rangle$ であり,

$\|\Psi(V)\| = 1$ となるよう再規格化せねばならない。この再規格化のもと、 $|\Psi\rangle$ が V について可分な状態のとき、相関の分離

$$\langle\Psi|\tilde{O}|\Psi\rangle = \langle\Psi|\tilde{O}_1|\Psi\rangle\langle\Psi|\tilde{O}_2|\Psi\rangle \quad (19)$$

が成立する。証明については、まず f_X が内積を保存することを補題 1 と同様の方法で示し、次に \tilde{O} 及び $|\Psi\rangle$ をスペクトル分解する。すると左辺が内積の和に書けるので、それらに f_X が内積を保存することを使えばよい [11]。これで今回提案した枠組みが、与えられた V のもとで異種粒子系の量子もつれの理論と完全に対応することが示せた。

3.3 測定手法との関係

ここまでで同種粒子系の状態空間に内在するテンソル積構造について解説した。この手法で特徴的なことは $\mathcal{H}_X(V)$ は我々が自由に選択可能な構造 V に依存しており、無限の任意性があるということである。この任意性をどう解釈すればよいだろうか？

これは具体例を考えると分かりやすい。同種粒子系のスピンと位置について先程解説した手法を適用してみるのである。位置については離散化し、左右それぞれに局在している状態 $|L\rangle, |R\rangle$ が正規直交基底を張るものとする。このとき 1 粒子状態空間は

$$\mathcal{H}_i = \text{span}\{|\uparrow\rangle, |\downarrow\rangle\} \otimes \text{span}\{|L\rangle, |R\rangle\} \cong \mathbb{C}^4 \quad (20)$$

となる。ここで $V = \{V_1, V_2\}$ として

$$V_1 = \text{span}\{|\alpha\rangle \otimes |L\rangle \mid \alpha = \uparrow, \downarrow\}, \quad V_2 = \text{span}\{|\beta\rangle \otimes |R\rangle \mid \beta = \uparrow, \downarrow\} \quad (21)$$

をとることにすれば、

$$\mathcal{H}_X(V) = \text{span}\{f_X(|\alpha, L\rangle_1 |\beta, R\rangle_2) \mid \alpha, \beta = \uparrow, \downarrow\} \quad (22)$$

となり、一方の粒子が左に、他方の粒子が右にいる状況のみを考えていることになる。また相関演算子 \tilde{O} は $\mathcal{H}_X(V)$ のみに台を持つので、「左右に測定器を置き、それぞれが粒子を一個観測した」ことに対応する。次に、上下に局在した状態²

$$|U\rangle = (|L\rangle + |R\rangle)/2, \quad |D\rangle = (|L\rangle - |R\rangle)/2 \quad (23)$$

を導入し、別の直交構造 $V' = \{V'_1, V'_2\}$ として

$$V'_1 = \text{span}\{|\alpha\rangle \otimes |U\rangle \mid \alpha = \uparrow, \downarrow\}, \quad V'_2 = \text{span}\{|\beta\rangle \otimes |D\rangle \mid \beta = \uparrow, \downarrow\} \quad (24)$$

を採用して同じ議論をすれば、 $\mathcal{H}_X(V')$ 上での相関演算子は「上下に測定器を置き、それぞれが粒子を一個観測した」と解釈できる。この例から、 V を変えることが測定器の配置等を変えることに対応することが分かるだろう。つまり V の任意性は相関検証実験をどのようにデザインするか の任意性に対応しており、量子もつれは検証実験のセットアップごとに定義されるのである。

² $|L\rangle, |R\rangle$ をパウリ行列 σ_x の固有状態に対応させれば、 $|U\rangle, |D\rangle$ は σ_x の固有状態に対応する。

3.4 普遍的に可分な状態

状態がひとつ与えられても量子もつれはそれだけでは決まらず、測定のセットアップに依存して変化する。これが今までの議論で得られた主要な知見であった。ではセットアップに応じた量子もつれの変化はどんな状態にも起こるのであるだろうか？この問いについては既に答が得られている。

命題 1 状態 $|\Psi\rangle = |\psi\rangle_1 \otimes |\psi\rangle_2$ は任意の V に対して可分。

証明. V を与えると $|\psi\rangle_i$ は

$$|\psi\rangle_i = |\psi_1\rangle_i + |\psi_2\rangle_i + (|\psi\rangle_i - |\psi_1\rangle_i - |\psi_2\rangle_i), \quad |\psi_j\rangle_i \in V_j(\mathcal{H}_i) \quad (25)$$

と直和分解できる。これを上記の $|\Psi\rangle$ に代入すれば

$$|\Psi(V)\rangle = \sqrt{2}\mathcal{S}(|\psi_1\rangle_1 \otimes |\psi_2\rangle_2) \quad (26)$$

となり、定義 1 より $|\Psi\rangle$ は V について可分。この議論は任意の V について成り立つので、題意が示せた。■

この命題は先程の問いの反例になっている。また、 V の任意性を考慮に入れた場合、状態空間が非常に豊かな構造を持つようになることを示唆しているとも解釈できよう。

4 まとめ

本報告では、同種粒子系の量子もつれをどのように定義するかについてを論じた。異種粒子系の量子もつれにはテンソル積構造に依拠した定式化がなされていることに着目し、同種粒子系の状態空間にテンソル積構造を入れる方法を与えることにより量子もつれを定義した。この定義は異種粒子系での定義と完全な対応を持っており、冒頭で述べた項目 ii) を満たしている。

最後に我々の定式化と他の定式化の関係を述べておこう。我々の定義は形式的には Ghirardi らが [7, 10] で与えた定義の多体量子もつれへの一般化になっている。ただし Ghirardi らはテンソル積構造の存在について言及しておらず、EPR 論文 [4] で提案された実在性の判定基準を論拠にして量子もつれの定義を与えている。一方、[6, 9] ではテンソル積構造を用いて、異種粒子系における一般化された量子もつれを定義し、それにもとづいて様々な量子アルゴリズムの整理・解析を行なっている。我々の定式化はこの二つの流儀を同種粒子系の文脈において統一したものといえる。さらに、[10] では Ghirardi らの流儀での量子もつれの定義と [5, 8] で与えられた定義で齟齬が生じる場合があることが報告されている。この齟齬は我々の定式化における測定セットアップへの量子もつれの依存性と再解釈することが可能であり、[5, 8] も我々の定式化が取り込んでいるということになる [12]。以上から、我々の定式化は既存の 3 つの流儀を含んでおり、冒頭の項目 i) も満足しているといえる。

参考文献

- [1] R. Hanbury Brown and R. Q. Twiss, *J. Astrophys. Astr.* **15** (1994) 13. *Nature* **177** (1956) 27 より転載.
- [2] E. M. Purcell, *J. Astrophys. Astr.* **15** (1994) 27. *Nature* **178** (1956) 1449 より転載.
- [3] R. Hanbury Brown, *J. Astrophys. Astr.* **15** (1994) 39.
- [4] A. Einstein, B. Podolsky and N. Rosen, *Phys. Rev.* **47** (1935) 777.
- [5] J. Schliemann, J. I. Cirac, M. Kuś, M. Lewenstein and D. Loss, *Phys. Rev.* **A64** (2001) 022303.
- [6] P. Zanardi, *Phys. Rev. Lett.* **87** (2001) 077901.
- [7] G.-C. Ghirardi, L. Marinatto and T. Weber, *J. Stat. Phys.* **108** (2002) 49.
- [8] K. Eckert, J. Schliemann, D. Bruß and M. Lewenstein, *Ann. Phys.* **299** (2002) 88.
- [9] P. Zanardi, D. A. Lidar and S. Lloyd, *Phys. Rev. Lett.* **92** (2004) 060402.
- [10] G.-C. Ghirardi and L. Marinatto, *Phys. Rev.* **A70** (2004) 012109.
- [11] Tsubasa Ichikawa, Toshihiko Sasaki and Izumi Tsutsui, *J. Math. Phys.* **51** (2010) 062202.
- [12] Toshihiko Sasaki, Tsubasa Ichikawa and Izumi Tsutsui, *Phys. Rev.* **A83** (2011) 012123.

FEYNMAN-KAC TYPE FORMULA WITH CÁDLÁG PATH AND SPIN \mathbb{Z}_p

廣島 文生 九大・数理

1 形式的な経路積分表示

この論考では Schrödinger 作用素が生成する一径数半群の経路積分表示について述べる．形式的なおはなしからはじめ、厳密な記述は2章からはじめる．

一般に時間に依存しない Schrödinger 作用素 H はあるヒルベルト空間上の『下から有界な自己共役作用素』¹ として定義され、その物理系はユニタリー群 e^{-itH} によって時間発展すると考えられている．ポテンシャル $V(x)$ をもった Schrödinger 作用素

$$H = -\frac{1}{2}\Delta + V(x) \quad (1.1)$$

を考えよう．Schrödinger 方程式

$$i\partial_t\Psi(x, t) = H\Psi(x, t), \quad \Psi_0(x) = \Psi(x, 0), \quad (1.2)$$

の解は $\Psi = e^{-itH}\Psi_0$ で与えられる． $e^{-itH}\Psi_0$ を作用積分をもちいて経路積分表示できるというのは R. P. Feynman の画期的なアイデアである．それを概観してみよう．

$$\mathcal{H}(p, q) = \frac{1}{2}p^2 + V(q) \quad (1.3)$$

として、そのラグランジアンは

$$\mathcal{L}(q, \dot{q}) = \dot{q}p - H(p, q) = \frac{1}{2}\dot{q}^2 - V(q) \quad (1.4)$$

で与えられる．ここで $\dot{q} = \frac{\partial H}{\partial p} = p$ である．そうすると作用積分は $q : [0, t] \rightarrow \mathbb{R}^3$ を経路とし、 \dot{q}_s をその微分とすれば

$$\int_0^t \mathcal{L}(q_s, \dot{q}_s) ds = \int_0^t \frac{1}{2}\dot{q}_s^2 - V(q_s) ds \quad (1.5)$$

¹ $H^* = H$, i.e., $\text{Dom}(H) = \text{Dom}(H^*)$ かつ $Hf = H^*f$ となるとき H を自己共役作用素という． $H \subset H^*$ となるとき H を対称作用素という．また $\inf_{f \in \text{Dom}(H)} (f, Hf) > -\infty$ となるとき「下から有界」という．

で与えられる. $\Omega_{xy} = \{q : [0, t] \rightarrow \mathbb{R}^3 | q_0 = x, q_t = y\}$ は x から y へ動く経路全体の集合とすると, ユニタリー群 e^{-itH} の積分核が

$$(e^{-itH})(x, y) = \text{定数} \times \int_{\Omega_{xy}} e^{i \int_0^t (\frac{1}{2} \dot{q}_s^2 - V(q_s)) ds} \prod_{0 \leq s \leq t} dq_s \quad (1.6)$$

と表せるという. これが Feynman 経路積分である. もちろん $\prod_{0 \leq s \leq t} dq_s$ に対応する数学的に厳密な測度² を Ω_{xy} 上に構成できるわけではないが, このような表現が量子力学に限らず, 量子論全般にわたって非常に有用であることは言うまでもない.

さて唐突ではあるが時間 t を純虚数時間 $-it$ に形式的に置き換えて, ユニタリー群のかわりに半群 e^{-tH} を考えることにする³. そうすると

$$e^{-tH}(x, y) = \text{定数} \times \int_{\Omega_{xy}} e^{-\int_0^t V(q_s) ds} e^{-\frac{1}{2} \int_0^t \dot{q}_s^2 ds} \prod_{0 \leq s \leq t} dq_s \quad (1.7)$$

となることがわかる. さらに Ω_{xy}^D をいたるところ微分できない経路全体とすれば形式的には Ω_{xy}^D 上で

$$\exp\left(-\int_0^t \frac{1}{2} \dot{q}_s^2 ds\right) \prod_{0 \leq s \leq t} dq_s = 0 \times \infty < \infty$$

となることが予想できなくもない. 実際, 左辺に対応する測度を連続関数空間 $\mathcal{C} = C([0, \infty))$ 上に定義できる. 厳密に言えば次のようになる. $(\mathcal{C}, \mathcal{F}, P^x)$ をウィナー空間とする. つまり \mathcal{F} は \mathcal{C} に局所一様位相をいれて位相空間としたときのボレル σ 代数, P^x はウィナー測度で $B_t(\omega) = \omega(t)$, $\omega \in \mathcal{C}$, が x から出発する 3 次元ブラウン運動になる. $\mathbb{E}^x[\dots] = \int \dots dP^x$ とすれば

$$(f, e^{-tH}g) = \int dx \mathbb{E}^x \left[f(B_0)g(B_t) e^{-\int_0^t V(B_s) ds} \right] \quad (1.8)$$

が成立する. これより $e^{-tH}(x, y)$ は

$$e^{-tH}(x, y) = \int_{\mathcal{C}} e^{-\int_0^t V(B_s) ds} dP^x(\cdot | X_t = y) \quad (1.9)$$

と表せる⁴.

次にベクトルポテンシャル a がある場合を考えよう. ストーリーは上の場合と全く同じである. Schrödinger 作用素は

$$H = \frac{1}{2}(-i\nabla - a(x))^2 + V(x) \quad (1.10)$$

² (Ω, \mathcal{F}) を可測空間とする. ここで Ω は集合, \mathcal{F} は Ω の集合族で σ 代数といわれるものである. このとき $P : \mathcal{F} \rightarrow \mathbb{R}$ が測度とは (a) $0 \leq P(A) \leq \infty$, $\forall A \in \mathcal{F}$, (b) $P(\sum A_n) = \sum_n P(A_n)$, $A_n \cap A_m = \emptyset$, (c) $P(\emptyset) = 0$ をみたす集合関数のことである. $P(A) = \int_{\Omega} 1_A dP$ と表し, 一般の $\int_{\Omega} f dP$, $f \geq 0$, は階段関数 $\int_{\Omega} \sum_j a_j 1_{A_j} dP$ の極限として定義される.

³一般に『群』の定義で逆元の存在を仮定しないものを半群という.

⁴確率変数 1_A の部分 σ 代数 $\sigma(B_t)$ による条件付期待値を $P(A|B_t)$ とかく (これは確率変数になる). これは \mathbb{R}^3 上のある関数 $h(y)$ によって $P(A|B_t) = h(B_t)$ と表せることが知られている. この $h(y)$ を $P(A|B_t = y)$ で表し, さらに $P(\cdot | B_t = y)$ は測度になることが知られている (regular conditional probability measure という).

であるから

$$\mathcal{H}(p, q) = \frac{1}{2}(p - a(q))^2 + V(q) \quad (1.11)$$

とすれば, ラグランジアンは

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2}\dot{q}^2 - V(q) + a(q) \cdot \dot{q} \quad (1.12)$$

となり先ほどと同様に $e^{-itH}(x, y)$ のファインマン経路積分をつくって, $t \rightarrow -it$ の置き換えをすれば

$$e^{-tH}(x, y) = \text{定数} \times \int_{\Omega_{xy}^D} e^{-\int_0^t V(q_s)ds - i \int_0^t a(q_s)\dot{q}_s ds} e^{-\frac{1}{2} \int_0^t \dot{q}_s^2 ds} \prod_{0 \leq s \leq t} dq_s \quad (1.13)$$

となる. ベクトルポテンシャルがある場合には $-i \int_0^t a(q_s)\dot{q}_s ds$ という積分項があらわれることに注意しよう. ただし q_s は微分できないのでこの積分にはこのままでは意味がないが, 後で見るように

$$a(q_s)\dot{q}_s ds = a(q_s)dq_s \quad (1.14)$$

と書いて確率積分とよび,

$$\mathcal{C} \ni q \rightarrow \left(\int_0^t a(q_s)dq_s \right)_{t \geq 0} \quad (1.15)$$

となる \mathcal{C} 上の確率過程としてとらえることができる. 厳密には

$$e^{-tH}(x, y) = \int_{\mathcal{C}} e^{-\int_0^t V(B_s)ds - i \int_0^t a(B_s) \circ dB_s} dP^x(\cdot | X_t = y) \quad (1.16)$$

と表せる. 次の章から, スピンをもった相対論的 Schrödinger 作用素の作る半群の経路積分表示の構成を厳密に述べることにする.

2 相対論的 Schrödinger 作用素

2.1 定義

ベクトルポテンシャル $a : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ とスピン $1/2$ をもった Schrödinger 作用素は $L^2(\mathbb{R}^3; \mathbb{C}^2)$ 上の自己共役作用素

$$h = \frac{1}{2}(\sigma \cdot (p - a))^2 \quad (2.1)$$

で定義される. ここで

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

h は形式的 (定義域の議論を無視して) には

$$h = \frac{1}{2}(p - a)^2 - \frac{1}{2}\sigma \cdot b, \quad b = \nabla \times a, \quad (2.2)$$

と表せる. さて本稿の主題である相対論的 Schrödinger 作用素は次で定義される:

$$h_{\text{rel}} = \sqrt{(\sigma \cdot (p - a))^2 + m^2} - m + V. \quad (2.3)$$

$\sqrt{\cdots}$ は自己共役作用素 $(\sigma \cdot (p - a))^2 + m^2$ のスペクトル分解⁵で定義する. 相対論的 Schrödinger 作用素 h_{rel} は h をもちいて

$$h_{\text{rel}} = \sqrt{2h + m^2} - m + V \quad (2.4)$$

と表せるが, ここで関数 $\Psi(u) = \sqrt{2u + m^2} - m$ は次をみたす: (1) $\Psi \in C^\infty((0, \infty))$, (2) $\Psi(u) \geq 0$, (3) $(-1)^n \Psi^{(n)} \leq 0$, $n \geq 1$. 一般に (1) - (3) を満たす関数を Bernstein 関数という.

この論考では Ψ を任意の Bernstein 関数として次のような自己共役作用素が生成する一径数半群の経路積分表示を構成する:

$$H^\Psi = \Psi \left(\frac{1}{2}(\sigma \cdot (p - a))^2 \right) + V. \quad (2.5)$$

もちろん H^Ψ は相対論的 Schrödinger 作用素も含んでいる.

1章ではFK公式の形式的な議論をしたが, そこに現れたのは連続な経路の空間 \mathcal{C} 上の測度であった. しかし H^Ψ のような非局所的な作用素の経路積分表示を得るためには, 次章以降でみるようにポアソン過程や subordinator が必要であり, それらは一般に連続な経路もたづ, 左連続かつ右極限が存在する経路 (càdlàg path) をもつにすぎない. このような非連続な経路をもったFK公式によって半群が表せるのである.

2.2 Bernstein 関数, subordinators そして Lévy 測度

Bernstein 関数の部分集合 \mathcal{B}_0 を次で定める: $\mathcal{B}_0 = \{f \in \mathcal{B} \mid \lim_{u \rightarrow 0+} f(u) = 0\}$. \mathcal{B}_0 の典型的な例は (a) $\Psi(u) = cu^\alpha$, $c \geq 0$, $\alpha \in (0, 1]$, (b) $\Psi(u) = 1 - e^{-au}$, $a \geq 0$ などである. いま, \mathcal{L} を $\mathbb{R} \setminus \{0\}$ 上のボレル測度 λ で次を満たすもの全体とする:

$$(1) \lambda((-\infty, 0)) = 0, \quad (2) \int_{\mathbb{R} \setminus \{0\}} (y \wedge 1) \lambda(dy) < \infty.$$

$\lambda \in \mathcal{L}$ は Lévy 測度的一种である. $\mathbb{R}_+ = [0, \infty)$ とおく. さて Bernstein 関数 $\Psi \in \mathcal{B}_0$ に対して, 実は $(b, \lambda) \in \mathbb{R}_+ \times \mathcal{L}$ で次を満たすものが存在する:

$$\Psi(u) = bu + \int_0^\infty (1 - e^{-uy}) \lambda(dy). \quad (2.6)$$

⁵自己共役作用素 H に対して射影 $E_H(\lambda)$ が存在して $(f, Hg) = \int_{\sigma(H)} \lambda d(E_H(\lambda)f, g)$ と表せる. $(f, e^{-tH}g) = \int_{\sigma(H)} e^{-t\lambda} d(E_H(\lambda)f, g)$ で定義する.

逆に $(b, \lambda) \in \mathbb{R}_+ \times \mathcal{L}$ に対して (2.6) の右辺は \mathcal{B}_0 に含まれる．つまり $\mathcal{B}_0 \rightarrow \mathbb{R}_+ \times \mathcal{L}$, $\Psi \mapsto (b, \lambda)$ は 1 対 1 上への対応になっている．次に Bernstein 関数と確率過程の関係をみてみよう．確率空間 $(\Omega_\nu, \mathcal{F}_\nu, \nu^0)$ 上の確率過程 $(T_t)_{t \geq 0}$ が次を満たすとき subordinator とよばれる：(1) $(T_t)_{t \geq 0}$ は ゼロから出発する Lévy 過程, (2) T_t は t について非減少． \mathcal{S} を $(\Omega_\nu, \mathcal{F}_\nu, \nu^0)$ 上の subordinator 全体の集合とする．以降, x からスタートするパス上の確率測度 m^x に関する期待値を $\mathbb{E}_m^x[\dots] = \int \dots dm^x$ で表す． $\Psi \in \mathcal{B}_0$ とする．このとき $(T_t)_{t \geq 0} \in \mathcal{S}$ で次を満たすものが一意的に存在する：

$$\mathbb{E}_\nu^0[e^{-uT_t}] = e^{-t\Psi(u)}, \quad u \geq 0. \quad (2.7)$$

逆に $(T_t)_{t \geq 0} \in \mathcal{S}$ としよう．このとき $\Psi \in \mathcal{B}_0$ で (2.7) を満たすものが存在する．つまり $\mathcal{S}, \mathcal{B}_0, \mathbb{R}_+ \times \mathcal{L}$ はお互いに同一視することができる：

$$\mathcal{S} \cong \mathcal{B}_0 \cong \mathbb{R} \times \mathcal{L}, \quad (T_t)_{t \geq 0} \iff \Psi \iff (b, \lambda)$$

記号 $T_t^\Psi \in \mathcal{S}$ は $\Psi \in \mathcal{B}_0$ に対応する subordinator を表すものとする．

非局所的な Schrödinger 作用素の経路積分を構成するためのアイデアは以下である． $\mathbb{E}_\nu^0[e^{-T_t^\Psi u}] = e^{-t\Psi(u)}$ を右から左に読んで, u が Schrödinger 作用素 h と思えば $e^{-t\Psi(h)}$ の経路積分表示はランダムな時間 T_t^Ψ を導入して, $e^{-T_t^\Psi h}$ の T_t^Ψ に関する期待値をとることで実現できる．つまり T_t^Ψ を Bernstein 関数 $\sqrt{2u + m^2} - m$ に対応する subordinator とすれば

$$(f, e^{-t(\sqrt{-\Delta + m^2} - m)} g) = \mathbb{E}_\nu^0 \left[(f, e^{-T_t^\Psi(-\Delta/2)} g) \right].$$

さらに N 粒子系に応用しようと思えば, N 個の独立な subordinator $(T_t^j)_{t \geq 0}, j = 1, \dots, N$, によって

$$(f, e^{-t \sum_{j=1}^N (\sqrt{-\Delta_j + m^2} - m)} g) = \mathbb{E}_{\nu_1 \times \dots \times \nu_N}^0 \left[(f, e^{-\sum_{j=1}^N T_t^j(-\Delta/2)} g) \right].$$

と表せる．また外場ポテンシャルのない Schrödinger 作用素

$$H = \sqrt{(p - a(x))^2 + m^2} - m$$

の作る半群の FK 公式を導くことも即座にできる． $e^{-t \frac{1}{2}(p - a(x))^2}$ の FK 公式は第 1 章で述べたとおりであるから e^{-tH} の FK 公式は簡単に

$$(f, e^{-tH} g) = \mathbb{E}_\nu^0[(f, e^{-T_t \frac{1}{2}(p - a(x))^2} g)] = \int dx \mathbb{E}_\nu^0 \mathbb{E}_P^x \left[\overline{f(B_{T_0})} g(B_{T_t}) e^{-i \int_0^{T_t} a(B_s) \circ dB_s} \right]$$

で与えられることが分かる．

3 経路積分表示

3.1 一般化された Schrödinger 作用素

次の条件を導入する .

仮定 3.1 a の各成分 a_μ は実数値関数とする .

$$(A1) \quad a \in (L^2_{\text{loc}}(\mathbb{R}^3))^3.$$

$$(A2) \quad a \in (L^2_{\text{loc}}(\mathbb{R}^3))^3, \nabla \cdot a \in L^1_{\text{loc}}(\mathbb{R}^3).$$

$$(A3) \quad a \in (L^4_{\text{loc}}(\mathbb{R}^3))^3, \nabla \cdot a \in L^2_{\text{loc}}(\mathbb{R}^3) \text{ かつ } \nabla \times a \in (L^2_{\text{loc}}(\mathbb{R}^3))^3.$$

(A1) は Schrödinger 作用素 $h = \frac{1}{2}(\sigma \cdot (p - a))^2$ を 2 次形式で定義するのに必要な条件である . (A2) は FK 公式のために必要になる . (A3) は h の自己共役性を調べるときに役立つ .

$D_\mu = p_\mu - a_\mu$ とする . 2 次形式 q を次で定める :

$$q(f, g) = \frac{1}{2} \sum_{\mu=1}^3 (\sigma_\mu D_\mu f, \sigma_\mu D_\mu g). \quad (3.1)$$

定義域は $Q(q) = \{f \in L^2(\mathbb{R}^3; \mathbb{C}^2) \mid \sigma_\mu D_\mu f \in L^2(\mathbb{R}^3; \mathbb{C}^2), \mu = 1, 2, 3\}$. (A1) を仮定する . このとき q は非負な閉 2 次形式になるので, 次を満たす自己共役作用素 h がただ一つ存在する :

$$(hf, g) = q(f, g), \quad f \in \text{Dom}(h), g \in Q(q). \quad (3.2)$$

ここで $\text{Dom}(h) = \{f \in Q(q) \mid q(f, \cdot) \in L^2(\mathbb{R}^3; \mathbb{C}^2)'\}$.

さて滑らかな関数空間上では

$$h = \frac{1}{2}(p - a)^2 - \frac{1}{2} \begin{bmatrix} b_3 & b_1 - ib_2 \\ b_1 + ib_2 & -b_3 \end{bmatrix}, \quad (3.3)$$

である . 特に h はベクトル値関数空間上に作用する自己共役作用素なので, 例えば $h + V$ が生成する熱半群の経路積分表示をトロツタ積公式

$$e^{-t(h+V)} = s - \lim_{n \rightarrow \infty} (e^{-(t/n)h} e^{-(t/n)V})^n$$

を用いて直接求めても, 積分核はもちろん 2×2 行列の無限積のようなものになり, あまり興味がわからない . そこで経路積分表示のために h を $L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ 上の作用素へユニタリー変換する [ALS83, ARS91, HIL09] . ここで $\mathbb{Z}_2 = \{-1, 1\} = \{\theta_1, \theta_2\}$. $L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ は言うまでもなく $x \in \mathbb{R}^3$ とスピン変数 $\theta \in \mathbb{Z}_2$ 上の L^2 -関数空間である . $L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ 上に作用する作用素 $h_{\mathbb{Z}_2}$ を次で定める :

$$(h_{\mathbb{Z}_2} f)(x, \theta) = (hf)(x, \theta) - \frac{1}{2} \theta b_3(x) f(x, \theta) - \frac{1}{2} (b_1(x) - i\theta b_2(x)) f(x, -\theta). \quad (3.4)$$

命題 3.2 (A3) を仮定する . このとき $h_{\mathbb{Z}_2}$ は $\ell^2(\mathbb{Z}_2) \otimes \text{Dom}(h)$ 上で自己共役作用素で, ユニタリー作用素 $F : L^2(\mathbb{R}^3; \mathbb{C}^2) \rightarrow L^2(\mathbb{R}^3 \times \mathbb{Z}_2)$ を

$$Ff = \begin{bmatrix} f(\cdot, +1) \\ f(\cdot, -1) \end{bmatrix}$$

で定義すれば $FhF^{-1} = h_{\mathbb{Z}_2}$ が従う .

$h_{\mathbb{Z}_2}$ を $\mathbb{Z}_2 \rightarrow \mathbb{Z}_p$ へ一般化する . \mathbb{Z}_p を 1 の原始 p 乗根の集合とする :

$$\mathbb{Z}_p = \{\theta_1^{(p)}, \dots, \theta_p^{(p)}\}, \quad \theta_\alpha^{(p)} = \exp\left(2\pi i \frac{\alpha}{p}\right). \quad (3.5)$$

以下で $p \geq 2$ を固定し $\theta_\beta^{(p)}$ を簡単に θ_β と書くことにする . $\ell^2(\mathbb{Z}_p) = \{f | \mathbb{Z}_p \rightarrow \mathbb{C}\}$ にスカラー積 $(f, g)_{\ell^2(\mathbb{Z}_p)} = \sum_{\beta=1}^p \overline{f(\theta_\beta)} g(\theta_\beta)$ を導入する .

スピン \mathbb{Z}_p をもった Schrödinger 作用素 を定義しよう .

定義 3.3 (1) (対角部分) $U : \mathbb{R}^3 \times \mathbb{Z}_p \rightarrow \mathbb{R}$ は $\max_{\theta \in \mathbb{Z}_p} |U(x, \theta)|$ が $\frac{1}{2}p^2$ に相対有界な掛け算作用素 .

(2) (非対角部分) $W_\beta : \mathbb{R}^3 \times \mathbb{Z}_p \rightarrow \mathbb{C}$, $1 \leq \beta \leq p-1$, は $\max_{\theta \in \mathbb{Z}_p} |W_\beta(x, \theta)|$ が $\frac{1}{2}p^2$ に相対有界となる掛け算作用素 . $U_\beta : \mathbb{R}^3 \times \mathbb{Z}_p \rightarrow \mathbb{C}$ を次で定める :

$$U_\beta(x, \theta_\alpha) = \frac{1}{2} \left(W_\beta(x, \theta_{\alpha+\beta}) + \overline{W_{p-\beta}(x, \theta_\alpha)} \right). \quad (3.6)$$

(3) (一般化されたスピン作用素) $M_{\mathbb{Z}_p} : L^2(\mathbb{R}^3 \times \mathbb{Z}_p) \rightarrow L^2(\mathbb{R}^3 \times \mathbb{Z}_p)$,

$$M_{\mathbb{Z}_p} : f(x, \theta_\alpha) \mapsto U(x, \theta_\alpha) f(x, \theta_\alpha) + \sum_{\beta=1}^{p-1} U_\beta(x, \theta_\alpha) f(x, \theta_{\alpha+\beta}) \quad (3.7)$$

を一般化されたスピン作用素 とよぶ .

$$h_{\mathbb{Z}_p} = \frac{1}{2}(p-a)^2 + M_{\mathbb{Z}_p} \quad (3.8)$$

とおく .

例 3.4 (スピン 1/2) $p = 2$ として $W_1(x, \theta) = -\frac{1}{2}(b_1(x) + i\theta b_2(x))$, $\theta \in \mathbb{Z}_2$ と定めれば, これはまさに (3.4) で与えたものに他ならない .

以下で $u_\beta(x) = \begin{cases} \max_{\theta \in \mathbb{Z}_p} |U(x, \theta)| & \text{if } \beta = p, \\ \max_{\theta \in \mathbb{Z}_p} |U_\beta(x, \theta)| & \text{if } 1 \leq \beta \leq p-1 \end{cases}$ とおく .

$$\|u_\beta f\| \leq c_\beta \|\frac{1}{2}p^2 f\| + b_\beta \|f\| \quad (3.9)$$

となる $c_\beta > 0$ と $b_\beta \geq 0$ が存在する .

定理 3.5 (A2) を仮定する . $\sum_{\beta=1}^p c_\beta < 1$ としよう . このとき $h_{\mathbb{Z}_p}$ は $\ell^2(\mathbb{Z}_p) \otimes \text{Dom}(h)$ 上で自己共役 で下から有界である .

自己共役作用素 H のスペクトルの下限は基底状態エネルギーといわれる . $E(H) = \inf \sigma(H)$ と表すことにする .

定義 3.6 (A2) を仮定する . $\sum_{\beta=1}^p c_\beta < 1$, $\Psi \in \mathcal{B}_0$ とし $\overline{h_{\mathbb{Z}_p}} = h_{\mathbb{Z}_p} - E(h_{\mathbb{Z}_p})$ とおく . このとき

$$H_{\mathbb{Z}_p}^\Psi = \Psi(\overline{h_{\mathbb{Z}_p}}) + V \quad (3.10)$$

を一般化された Schrödinger 作用素という .

Ψ は \mathbb{R}_+ 上で定義された関数なので $\Psi(h_{\mathbb{Z}_p})$ は一般に定義されない . そこで $\Psi(\overline{h_{\mathbb{Z}_p}})$ とした .

命題 3.7 (A2) を仮定し $\sum_{\beta=1}^p c_\beta < 1$ とする . もし $\Psi \in \mathcal{B}_0$ ならば, $\ell^2(\mathbb{Z}_p) \otimes C_0^\infty(\mathbb{R}^3)$ は $\Psi(\overline{h_{\mathbb{Z}_p}})$ の芯である .

3.2 FK 公式

ここでは以下の 3 つの確率過程の組み合わせによって e^{-tH^Ψ} の経路積分表示を構成する . 詳しくは [HIL09] または [Hir09, Hir10a, HL08b] を参照せよ .

$$\begin{aligned} (B_t)_{t \geq 0} & \text{ ブラウン運動} & \implies & -\Delta/2 \text{ に対応する確率過程} \\ (-1)^{N_t} & \text{ ポアソン過程 } (N_t)_{t \geq 0} & \implies & \text{ スピンに対応する確率過程} \\ (T_t^\Psi)_{t \geq 0} & \text{ Subordinator} & \implies & \Psi(\cdot) \text{ に対応する確率過程} \end{aligned}$$

$(N_t^\beta)_{t \geq 0}$, $\beta = 1, \dots, p-1$, は $(\Omega_N, \mathcal{F}_N, \mu)$ 上の $p-1$ 個の独立なポアソン過程で intensity 1 をもつとする . 即ち $\mu(N_t^\beta = n) = e^{-t} t^n / n!$. 確率過程 $(N_t)_{t \geq 0}$ を

$$N_t = \sum_{\beta=1}^{p-1} \beta N_t^\beta \quad (3.11)$$

によって定める .

$$\int_v^{w+} g(N_{s-}) dN_s^\beta = \sum_{\substack{v \leq r \leq w \\ N_{r+}^\beta \neq N_{r-}^\beta}} g(N_{r-}) \quad (3.12)$$

と定義する . 即座に $\mathbb{E}_\mu \left[\int_v^{w+} g(N_{s-}) dN_s^\beta \right] = \mathbb{E}_\mu \left[\int_v^w g(N_s) ds \right]$ がわかる .

ここでセミマルチンゲールに関する伊藤の公式を復習しておこう . フィルトレーションを $\mathcal{F}_t = \sigma((B_s, N_s^\beta), 0 \leq s \leq t, \beta = 1, \dots, p)$ とする . 次を考える

$$L_t^i = \int_0^t f^i(s, \omega) ds + \int g^i(s, \omega) \cdot dB_s + \sum_{\beta=1}^{p-1} \int_0^{t+} h_\beta^i(s, \omega) dN_s^\beta, \quad i = 1, \dots, n.$$

ここで⁶ $f^i(\cdot, \omega) \in L^1_{\text{loc}}(\mathbb{R})$ a.s., $g^i \in \mathcal{E}_{\text{loc}}$. そして $h^i_\beta(s, \omega)$ は \mathcal{F}_t に関して適合していて, s について左連続で $\int_0^{t+} |h^i_\beta(s, \omega)| dN_s^\beta < \infty$ a.s. $F \in C^2(\mathbb{R}^n)$ のとき

$$\begin{aligned} F(L_t) - F(L_0) &= \sum_{i=1}^n \int_0^t F_i(L_s) f^i(s) ds + \sum_{i,j=1}^n \int_0^t \frac{1}{2} F_{ij}(L_s) g^i(s) \cdot g^j(s) ds \\ &\quad + \sum_{i=1}^n \int_0^t F_i(L_s) g^i(s) \cdot dB_s + \sum_{\beta=1}^{p-1} \int_0^{t+} (F(L_{s-} + h_\beta(s)) - F(L_{s-})) dN_s^\beta \end{aligned}$$

となる. ここで $F_i = \partial_i F$, $F_{ij} = \partial_i \partial_j F$. この公式は簡潔に

$$dF(L_t) = F_i dL_t^i + \frac{1}{2} F_{ij} dL_t^i dL_t^j$$

と表される.

$\mathbb{R}^3 \times \mathbb{Z}_2$ -値確率過程 Q_t を $Q_t = (B_t, \theta_{N_t})$ で定める.

補題 3.8 (スピンをもった Schrödinger 作用素の FK 公式) $\sum_{\beta=1}^p c_\beta < 1$ とする. (A2) を仮定し

$$\int_0^t ds \int_{\mathbb{R}^3} dy (2\pi s)^{-d/2} e^{-|x-y|^2/(2s)} |\log u_\beta(y)| < \infty \quad (3.13)$$

としよう. このとき

$$(f, e^{-th_{\mathbb{Z}_p}} g) = e^{(p-1)t} \sum_{\alpha=1}^p \int_{\mathbb{R}^3} dx \mathbb{E}_{P \times \mu}^{x, \alpha} \left[\overline{f(Q_0)} g(Q_t) e^S \right] \quad (3.14)$$

が従う. ただし $S = S_a + S_{\text{spin}}$ で,

$$S_a = -i \int_0^t a(B_s) \circ dB_s,$$

$$S_{\text{spin}} = - \int_0^t U(Q_s) ds + \sum_{\beta=1}^{p-1} \int_0^{t+} \log(-U_\beta(Q_{s-})) dN_s^\beta.$$

ここで $\log z$ は主値をとるものと約束する.

証明: 証明の概略を述べる. (3.14) の右辺を $(f, S_t g)$ とおいて有界作用素 S_t を $t \geq 0$ ごとに定義する. ブラウン運動, ポアソン過程, subordinator のマルコフ性をつかって S_t が強連続一径数半群の性質を満たすことを示す. そうすると Stone の定理よりある自己共役作用素 H で $S_t = e^{-tH}$ と書けることが分かる. さらにセミマルチンゲールに関する伊藤の公式を用いて $-H = \lim_{t \rightarrow 0} (S_t f - f)/t$ を計算して $H = h_{\mathbb{Z}_p}$ となることを示す. 実際 $S_t f(x, \alpha) = \mathbb{E}_{P \times \mu}^{x, \alpha} [f(Q_t) e^S]$ となり,

$$S_t f - S_0 f = \mathbb{E}_{P \times \mu}^{x, \alpha} [f(Q_t) e^S - f(Q_0)]$$

を計算するために次の等式⁷

$$f(Q_t) e^S - f(Q_0) = df \cdot e^S + f \cdot de^S + df \cdot de^S$$

と df, de^S に伊藤の公式を当てはめる.

終り

⁶ $f \in \mathcal{E}_{\text{loc}}$ とは $P^x(\int_0^t |f(B_s)|^2 ds < \infty) = 1$ となること.

⁷ 伊藤の積公式という.

次が主定理である.

定理 3.9 (一般化された Schrödinger 作用素の FK 公式) $\Psi \in \mathcal{B}_0$ としよう. T_t^Ψ の \mathbb{R} 上の分布を $\rho(r, t)$ とする. $u_\beta \in L^\infty(\mathbb{R}^3)$, $V \in L^\infty(\mathbb{R}^3)$ とし, (A2) を仮定する. さらに

$$\int_{\mathbb{R}} dr \rho(r, t) \int_0^r ds \int_{\mathbb{R}^3} dy (2\pi s)^{-d/2} e^{-|x-y|^2/(2s)} |\log u_\beta(y)| < \infty \quad (3.15)$$

が成立しているとする. このとき

$$(f, e^{-tH_{\mathbb{Z}_p}^\Psi} g) = \sum_{\alpha=1}^p \int_{\mathbb{R}^3} dx \mathbb{E}_{P \times \mu \times \nu}^{x, \alpha, 0} \left[e^{(p-1)T_t^\Psi} \overline{f(Q_0)} g(Q_{T_t^\Psi}) e^{S^\Psi} \right]. \quad (3.16)$$

ここで $S^\Psi = S_V^\Psi + S_a^\Psi + S_{\text{spin}}^\Psi$ は $S_V^\Psi = -\int_0^t V(B_{T_s^\Psi}) - E(h_{\mathbb{Z}_p}) ds$, S_a^Ψ , S_{spin}^Ψ は S_a , S_{spin} で t に T_t^Ψ を代入したもの.

証明: はじめに $V = 0$ とおく. このとき

$$(f, e^{-tH_{\mathbb{Z}_p}^\Psi} g) = \mathbb{E}_\nu^0[(f, e^{-th_{\mathbb{Z}_p}} g)]$$

なので, 補題 3.8 により FK 公式をえる. $V \neq 0$ のときはトロツタ積公式による. 終り

$h_{\mathbb{Z}_p}^0$ を $h_{\mathbb{Z}_p}$ で a と U_β を各々 0 と $|U_\beta|$ に置き換えたものとして定義する. 定理 3.9 から直ちに次の系が従う.

系 3.10 (双極不等式) 定理 3.9 の条件を仮定する. このとき $h_{\mathbb{Z}_p} - E(h_{\mathbb{Z}_p}^0) \geq 0$ となり, さらに次の不等式が成り立つ:

$$\left| (f, e^{-t(\Psi(h_{\mathbb{Z}_p} - E(h_{\mathbb{Z}_p}^0)) + V)} g) \right| \leq (|f|, e^{-t(\Psi(\overline{h_{\mathbb{Z}_p}^0}) + V)} |g|). \quad (3.17)$$

4 場の理論など

前章では Schrödinger 作用素の FK 公式を与えるという事実のみを概説したが, この表示を利用して Schrödinger 作用素のスペクトルも解析することができる. Schrödinger 作用素の固有ベクトルの減衰性, 一意性, 固有値の個数や大きさの評価など. また半群の hypercontractivity やその積分核の連続性なども FK の公式から導くことができる [Sim04]. これらのことは場の量子論でも同様に経路積分 / 汎関数積分を用いて示せることが近年分かってきた [LHB11]. 場の量子論のスペクトルの解析は Schrödinger 作用素のそれと大きく異なり, 非常に困難である. その理由は固有値が連続固有値に埋め込まれていることにある. さらに threshold になっているために complex dilatation がうまく働かない. そのために離散固有値の摂動理論が使えず, 非摂動的な方法が必要になる. そのひとつの有力な武器が汎関数積分や作用素論的なくりこみ群である. 例えば近年は相対論的 Pauli-Fierz 模型と言われる, 相対論的な Schrödinger 粒子が輻射

場とミニマル結合した模型のスペクトルが盛んに調べられている．それは次のように $L^2(\mathbb{R}^3) \otimes \mathcal{F}$ 上に定義されるものである:

$$\sqrt{(p \otimes \mathbb{1} + \sqrt{\alpha} A_{\text{rad}}(x))^2 + M^2} - M + V \otimes \mathbb{1} + \mathbb{1} \otimes H_f.$$

\mathcal{F} はボゾンフォック空間, $A_{\text{rad}}(x)$ は輻射場

$$A_{\text{rad}}(x) = \frac{1}{\sqrt{2}} \sum_{j=1}^{d-1} \int \left(a^\dagger(k, j) \frac{\hat{\varphi}(k)e(k, j)}{\sqrt{|k|}} e^{-ikx} + a(k, j) \frac{\hat{\varphi}(k)e(k, j)}{\sqrt{|k|}} e^{+ikx} \right) dk,$$

$\hat{\varphi}$ はカットオフ関数, $e(k, 1), \dots, e(k, d-1)$ は極ベクトル, H_f は自由ハミルトニアン

$$H_f = \sum_{j=1}^{d-1} \int |k| a^\dagger(k, j) a(k, j) dk$$

である. a, a^\dagger は $[a(k, j), a^\dagger(k', j')] = \delta(k - k') \delta_{jj'}$ を満たす生成消滅作用素である. 場の理論に現れる非局所的 (non-local) な作用素を解析するのは一般には非常に困難であるが, 汎関数積分表示を用いれば, ある部分で局所的な作用素と同様の解析ができるという有利な点がある. 詳しくは [Hir10b, HL08a, LHB11]などを参照せよ.

参考文献

- [ALS83] G. F. De Angelis, G. Jona-Lasinio, and M. Sirugue, Probabilistic solution of Pauli-type equations, *J. Phys.* **A16** (1983), 2433–2444.
- [ARS91] G. F. De Angelis, A. Rinaldi, and M. Serva, Imaginary-time path integral for a 相対論的 spin-(1/2) particle in a magnetic field, *Europhys. Lett.* **14** (1991), 95–100.
- [Hir09] F. Hiroshima, Path integrations of relativistic Schrödinger operators and Bernstein functions, *京大数理研講究録* **1658**, (2009), 18–34.
- [Hir10a] F. Hiroshima, Feynman-Kac type formulas for Schrödinger semigroup with Bernstein functions of Laplacian, *京大数理研講究録* **1696**, *スペクトル散乱理論とその周辺* (2010), 119–143.
- [Hir10b] F. Hiroshima, Relativistic Pauli-Fierz model, COE Lecture Note (Kyushu University) **30** (2011), 145–161.
- [HIL09] F. Hiroshima, T. Ichinose, and J. Lőrinczi, J. Path integral representation of Schrödinger operator with Bernstein functions of the Laplacian, preprint 2009.
- [HL08a] F. Hiroshima and J. Lőrinczi, Functional integral representation of the Pauli-Fierz model with spin 1/2, *J. Funct. Anal.* **254** (2008), 2127–2185.
- [HL08b] F. Hiroshima and J. Lőrinczi, Functional integral representations of nonrelativistic QED, *京大数理研講究録* **1600** (2008), 68–91.
- [LHB11] J. Lőrinczi, F. Hiroshima and V. Betz, Feynman-Kac-Type Theorems and Gibbs Measures on Path Space. With Applications into Rigorous Quantum Field Theory to appear in *de Gruyter Studies in Mathematics* **34**.
- [Sim04] B. Simon, *Functional integration and quantum physics*, AMS Chelsea Publishing, 2004.

On Robust Quantum Control

近畿大学 量子コンピュータ研究センター	坂東 将光 ¹
近畿大学 量子コンピュータ研究センター	市川 翼
近畿大学理工学部	近藤 康
近畿大学理工学部	中原 幹夫

1 はじめに

本報告書ではロバストな量子ゲートを設計する為の、体系的な枠組みの基礎を与える。量子ゲートは量子コンピュータに於ける基本素子であり、時間発展演算子で実現される。従って量子コンピュータの実現には、正確な量子制御が必要不可欠となる。エラーに対してロバストな量子ゲートいかに設計するか、ということは重要な問題である。

背景 NMR の分野では、複合パルスと呼ばれるものがしばしば用いられている [1]。これはエラーの影響を受けにくくするため、ある 1 つのスピ操作をそのまま行うのではなく、複数の単純なスピ操作を組み合わせて行うというものである。スピ操作は時間発展によって行われるので、複合スピ操作は量子ゲートにおいても有効であると期待される。つまり本来 1 つの量子ゲートを、複数の量子ゲートを用いて作ることで、エラーの影響を受けにくい量子ゲートを実現することができよう。そのような量子ゲートを複合量子ゲートと呼ぶ。これまでに幾つか NMR 量子コンピュータを念頭に、ロバストな 1 量子ビットの複合量子ゲートが設計されている [2][3]。しかし、それらの複合量子ゲートの設計においてはエラーの具体形に依存しており、より一般的な、任意のエラーに対応可能な枠組みを構成しようという試みは決して多くはない。

我々は、統一的な複合量子ゲートの設計手法を作ることを目的としている。一方、これまでの研究で我々は、pulse strength エラーと呼ばれるエラーに対してロバストな複合量子ゲートを設計する為には、動的位相 [4] が 0 になるように量子ゲートを設計すれば良いことを示した [5]。これは全ての幾何学的位相ゲートが pulse strength エラーに対してロバストであることも示している。pulse strength エラーとは制御パラメータの大きさへの系統的なエラーのことで、off-resonance エラーと共に NMR 量子コンピュータに於いて特に重要なエラーある。

本報告書ではエラーを摂動的に取り扱い、“エラーの一次の項を打ち消すことが可能な複合量子ゲート”を設計する為の体系的な枠組みを与える。この体系は任意のエラーに対して適用可能であり、我々のものを含む諸々の先行研究を自然に導出できることを示す。

¹E-mail: bando@alice.math.kindai.ac.jp

2 量子ゲートとロバスト性

2.1 ロバストな量子ゲート

この節は量子ゲートの概観を与える。量子ゲートとは時間発展演算子である。一量子ビットゲートは $SU(2)$ 、 N 量子ビットゲートは $SU(2^N)$ で表される。ここでは時間発展演算子 $U_\lambda(t, 0) \in SU(n)$ を考える。この時間発展演算子を

$$U_\lambda(t, 0) := \mathcal{T} \exp \left[-i \int_0^t ds H(\lambda(s)) \right] \quad (1)$$

と書こう。ここで t は時間、

$$H(\lambda(t)) := \lambda_\mu(t) \tau_\mu \quad (2)$$

はハミルトニアン、 $\lambda(t) = (\lambda_1(t), \dots, \lambda_{n^2-1}(t))$ は $n^2 - 1$ 次元のパラメータ多様体 \mathcal{M} における座標、 τ_μ は $n^2 - 1$ 次元直交基底で $\mu = 1, \dots, n^2 - 1$ であり、 $\hbar = 1$ としている。例えば $n = 2$ では、 τ_μ は $SU(2)$ の生成元 $\tau_\mu = \sigma_\mu/2$ となる。ここで σ_μ はパウリ行列である。さて今、 $t \in [0, 1]$ とし、

$$U := U_\lambda(1, 0) \quad (3)$$

が我々の求める量子ゲートを表すとしよう。我々は量子ゲート U を得たいが、実際には $\lambda(t)$ に対して何らかのエラー $\delta\lambda(t)$ があり、望む量子ゲートそのものを得ることができない。しかし上手くパラメータ $\lambda(t)$ を選ぶことによって、大きさのわからない、与えられたエラーの1次の寄りを打ち消すことができる場合がある。このとき、つまり

$$U_{\lambda+\delta\lambda}(1, 0) = U + \mathcal{O}(|\delta\lambda|^2) \quad (4)$$

となるとき、そのような $\lambda(t)$ によって得られる U をロバストな量子ゲートと呼ぶ。

2.2 ロバストな量子ゲートの条件

式 (4) を満たすロバストな量子ゲートを設計する為に、その左辺を相互作用描像を用いて書き直すと、

$$U_{\lambda+\delta\lambda}(1, 0) = U U_{\delta\lambda}^I(1, 0) \quad (5)$$

となる。ここで

$$U_{\delta\lambda}^I := \mathcal{T} \exp \left[-i \int_0^1 dt H_I(\delta\lambda(t)) \right] \quad (6)$$

であり、

$$H_I(\delta\lambda(t)) := U_\lambda(t, 0)^\dagger H(\delta\lambda(t)) U_\lambda(t, 0) \quad (7)$$

は相互作用描像におけるハミルトニアンである。また式 (5) の右辺は

$$U_{\lambda+\delta\lambda}(1,0) = U - iU\Delta W + \mathcal{O}(|\delta\lambda|^2) \quad (8)$$

と展開することができる。ここで

$$\Delta W := \int_0^1 dt H_I(\delta\lambda(t)) \quad (9)$$

である。式 (4) と式 (8) を比べることで、ロバストな量子ゲートを設計するための条件は

$$\Delta W = 0 \quad (10)$$

であることがわかる。これは一般のエラー $\delta\lambda(t)$ に対して量子ゲートがロバストである条件である。次節からは系統的なエラーに焦点を当て、この条件をどのように理解すれば良いのかを調べる。

2.3 エラーの分類

以降は簡単のために、誤解のない範囲で、 $\lambda(t)$ など時間に依存する関数を、単に λ のように書くことにする。 $\delta\lambda$ が λ の系統的なエラーである場合、

$$\delta\lambda_\mu = F_\mu(\lambda) \quad (11)$$

と書くことができる。ここで、 $F = (F_1, \dots, F_{n^2-1})$ は未知の、しかし決定論的な関数である。 $F_\mu(\lambda)$ を

$$F_\mu(\lambda(t)) = f_\mu + f_{\mu\nu}\lambda_\nu + f_{\mu\nu\rho}\lambda_\nu\lambda_\rho \quad (12)$$

と展開する。 $f_\mu, f_{\mu\nu}, f_{\mu\nu\rho}, \dots$ は定テンソルである。

$$\int_0^1 dt \tilde{\tau}_\mu = 0, \int_0^1 dt \tilde{\tau}_\mu \lambda_\nu = 0, \int_0^1 dt \tilde{\tau}_\mu \lambda_\nu \lambda_\rho = 0, \dots \quad (13)$$

が成り立つとき、式 (10) は任意の F に対して成立する。ここで、 $\tilde{\tau}_\mu(t)$ は相互作用描像におけるジェネレーター

$$\tilde{\tau}_\mu(t) := U_\lambda(t,0)^\dagger \tau_\mu U_\lambda(t,0) \quad (14)$$

である。式 (13) を全て満たすと、その量子ゲートはあらゆる系統的なエラーに対してロバストになるが、それは現実的ではない。大抵の物理的に意味のある系統的なエラーは、式 (13) の2つ目の条件に含まれている。以下ではこの式 (13) の2つ目の条件式に注目する。 $f_{\mu\nu}$ を

$$f_{\mu\nu} = \delta_{\mu\nu} \frac{f_{\rho\rho}}{N} + \frac{f_{\mu\nu} - f_{\nu\mu}}{2} + \left[\frac{f_{\mu\nu} + f_{\nu\mu}}{2} - \delta_{\mu\nu} \frac{f_{\rho\rho}}{N} \right] \quad (15)$$

と展開することによって、この条件は

$$\int_0^1 dt \tilde{\tau}_\mu \lambda_\mu = 0 \quad (16a)$$

$$\int_0^1 dt (\tilde{\tau}_\mu \lambda_\nu - \tilde{\tau}_\nu \lambda_\mu) = 0 \quad (16b)$$

$$\int_0^1 dt \left[\frac{\tilde{\tau}_\mu \lambda_\nu + \tilde{\tau}_\nu \lambda_\mu}{2} - \delta_{\mu\nu} \frac{\tilde{\tau}_\rho \lambda_\rho}{N} \right] = 0 \quad (16c)$$

と3つの条件に分けることができる。ここで1つ目の条件(16a)はパラメータ $\lambda(t)$ のノルムに対するエラーに対して、ゲートがロバストになる条件であることが分かる。またこの条件式は式(14)から

$$\int_0^1 dt U_\lambda(t, 0)^\dagger H(\lambda(t)) U_\lambda(t, 0) = 0 \quad (17)$$

と書くことができる。これは $|\psi(t)\rangle = U_\lambda(t, 0) |\psi(0)\rangle$ として $|\psi(0)\rangle$ で期待値をとると

$$- \int_0^1 dt \langle \psi(t) | H(\lambda(t)) | \psi(t) \rangle = 0 \quad (18)$$

となる。この式(18)の左辺は動的位相[4]の定義そのものである。つまり1つ目の条件式(16a)は、動的位相が0になるように構成された量子ゲートは、 λ のノルムに対するエラーに対してロバストであるということを示している。これは我々が以前の研究で明らかにしたことである[5]。2つ目の条件(16b)は、パラメータ λ が、パラメータ空間で回転をするようなエラーに対して量子ゲートがロバストな条件である。また3つ目の条件(16c)は同じくパラメータ λ に対する応力のようなエラーに対してロバストな条件である。

3 ロバストな複合量子ゲートの設計

前章では、様々なエラーに対して量子ゲートがロバストな条件を導出した。実際にこの条件を用いて量子ゲートを設計するには、複合量子ゲート、つまり離散化して複数の単純な量子ゲートの組合せとして設計する方法が簡単である。この章では、ゲートがロバストである条件(10)を、離散化した時間発展演算子に対して適用する。

3.1 離散化

まず、ターゲットとなるユニタリ行列 U を

$$U = U_{\lambda^k}(t_k, t_{k-1}) \dots U_{\lambda^1}(t_1, t_0) \quad (19)$$

と離散化する。ここで $1 = t_k > t_{k-1} > \dots > t_1 > t_0 = 0$ であり、 λ^i は i 番目のユニタリ行列のパラメータである。各々のユニタリ行列を

$$U_{\lambda^j}(t_j, t_{j-1}) \rightarrow R(m^j) := \exp(-im_\mu^j \tau_\mu) \quad (20)$$

と書き直すことにする。今、エラーのあるユニタリ一行列 $U_{\lambda+\delta\lambda}$ を

$$U_{\lambda+\delta\lambda} = U_{\lambda^k+\delta\lambda^k}(t_k, t_{k-1}) \dots U_{\lambda^1+\delta\lambda^1}(t_1, t_0) \quad (21)$$

と離散化できるとしよう。エラーの無いときと同様に、各々のユニタリ一行列を

$$U_{\lambda^j+\delta\lambda^j}(t_j, t_{j-1}) \rightarrow R(m^j + \delta m^j) := \exp[-i(m_\mu^j + \delta m_\mu^j)\tau_\mu] \quad (22)$$

と書き直す。ここで、

$$W^j := m_\mu^j \tau_\mu, \quad \delta W^j := \delta m_\mu^j \tau_\mu \quad (23)$$

とし、2つの行列 A, B に関するよく知られた関係式

$$e^{A+B} = \int_0^1 dx \delta(1-x) e^{Ax} + \int_0^1 dx \int_0^1 dy \delta(1-x-y) e^{Ax} B e^{Ay} + \dots \quad (24)$$

を用いると、

$$R(m^j + \delta m^j) \approx R(m^j) - iR(m^j)\delta W_I^j \quad (25)$$

となる。ここで

$$\delta W_I^j := \int_0^1 dx e^{ixW^j} \delta H^j e^{-ixW^j} \quad (26)$$

である。よって、(8) 式で定義した ΔW は

$$\Delta W = \sum_{j=1}^k V^{j-1\dagger} \delta W_I^j V^{j-1} \quad (27)$$

となる。ここで

$$V^j := R(m^j) \dots R(m^1) \quad (28)$$

であり、 $R(m^k) \dots R(m^j) = UV^{j-1\dagger}$ の関係を使った。この ΔW が 0 になるように m^j を決定することで、ロバストな量子ゲートを設計することができる。

3.2 例

例として、前節までの方法を用いて $n=2$ 、 $k=3$ での、off-resonance エラーに対してロバストな量子ゲートを設計する。簡単の為、

$$R(m^1) = R(m^3), \quad m^j = \theta_j(\cos \phi_j, \sin \phi_j, 0) \quad (29)$$

の場合を考えよう。 $m_3^j = 0$ という制限は、NMR 量子コンピュータの制限と同じである。設計したい量子ゲートを

$$R(m) := R(m^3)R(m^2)R(m^1), \quad m = \theta(\cos \phi, \sin \phi, 0) \quad (30)$$

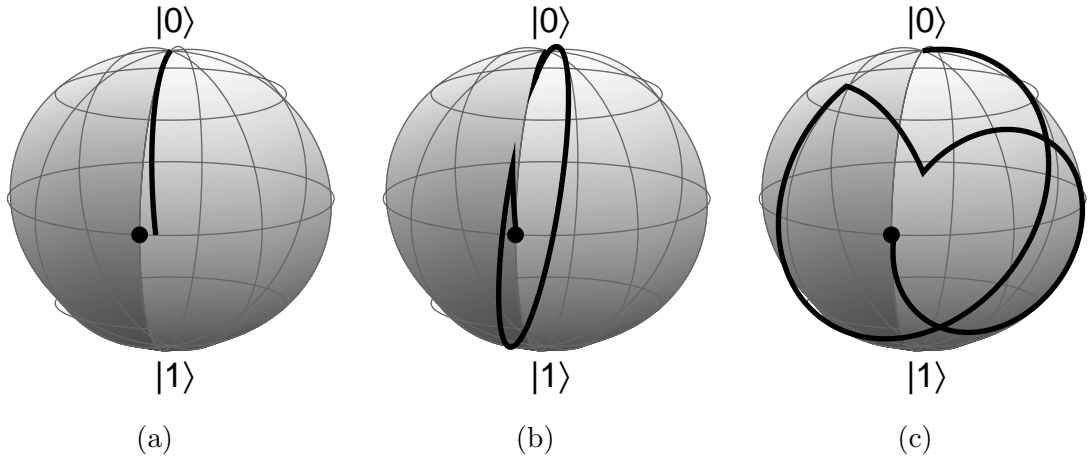


図 1: 量子ゲートによる状態の時間発展を Bloch 球上の軌跡で表したものである。(a) シンプルな量子ゲート、(b) CORPSE 及び (c) (37) 式の新しい条件から設計した複合量子ゲート。何れも $\theta = \pi/2$ 、off-resonance エラー $\varepsilon = 0.1$ で、時刻 $t=0$ では σ_3 の $+1$ の固有値に対応する固有状態 $|0\rangle$ をとる。黒点はエラーの無い、理想的な量子ゲートを用いた場合に於ける量子状態。(b)、(c) の複合量子ゲートは off-resonance エラーに対してロバストであることがわかる。

と書く。off-resonance エラーとは m^j が、

$$m^j = \theta_j(\cos \phi_j, \sin \phi_j, 0) \longrightarrow \theta_j(\cos \phi_j, \sin \phi_j, \varepsilon) \quad (31)$$

となるエラーである。 $R(m^j) = \exp(-im_\mu^j \sigma_\mu / 2)$ であったことを思い出すと、(27) 式からただちにエラーの一次の項は

$$U\Delta W = \varepsilon \left[\sin \frac{\theta_1}{2} R(m^1) R(m^2) \sigma_3 + \sin \frac{\theta_1}{2} \sigma_3 R(m^2) R(m^1) + \sin \frac{\theta_2}{2} R(m^1) \sigma_3 R(m^1) \right] \quad (32)$$

となることが分かる。 $m_3^j = 0$ から

$$R(m^j) \sigma_3 = \sigma_3 R(m^j)^\dagger \quad (33)$$

の関係があるので結局、ロバストな量子ゲートの条件は

$$2 \sin \frac{\theta_1}{2} \left[\cos \frac{\theta_1}{2} \cos \frac{\theta_2}{2} - \sin \frac{\theta_1}{2} \sin \frac{\theta_2}{2} \cos(\phi_1 - \phi_2) \right] + \sin \frac{\theta_2}{2} = 0 \quad (34)$$

となる。 $\phi_1 = \phi_2$ や $\phi_1 = \phi_2 + \pi$ 、 $\phi_1 = \phi_2 + \pi/2$ のとき、(34) 式は解くことができ、

$$\theta_1 + \theta_2 + \theta_3 = 2\theta_1 + \theta_2 = 0, \quad \phi_1 = \phi \quad (\phi_1 = \phi_2) \quad (35)$$

$$\theta_1 = \frac{\theta}{2} - k, \quad \theta_2 = -2k, \quad k := \arcsin \left(\frac{1}{2} \sin \frac{\theta}{2} \right), \quad \phi_1 = \phi \quad (\phi_1 = \phi_2 + \pi) \quad (36)$$

$$\theta_1 = -\arcsin \left(\tan \frac{\theta_2}{2} \right), \quad \theta_2 = 2 \arcsin \left(\frac{1}{\sqrt{2}} \sin \frac{\theta}{2} \right), \quad \phi_1 = \phi + \frac{3}{4}\pi \quad (\phi_1 = \phi_2 + \pi/2) \quad (37)$$

となる。 θ_1 、 θ_2 に関する条件からは mod 2π を省略している。(35) 式の条件を満たすものには、Bang-Bang コントロールや MLEV-4、MLEV-8 といった NMR でよく使われているものが含ま

れる。また (36) 式の条件は、既に off-resonance エラーに対してロバストな複合量子ゲートとして知られている CORPSE [3] の条件である。(37) 式の条件は、これまでに知られていない新しい条件である。図 (1) は $\theta = \pi/2$ 、 $\varepsilon = 0.1$ のときのシンプルな量子ゲート、CORPSE、そして (37) 式の新しい条件から設計した複合量子ゲートによる量子状態の時間発展を表したものである。ロバストな量子ゲートを設計する為には、このように (27) 式からエラーの一次の項をまず計算し、これが 0 になるように各々のパラメータを決定すれば良い。

4 まとめ

ロバストな量子ゲートを設計するための体系的な枠組みを作成した。さらに離散化することで、その条件を用いて 1 量子ビットにおける複合量子ゲートを設計できることを示した。これは NMR 量子コンピュータで既に知られている CORPSE などの複合量子ゲートも設計可能な、より一般的な複合量子ゲートの設計方法である。

参考文献

- [1] T. D. W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, (Elsevier, Oxford, 1999).
- [2] H. K. Cummins and J. A. Jones: *New J. Phys.* **2** (2000) 6; H. K. Cummins and J. A. Jones: *J. Magn. Reson.* **148** (2001) 338.
- [3] H. K. Cummins, G. Llewellyn, and J. A. Jones: *Phys. Rev. A* **67** (2003) 042308; W. G. Alway, and J. A. Jones: *J. Magn. Reson.* **189** (2007) 114.
- [4] Y. Aharonov and J. Anandan: *Phys. Rev. Lett.* **58** (1987) 1593; D. N. Page: *Phys. Rev. A* **36** (1987) 3479.
- [5] Y. Kondo and M. Bando, *J. Phys. Soc. Jpn* **80** (2011) 054002.

非インスタントトンネル効果：半古典論および量子論的解釈

高橋公也（九州工業大学大学院 情報工学研究院）， 池田研介（立命館大学 理工学部）

平成 23 年 5 月 1 日

概要

周期摂動の入ったステップポテンシャルではインスタント型のトンネル効果が実質的に禁止される。この場合に観測される非インスタントトンネル効果の半古典的および量子論的な解釈を試みた。摂動周波数を変化させると、トンネル確率は、 $\omega \rightarrow 0$ と $\omega \rightarrow \infty$ の極限でゼロになるが、その中間領域で著しく大きな値を取り、その値は摂動強度とともに増大する。 $\hbar\omega$ がポテンシャルの高さよりも十分に小さい時の非インスタントトンネル効果は、半古典的には安定・不安定多様体誘導トンネル効果で説明でき、量子論的には、超多重フォトアシストトンネル効果と解釈される。一方、 $\hbar\omega$ がポテンシャルの高さよりも大きい場合のトンネル過程は、一量子 $\hbar\omega$ の吸収によるエネルギー遷移で説明でき、摂動論が有効である。

1 Introduction

過去 20 年の間、古典的に非可積分な系のトンネル効果は、量子カオスの観点から多くの研究者によって研究され、可積分系には見られない複雑なトンネル現象が見つかってきている [1, 2, 3, 4]。また、複雑なトンネル現象に物理的な解釈を与える理論として、カオス・アシスト・トンネル効果、レゾナンス・アシスト・トンネル効果、ジュリアセット・アシスト・トンネル効果等の理論の提案もなされてきた [5, 6, 7, 8, 9]。

我々は、近年の論文で、インスタントでは説明できないトンネル効果を半古典的に説明するために、安定・不安定多様体誘導トンネル効果の理論を提案した [10, 11]。この理論は、首藤・池田のジュリアセット・アシスト・トンネル効果と本質的に同じものと考えられるが [8, 9]、彼らの理論が離散力学系に限られているのに対し、安定・不安定多様体誘導トンネル効果理論は、より一般的な時間連続な系に適用可能である。インスタントは、虚時間の時間発展によって生成されるので [12, 13]、複素領域に解析接続ができない離散系では取り扱う事が難しい。一方、時間連続な系では時間を複素化可能であり、インスタントを含めた形で多自由度系のトンネル効果を半古典論を用いて取り扱う事が可能である [14, 15]。

この報告では、インスタント型のトンネル効果が実質的に禁止される周期摂動の入ったステップ型ポテンシャルを用いて、非インスタントトンネル効果の半古典的および量子論的な解釈を試みる。安定・不安定多様体誘導トンネル効果理論が適用可能な解析的なポテンシャルとして、周期摂動を入れた滑らかなステップポテンシャル系を導入する [16]。この系で、摂動周波数を変化させた時のトンネル確率の変化を調べ、安定・不安定多様体誘導トンネル効果がどの領域でどの程度適用可能かを明らかにする。また、数値的にはあるが厳密に量子的な解が計算できる系として摂動の入ったステップポテンシャル系を導入する。この系には、複素半古典論は適用不可能であるが、類似の量子力学的なトンネル現象が起きる事が期待できる。この系を用いて、安定・不安定多様体誘導トンネル効果の量子力学的な解釈を試みる。このようにして、非インスタント型のトンネル効果の半古典論と量子論の双方からの相補的な解釈を与える。

2 モデルと量子論的な結果

インスタントン型のトンネル効果が実質的に禁止されるモデルとして、以下のような周期摂動を入れた滑らかな角を持つステップポテンシャルを考える [16].

$$\hat{H}(Q, P, \omega t) = \frac{1}{2} \hat{P}^2 + (1 + \epsilon \sin \omega t) \frac{1}{1 + \exp Q} \quad (1)$$

このポテンシャルに運動量 $P_I (< 0)$ を持つ平面波が右側から入射させた場合の準定常状態を考える [17]. そのような準定常状態は、以下のように波動演算子 $\hat{\Omega}_1^+$ を用いて表せる.

$$\langle Q_O | \hat{\Omega}_1^+(t_O) | P_I \rangle = \lim_{|Q_I| \rightarrow \infty} \sqrt{\frac{|P_I|}{2\pi\hbar}} e^{iP_I Q_I / \hbar} \int_0^\infty ds \langle Q_O | \hat{U}(\omega t_O, \omega t_O - \omega s) | Q_I \rangle \exp\left\{\frac{i}{\hbar} E_I s\right\} \quad (2)$$

ここで、 \hat{U} は時間推進演算子 $\hat{U}(\theta + \omega t : \theta) = \mathcal{T} \exp -\frac{i}{\hbar} \int_0^t \hat{H}(\theta + \omega s) ds$ (\mathcal{T} : 時間順序積) である. 添字 'I' と 'O' は、入射点と観測点の物理量である事を示す.

図1に、入射エネルギーを $E_I (= P_I^2/2)$ に取り、摂動強度 ϵ を $\epsilon = 0, 0.05, 0.1, 0.2$ とした時の準定常状態を示す. 無摂動 ($\epsilon = 0$) の場合には、トンネルはインスタントン型であり、トンネル確率はポテンシャル内で急速に減衰しトンネルによる透過は実質的に禁止されていると考えてよい. しかし、摂動が加わると、トンネル確率は透過領域でほぼ一定の値を取り、その値は摂動強度の増加とともに増大する. このようなトンネル現象は、安定・不安定多様体誘導トンネル効果によって説明可能である [16].

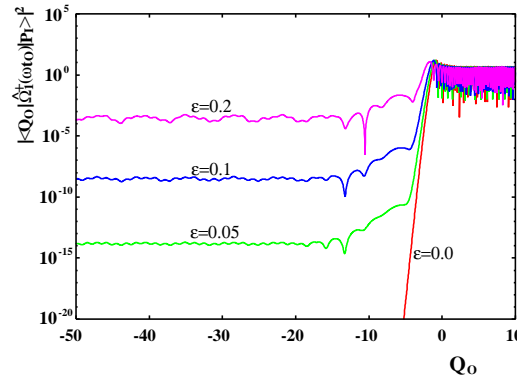


図1: 量子論的计算によって得られた摂動強度 $\epsilon = 0, 0.05, 0.1, 0.2$ の時の準定常状態. パラメータは、 $E_I = 0.75$, $\omega = 0.3$, $\hbar = 1000/(3\pi \times 2^{10}) \sim 0.1036$ である.

3 波動演算子の半古典表示

波動演算子の半古典表示は以下のように与えられる [17].

$$\langle Q_O | \hat{\Omega}_1^+(t_O) | P_I \rangle \sim \sum_{\text{c.t.}} \lim_{Q_I \rightarrow \infty} \sqrt{\frac{|P_I|}{2\pi\hbar}} e^{iP_I Q_I / \hbar} \left[\frac{1}{P_I} \frac{\partial^2 S_\Omega}{\partial E_I \partial Q_O} \right]^{1/2} \exp\left\{\frac{i}{\hbar} S_\Omega(Q_O, t_O, Q_I, E_I)\right\} \quad (3)$$

ここで, $S_{\Omega}(Q_O, t_O, Q_I, E_I)$ は, 作用積分

$$S_{\Omega}(Q_O, t_O, Q_I, E_I) \equiv \int_{Q_I}^{Q_O} PdQ - \int_{t_I}^{t_O} H(Q, P, \omega t) dt + E_I(t_O - t_I(Q_O, t_O, Q_I, E_I)) \quad (4)$$

である. (3) 式における $\sum_{\text{c.t.}}$ は, 半古典近似に寄与するすべての古典軌道について和を取る事を意味する.

作用積分 $S_{\Omega}(Q_O, t_O, Q_I, E_I)$ の引数は, 量子力学的な観測量になるので, Q_I, E_I, Q_O, t_O は, 実数値を取らなければならない. これらの力学量を固定すると軌道の境界条件が定まる. 一方, 初期時刻 t_I は, E_I と正準共役なので不確定である. Miller によって与えられた作業仮説では, 不確定な力学量は任意の複素値を取る事が許される [13]. したがって, 複素初期時刻 t_I を境界条件を満たす軌道を見つけるための複素探索パラメータとして使う事ができる. 初期時刻 t_I がサーチする複素初期時刻面 \mathcal{I} は,

$$\mathcal{I} \equiv \{Q, P, t_I | Q = Q_I, P = P_I, t_I \in \mathbf{C}\} \quad (5)$$

で与えられる. 一方, 観測面 \mathcal{F} は,

$$\mathcal{F} \equiv \{Q, P_O, t | Q = Q_O, P_O \in \mathbf{C}, t = t_O\} \quad (6)$$

で与えられる. Q_I, E_I, Q_O, t_O を固定するごとに, t_I を調節し \mathcal{I} を出発し \mathcal{F} に到達する (複素) 軌道を求めれば, それが半古典波動演算子に寄与する軌道になる (注: 厳密には Stokes 問題を考慮する必要がある). 準定常散乱状態の波動関数を Q_O の関数として求めるには, 他の観測量を固定し Q_O を実軸に沿って動かし, 各点ごとに寄与する軌道を求めればよい.

4 安定・不安定多様体誘導トンネル効果

4.1 メカニズム

量子系 (1) に対応する古典系を考える. 無摂動状態 ($\epsilon = 0$) では, エネルギーを $E = 1$ に取った場合の軌道はセパトラックスを構成し, 不安定固定点が $Q \rightarrow -\infty$ にあると考えれば, それに漸近する安定多様体または不安定多様体と考えられる. このような安定・不安定多様体は, 摂動が入っても, 変形されるが安定に存在し続ける. したがって, 障壁ポテンシャルで見られた安定・不安定多様体誘導トンネル効果と同じものが, ステップポテンシャルにも見られる [10, 11, 16].

図 2 に, 周期摂動下の滑らかなステップポテンシャルの場合の Poincare 写像 (概念図) を示す. W_s と W_u は, それぞれ安定多様体と不安定多様体である. 一周期分の初期面を写像したものを \mathcal{I} で表す. 入射エネルギー E_I が十分に小さく透過領域の確率波がトンネル効果によってのみ作られる場合には, $\text{Re}\mathcal{I}$ と $\text{Re}W_s$ は交差せず, $\text{Re}\mathcal{I}$ を写像すると細い実線で描かれた不変多様体を入射および反射領域に作り出す. しかし, 複素領域では, W_s は点線に示されたように変形され, \mathcal{I} との交点を持つ. 交点は孤立した点 $t_I = t_{Ic}$ になり臨界点と呼ばれる [10, 11]. 安定・不安定多様体誘導トンネル効果に寄与する (複素) 古典軌道は, 臨界点のごく近傍に初期点を持ち, 複素安定多様体 W_s に沿って $Q \rightarrow -\infty$ に向かって進む. それらの中で相対的にエネルギーが小さい軌道は, 途中で反転し不安定多様体 W_u に沿って戻り, 反射波に寄与するが, それ以外の軌道は, 安定多様体 W_s に沿って進み無限遠方 ($Q \rightarrow -\infty$) に到達し, トンネル確率を作り出す.

このようなトンネル機構は, インスタント型のトンネル効果とは全く異なるものである. ポテンシャル内部を進むインスタント軌道は, 無限遠方に到達するのに負の無限大の虚時間を必

要とし、そのため、トンネル確率は指数関数的に減少し、ポテンシャルの縁から極めて短い距離で実質的にゼロとなる。この意味で、ステップ型のポテンシャルでは、インスタント型のトンネル効果は実質的に禁止されていると行ってよい。そのような場合でも、摂動が加わると安定・不安定多様体誘導トンネル効果によるトンネル現象が観測される。

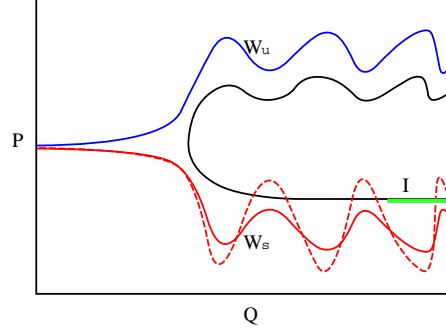


図 2: 安定・不安定多様体誘導トンネル効果の概念図

4.2 Melnikov 積分による臨界点の深さの評価

臨界点の存在は、Melnikov 積分法を用いて証明することができる [10, 11]. まず、安定多様体 W_s 上の軌道 $Q_s(t)$ のエネルギーは、以下のように与えられる。

$$H(Q_s(t), P_s(t), t) = 1 + \epsilon \sin \omega t + \Delta H_M(t) \quad (7)$$

ここで、

$$\Delta H_M(t) = \int_{-\infty}^t \frac{\partial H}{\partial t'} - \epsilon \omega \cos \omega t' dt' = \int_{-\infty}^t (V_0(Q_s(t)) - 1) \epsilon \omega \cos \omega t' dt' \quad (8)$$

で、 V_0 は無摂動系のポテンシャル $V_0(Q) \equiv \frac{1}{1 + \exp Q}$ である。Melnikov の方法では、 W_s 上の軌道 $Q_s(t)$ を無摂動系の解

$$t - t_0 = \frac{1}{\sqrt{2}} \log \left(4Y - 2 - 4\sqrt{Y^2 - Y} \right) + \frac{\sqrt{2(Y^2 - Y)}}{Y - 1} \quad (9)$$

に置き換えて積分を評価する。ここで、 $Y = 1 + \exp Q$ で、 t_0 は無摂動解の時間の原点で、軌道の初期点 (t_I, Q_I) と以下のように関係づけられる。

$$t_I = t_0 - Q_I / \sqrt{2} + \sqrt{2} \quad (10)$$

積分を実行すると $\text{Ret}_I \rightarrow -\infty$ とした時のエネルギー E_I が以下のように求まる。

$$E_I = 1 + \epsilon A(\omega) \sin(\omega t_{0c} + \phi) \in \mathbf{R} \quad (11)$$

ここで、

$$\tan \phi = \frac{\omega \chi_{\cos}}{1 - \omega \chi_{\sin}} \quad (12)$$

および

$$A(\omega) = \sqrt{(1 - \omega\chi_{\sin})^2 + \omega^2\chi_{\cos}^2} \quad (13)$$

で, χ_{\cos} と χ_{\sin} は, 以下の積分で定義される.

$$\begin{aligned} \chi_{\{\cos, \sin\}} &\equiv \int_{-\infty}^0 [V_0(Q_{s0}(\tau, t_0 = 0)) - 1] \{\cos, \sin\} \omega \tau \, d\tau \\ &\quad + \int_0^{-\infty} V_0(Q_{s0}(\tau, t_0 = 0)) \{\cos, \sin\} \omega \tau \, d\tau \end{aligned} \quad (14)$$

E_I が実数値を取る事を要請すると, 時間の原点 t_{0c} は孤立した点として求まり, (10) 式より対応する t_{Ic} も孤立点となり, 初期面 \mathcal{I} と安定多様体 W_s との交点 (臨界点) になる. $E_I < 1 - \epsilon A(\omega)$ の場合は, 実領域に交点はないが, その場合でも複素領域に交点は存在し, その深さは,

$$\text{Im}t_{Ic} = \frac{1}{\omega} \cosh^{-1}((1 - E_I)/\epsilon A(\omega)) \quad (15)$$

と評価される. $\omega \rightarrow 0$ の極限では, $A(\omega) \rightarrow 1$ となるので, 漸近形は,

$$\text{Im}t_{Ic} \sim \frac{1}{\omega} \cosh^{-1}\left(\frac{1 - E_I}{\epsilon}\right) \quad (16)$$

$$\sim \frac{1}{\omega} \log\left(\frac{2(1 - E_I)}{\epsilon}\right), \quad (\epsilon \ll 1 - E_I) \quad (17)$$

となる. 図3に, ω を変化させた時の $\text{Im}t_{Ic}$ の変化を示す. $\omega \ll 1$ の領域では, (16) 式 (または (17) 式) による評価通り, ω に逆比例する. 一方, $\omega \rightarrow \infty$ では, ϵ によらず一定値 $\pi/\sqrt{2}$ に収束しているように見える. 障壁ポテンシャルにおいても, 同様の臨界点の振る舞いが観測される [15]. したがって, 臨界点の深さの摂動周波数に対する変化には, 普遍性があると考えられる.

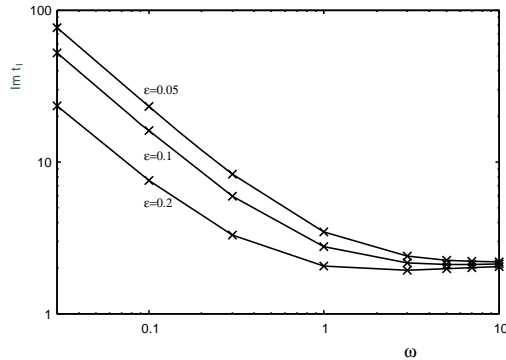


図 3: 臨界点の深さの ω に対する変化 ($\epsilon = 0.05, 0.1, 0.2$).

4.3 安定・不安定多様体誘導トンネル効果理論を用いたトンネル確率の評価

3章で導入した複素半古典論を実行すると, 図1に示す準定常散乱状態をほぼ忠実に再現できる [16]. そのためには, 寄与する複素半古典軌道をすべて見つけ出し, それらの確率的な重みを計算し足し合わせなければならない. これは極めて困難で時間のかかる作業であり現実的ではない.

ここでは、より簡単な方法で大まかなトンネル確率を評価する方法を提示する。安定・不安定多様体誘導トンネル効果では、臨界点 t_{Ic} を初期点とし複素安定多様体 W_s 上を動く臨界軌道が、複素トンネル軌道を誘導し、それらの大まかな性質を決定するので、トンネル確率を臨界軌道を用いて評価することが可能である [14, 15, 16].

具体的には、トンネル軌道の作用積分 (4) を臨界軌道のそれによって評価する。波動演算子の半古典表現 (3) で、振幅成分 $\sqrt{\frac{\partial^2 S_\Omega}{\partial E_I \partial Q_O}}$ の寄与を無視すれば、トンネル確率は以下のように評価される。

$$W_{SUMGT} \sim \frac{1}{2\pi\hbar} \exp\left\{-\frac{2}{\hbar} \text{Im} S_\Omega\right\} \quad (18)$$

トンネル軌道の作用積分 S_Ω を臨界軌道のそれに置き換え、Melnikov の方法を用いて評価する。Melnikov の方法が適用しやすいように、 S_Ω を以下のように書き換える。

$$S_\Omega(Q_O, t_O, Q_I, E_I) = \int_{t_{Ic}}^{t_O} (H - 2V) dt + E_I(t_O - t_{Ic}) \quad (19)$$

ここで、 $V(Q, \omega t) = a(t)V_0(Q)$, $a(t) = 1 + \epsilon \sin \omega t$ である。Melnikov の方法では、臨界軌道を、(11) 式で与えられる時間の原点 t_{0c} を持つ無摂動な W_s 上の軌道 $Q_{0c}(t)$ に置き換え、積分を評価する。細かな計算過程は省略するが、結果として以下のような評価が得られる。

$$\begin{aligned} \text{Im} S_\Omega(Q_O, t_O, Q_I, E_I) &\sim (1 - E_I) \text{Im} t_{Ic} - \frac{\epsilon}{\omega} \text{Im} \cos(\omega t_{0c}) \\ &+ \text{Im} \int_{t_{0c}}^{t'_O} dt \int_{t'_O}^t ds \dot{a}(s)(V_0(Q_{0c}(s)) - 1) + \text{Im} \int_{t_{Ic}}^{t_{0c}} dt \int_{t_{Ic}}^t ds \dot{a}(s)V_0(Q_{0c}(s)) \\ &- 2\text{Im} \int_{t_{0c}}^{t'_O} dt \epsilon \sin \omega t (V_0(Q_{0c}(t)) - 1) - 2\text{Im} \int_{t_{Ic}}^{t_{0c}} dt \epsilon \sin \omega t V_0(Q_{0c}(t)) dt \end{aligned} \quad (20)$$

ここで、 t'_O は $\text{Re} t'_O = t_O$, $\text{Im} t'_O = \text{Im} t_{Ic}$ を満たす点である。(20) 式を数値計算で評価し、(18) 式に代入すればトンネル確率が評価できる。低い周波数の領域 ($\omega \ll 1$) では、(20) 式の右辺の最初の 2 つの項が支配的になるので、(16), (17) 式を用いると、

$$\text{Im} S_\Omega \sim \frac{1 - E_I}{\omega} \cosh^{-1}\left(\frac{1 - E_I}{\epsilon}\right) - \frac{\epsilon}{\omega} \sinh\left(\cosh^{-1}\left(\frac{1 - E_I}{\epsilon}\right)\right) \quad (21)$$

$$\sim \frac{1 - E_I}{\omega} \log\left(\frac{2(1 - E_I)}{\epsilon}\right) - \frac{1 - E_I}{\omega} \quad (22)$$

と評価される。したがって、トンネル確率は $\propto \exp(-C/\omega)$ と評価され、 $\omega \rightarrow 0$ の極限で指数関数的に減衰する。

5 トンネル確率の ω による変化

図 4 に量子力学的な計算で得られたトンネル確率 ($\max_{-50 \leq Q_O \leq -30} |\langle Q_O | \hat{\Omega}_1^+(t_O) | P_I \rangle|^2$) と (18) 式と (20) 式から計算される半古典論による予測確率 W_{SUMGT} の ω による変化を摂動強度 $\epsilon = 0.05, 0.1, 0.2$ に対して示す。図中の $\omega_{cq} \equiv \frac{1 - E_I}{\hbar}$ は、半古典的領域と純量子力学的領域の境界を特徴づける周波数で、この周波数を越えると入射粒子は 1 つの量子 $\hbar\omega$ の吸収でポテンシャルステップを越えることができる。

$\omega < \omega_{cq}$ の領域では、トンネル確率は W_{SUMGT} とよく一致する。特に $\omega < 1$ 場合、(18) 式に (21) 式または (22) 式を代入すると、トンネル確率が $\propto \exp(-C/\omega)$ と評価され $\omega \rightarrow 0$ で指数関数

的に減少することがわかる. $\epsilon \ll 1 - E_I$ の場合には, (18) 式と (22) 式より, ϵ に対する依存性が $\propto \epsilon^{2\omega_{cq}/\omega}$ と評価され, 低周波数ではトンネル確率の摂動強度による違いが増大することがわかる.

$\omega > \omega_{cq}$ の領域では, 1つの量子 $\hbar\omega$ の吸収で, 反応粒子はポテンシャルステップを越えることが可能となる. このように, $\hbar\omega$ が大きな領域では半古典論 (安定・不安定多様体誘導トンネル効果理論) はもはや有効ではない. 実際, トンネル確率は, ω の増加とともに指数関数的に減少するが, W_{SUMGT} は ϵ に無関係に一定値に収束する. このような場合には, 量子力学的な摂動論が有効である [15]. トンネル確率は, 1つの量子 $\hbar\omega$ の吸収過程でほぼ決まり, 以下のように評価される.

$$\lim_{Q_O \rightarrow -\infty} |\langle Q_O | \hat{\Omega}_1^+(t_0) | P_I \rangle|^2 \sim \lim_{Q_O \rightarrow -\infty} \epsilon^2 \pi^2 D_\omega |\langle u_{E_I + \hbar\omega} | V_0 | u_{E_I} \rangle|^2 \quad (23)$$

ここで, u_E は無摂動系の散乱固有状態で, 係数 D_ω は, 観測点 $Q = Q_O$ での $u_{E_I + \hbar\omega}$ の確率振幅で以下のように与えられる [19].

$$\begin{aligned} D_\omega &= \lim_{Q_O \rightarrow -\infty} |\langle Q_O | u_{E_I + \hbar\omega} \rangle|^2 \\ &= \frac{1}{2\pi\hbar|P_T|} \frac{\sinh^2(\pi(P_R + |P_T|)/\hbar) - \sinh^2(\pi(P_R - |P_T|)/\hbar)}{\sinh^2(\pi(P_R + |P_T|)/\hbar)} \end{aligned} \quad (24)$$

ここで, $P_R = \sqrt{2(E_I + \hbar\omega)}$ および $P_T = -\sqrt{2(E_I + \hbar\omega - 1)}$ である. したがって, (23) 式で与えられるトンネル確率は, フォト・アシスト・トンネル効果に他ならない. 図4に示すように, 量子摂動論による近似 (細い破線) はトンネル確率の変化をよく近似する. D_ω は $\hbar\omega \rightarrow \infty$ の極限で $D_\omega \rightarrow \frac{1}{2\pi\hbar|P_T|}$ となるので, 指数関数的な減衰は, 主に, 遷移係数 $|\langle u_{E_I + \hbar\omega} | V_0 | u_{E_I} \rangle|^2$ による.

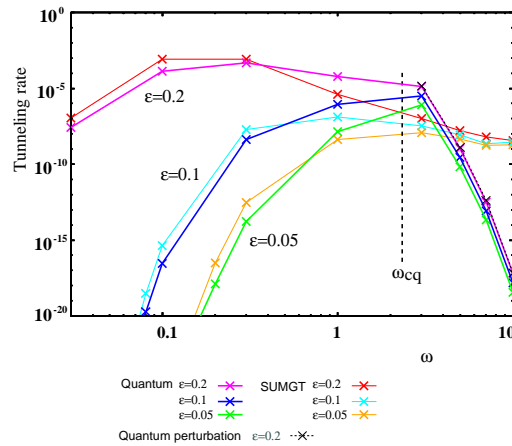


図 4: トンネル確率と安定・不安定多様体誘導トンネル効果理論による予測確率 (18) の周波数領域 ($0.03 \leq \omega \leq 10$) における比較. 3つの摂動強度 $\epsilon = 0.05, 0.1, 0.2$ の場合を示す. 細い破線は, $\epsilon = 0.2$ の時の量子摂動論 (23) の結果である.

6 非インスタントトンネル効果の量子論的解釈

6.1 モデルと準固有状態

この章では、安定・不安定多様体誘導トンネル効果の量子力学的な解釈を試みる。そのために、ステップポテンシャルに摂動を入れた系を考える。

$$H(Q, P, \omega t) = \frac{1}{2}P^2 + (1 + \epsilon \sin \omega t)H(-Q) \quad (25)$$

ここで、 $H(x)$ は、Heaviside 関数である。この系はポテンシャルの角が直角なので、複素領域に解析接続する事は不可能であり、複素半古典論を適用する事はできない。しかし、量子力学的な振る舞いはほぼ同じであると期待され、また、摂動論を使わずに厳密に準定常散乱状態を書き下せるので、量子力学的な解釈を考えるのに都合が良い。

エネルギー E_I を持った入射波 Ψ_I は、

$$\Psi_I(Q) = e^{-\frac{i}{\hbar}E_I t} e^{\frac{i}{\hbar}P_I Q} \quad (P_I = -\sqrt{2E_I}) \quad (26)$$

と書ける。図5に示すように、振動するポテンシャルとの相互作用による量子の多重吸収・放出のために、反射波はエネルギー $E_n = E_I + n\hbar\omega$ (運動量: $P_{Rn} = \sqrt{2E_n}$) の平面波の重ね合わせで書ける。

$$\Psi_R(Q) = \sum_n R_n e^{-\frac{i}{\hbar}E_n t} e^{\frac{i}{\hbar}P_{Rn} Q} \quad (27)$$

ここで、 R_n は反射係数である。一方、透過波は、

$$\Psi_T(Q) = \sum_n T_n e^{\frac{i}{\hbar}\epsilon \cos \omega t} e^{-\frac{i}{\hbar}E_n t} e^{\frac{i}{\hbar}P_{Tn} Q} \quad (28)$$

と書ける [18]。ここで、 T_n は透過係数で、運動量 P_{Tn} は $P_{Tn} = -\sqrt{2(E_n - 1)}$ で与えられる。したがって、透過波は、振動するポテンシャルの影響で時間的な変調 $e^{\frac{i}{\hbar}\epsilon \cos \omega t}$ を受けた運動量 P_{Tn} の平面波の重ね合わせで表せる。

$E_n < 1$ ならば、運動量 P_{Tn} は負の純虚数になり、

$$\lim_{Q \rightarrow -\infty} e^{\frac{i}{\hbar}P_{Tn} Q} \rightarrow 0 \quad (29)$$

となる。したがって、 Q が大きな負の実数値を取るならば、透過波は、

$$\Psi_T(Q) \sim \sum_{n \geq n^*} T_n e^{\frac{i}{\hbar}\epsilon \cos \omega t} e^{-\frac{i}{\hbar}E_n t} e^{\frac{i}{\hbar}P_{Tn} Q} \quad (30)$$

と近似できる。ここで、 $n^* = \min\{n | E_n > 1\}$ である。

$Q = 0$ における解の接続条件から反射係数 R_n と透過係数 T_n が決まる。まず、波動関数そのものの接続条件 $\Psi_I + \Psi_R = \Psi_T$ より、

$$1 + \sum_n R_n e^{-in\omega t} = \sum_n T_n e^{\frac{i}{\hbar}\epsilon \cos \omega t} e^{-in\omega t} \quad (31)$$

となる。次に、一階微分の連続性 $\frac{\partial}{\partial Q}\Psi_I + \frac{\partial}{\partial Q}\Psi_R = \frac{\partial}{\partial Q}\Psi_T$ は、

$$P_I + \sum_n R_n P_{Rn} e^{-in\omega t} = \sum_n T_n P_{Tn} e^{\frac{i}{\hbar}\epsilon \cos \omega t} e^{-in\omega t} \quad (32)$$

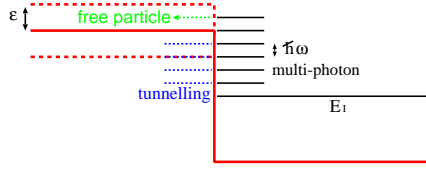


図 5: 周期的に変動するステップポテンシャル.

となる. ここで, $-P_I = P_{R0}$ である.

Bessel 関数 $J_n(z)$ を用いると, 時間変調項 $e^{\frac{i}{\hbar} \frac{\epsilon}{\omega} \cos \theta}$ は,

$$e^{\frac{i}{\hbar} \frac{\epsilon}{\omega} \cos \theta} = \sum_n i^n J_n\left(\frac{\epsilon}{\hbar\omega}\right) e^{in\theta} \quad (33)$$

と書け, (31) 式と (32) 式は, それぞれ,

$$\delta_{n,0} + R_n = \sum_{n'} T_{n'} i^{n'-n} J_{n'-n}\left(\frac{\epsilon}{\hbar\omega}\right) \quad (34)$$

と

$$-\delta_{n,0} P_{Rn} + R_n P_{Rn} = \sum_{n'} T_{n'} P_{Tn'} i^{n'-n} J_{n'-n}\left(\frac{\epsilon}{\hbar\omega}\right) \quad (35)$$

となる. したがって, (34) 式と (35) 式から, $\{T_n\}$ の連立方程式

$$\sum_{n'} T_{n'} (P_{Rn} - P_{Tn'}) i^{n'-n} J_{n'-n}\left(\frac{\epsilon}{\hbar\omega}\right) = 2P_{Rn} \delta_{n,0} \quad (36)$$

が得られる.

6.2 準固有状態と透過係数

(36) 式を解くと, 透過係数 $\{T_n\}$ が求まる. それを (34) 式に代入すると反射係数 $\{R_n\}$ が求まる. 図 6(a) に, そのようにして求めた $\{R_n\}$ と $\{T_n\}$ を用いて再現した $\omega = 0.3$ の準定常状態を 4 つの摂動強度 $\epsilon = 0, 0.05, 0.1, 0.2$ に対して示す. 図 6(a) の波動関数は, 図 1 に示した滑らかなステップポテンシャルの波動関数によく似た振る舞いをしているのがわかる. $\epsilon = 0$ の無摂動な場合には, インスタントン型のトンネルで, トンネル確率はポテンシャル障壁の中で指数関数的に減衰する. 一方, 摂動が入るとトンネル確率は安定・不安定多様体誘導トンネル効果のトンネル確率とほぼ同じ振る舞いを示し, $Q < 0$ の領域ではほぼ一定値になり, その値は ϵ とともに増加する.

図 6(b) に, $\epsilon = 0.05, 0.1, 0.2$ の場合の透過係数 $|T_n|^2$ をエネルギー E の関数として示す. $|T_n|^2$ は, 中心が $E = E_I$ で幅が 2ϵ のプラトー型の分布になる. しかし, プラトー上部の成分は, (28) 式の和にほとんど寄与しない. なぜなら, プラトー上部のエネルギー E_n は $E_n \leq E_I + \epsilon < 1$ の領域にあるので, $e^{\frac{i}{\hbar} P_{Tn} Q}$ で評価されるトンネル確率は, (29) 式で示したように $Q \rightarrow -\infty$ の極限で指数関数的に減衰するからである. これが, インスタントン型のトンネル効果が実質的に禁止されるメカニズムの量子力学的な解釈である.

したがって、 $n \geq n^*$ の成分がトンネルに寄与するが、 $n = n^*$ の成分が支配的になる。なぜならば、(30) 式より、各成分の寄与は $|T_n|^2$ の大きさでほぼ決まり、 $n \geq n^*$ における $|T_n|^2$ の中には、 $|T_n|^2$ が最も大きな値を取るからである (図 6(b)). 実際、図 6(a) と (b) を比較すると、各摂動強度において $|T_n|^2$ とトンネル確率がほぼ一致しているのがわかる。摂動強度 ϵ の変化によってプラトーの形が変わり、トンネル確率も変化する。 $n = n^*$ の成分に対応する古典軌道は、 $n \geq n^*$ の古典軌道の中で、最も小さな負の運動量 P_{Tn} を持ち、安定・不安定多様体誘導トンネル効果に寄与する複素トンネル軌道と同じ役割をする。ただし、この場合は安定多様体上の臨界軌道は存在しない。

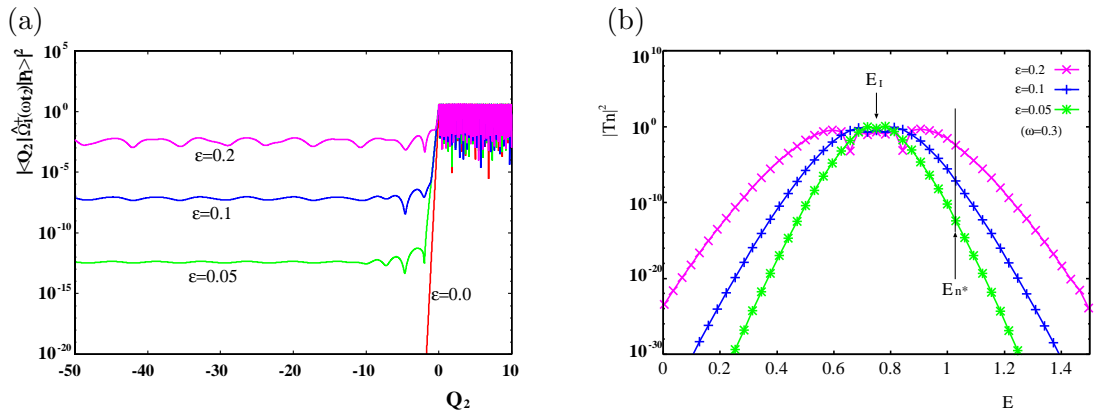


図 6: ステップポテンシャルの $\omega = 0.3$, $\epsilon = 0, 0.05, 0.1, 0.2$ における計算結果. 他のパラメータは、図 1 と同じである. (a) 準定常状態. (b) 透過係数.

6.3 トンネル確率の ω による変化

図 7 に、トンネル確率 ($\max_{-50 < Q < -20} |\Psi_T(Q)|^2$) の ω による変化を 3 つの摂動強度 $\epsilon = 0.05, 0.1, 0.2$ について示す。 $\omega < \omega_{cq}$ の領域では、トンネル確率は $\omega \rightarrow 0$ の極限で、 $\propto \exp(-C/\omega)$ のように指数関数的に減衰し、図 4 に示す滑らかなステップポテンシャルとほぼ同じ傾向を示す。したがって、ステップ型のポテンシャルのこの領域でのトンネルメカニズムは、ポテンシャル形状の詳細によらず同じであると考えてよい。この領域の大部分で、 $\hbar\omega \ll 1 - E_I$ と考えてよいので、粒子がトンネル $E_n > 1$ のエネルギーを獲得するためには、たくさんの量子を吸収する必要がある。これは、超多重フォト・アシスト・トンネル効果に他ならない。この過程を半古典的に解釈すると、安定多様体によるトンネル軌道の誘導によって粒子が大きなエネルギーを獲得すると解釈できる。

$\omega > \omega_{cq}$ の領域では、トンネル確率は ω_{cq} の少し上で最大値を取った後、 ϵ^2/ω^2 に比例したベキ的な減衰を示し、図 4 の滑らかなステップポテンシャルで見られる指数関数的な減衰とは異なる。この違いはポテンシャルの形状によるもので、この 2 つのポテンシャルの中間的な形状のポテンシャルでは、指数関数的ではあるが減衰率が小さくなり、中間的な減衰になる。 ϵ^2/ω^2 型のベキ的な減衰は箱形ポテンシャルでも見られ [18]、ステップや箱形の共通の性質と考えられる。

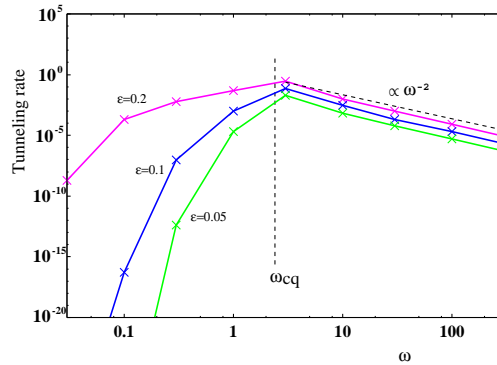


図 7: ステップポテンシャルにおけるトンネル確率の ω による変化. 3つの摂動強度 $\epsilon = 0.05, 0.1, 0.2$ の場合を示す.

7 Discussion

この報告では、インスタントン型のトンネル効果が実質的に禁止される周期摂動の入ったステップ型ポテンシャルを用いて [16], 非インスタントントンネル現象の半古典的および量子論的な解釈を試みた. 特に、摂動周波数を変化させたときのトンネル確率の変化を調べた. その結果、トンネル確率は、 $\omega \rightarrow 0$ と $\omega \rightarrow \infty$ の極限でゼロになるが、その中間領域で著しく大きな値を取り、摂動強度とともに増大することがわかった.

半古典領域 $\omega < \omega_{cq} (\equiv (1 - E_I)/\hbar\omega)$ のトンネル現象は、周期摂動の入った滑らかなステップポテンシャルでは安定・不安定多様体誘導トンネル効果で説明が可能である [16]. $\omega \rightarrow 0$ の極限で、トンネル確率は $\propto \exp(-C/\omega)$ のように指数関数的に減衰するが、これは、複素領域の臨界点 t_{Ic} (半古典近似の初期面 \mathcal{I} と安定多様体 W_s の交点) が $\omega \rightarrow 0$ の極限で $\text{Im}t_{Ic} \propto 1/\omega$ の様に発散し、そのため、複素安定多様体に誘導されるトンネル軌道の作用積分 S_Ω が $\text{Im}S_\Omega \propto 1/\omega$ のように発散するためである. 周期摂動の入ったステップポテンシャルを用いると、量子力学的な解釈が可能である. この場合も $\omega \rightarrow 0$ の極限で、トンネル確率は $\propto \exp(-C/\omega)$ のように指数関数的に減衰する. トンネル過程を量子力学的に見ると、粒子が、摂動により多数の量子を吸収することで、入射エネルギー $E_I \ll 1$ からポテンシャルステップより大きなエネルギー領域 $E_n > 1$ へジャンプする過程と解釈される. したがって、超多重フォト・アシスト・トンネル効果と解釈可能である. トンネル確率は、透過係数 $|T_n|^2$ で決まるが、古典的に許されるエネルギー変動領域 $E_I - \epsilon \leq E_n \leq E_I + \epsilon$ を越えた遷移なので、 $|T_n|^2$ は n の増加とともに指数関数的に小さくなり、 $n^* = \min\{n | n\hbar\omega > 1 - E_I\}$ で決まる n^* 状態の透過係数 $|T_{n^*}|^2$ がトンネル確率を支配する. $n = n^*$ の成分に対応する古典軌道は、 $n \geq n^*$ の古典軌道の中で最も小さな負の運動量 P_{Tn} を持ち、安定・不安定多様体誘導トンネル効果に寄与する複素トンネル軌道と同じ役割をすると考えられる.

純量子力学的な領域 $\omega > \omega_{cq}$ では、1つの量子 $\hbar\omega$ の吸収で、粒子はポテンシャルステップを越えるエネルギーを獲得するので、半古典論 (安定・不安定多様体誘導トンネル効果) は有効でなくなる. この場合、量子力学的摂動計算で E_I から $E_I + \hbar\omega$ への遷移確率を計算すればトンネル確率を評価できる. トンネル確率の摂動周波数 ω による変化はポテンシャルの形状に依存し、滑らかなステップポテンシャルでは指数関数的に減少するのに対し、ステップポテンシャルでは $\propto 1/\omega^2$ のようにべき的に減衰する. いずれにせよ、 $\omega \rightarrow \infty$ でゼロになる.

最後に、インスタントンと非インスタントントンネル効果が共存する周期摂動が入った障壁ポテ

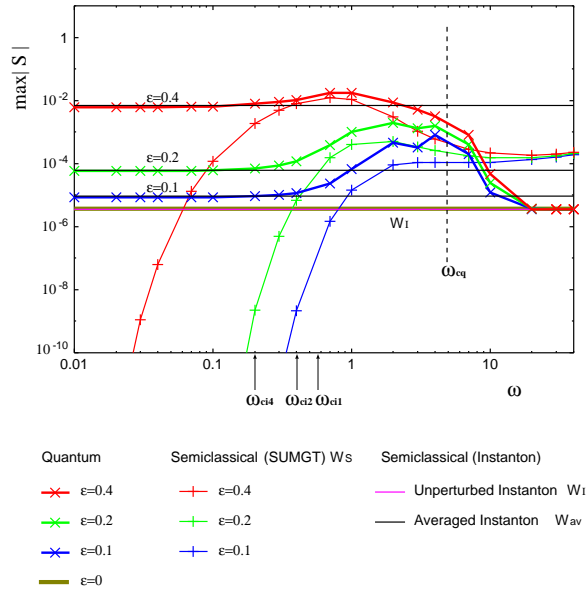


図 8: 障壁ポテンシャルにおけるトンネル確率の周波数による変化 [15].

ンシャルの場合と比較してみよう [15]. 障壁ポテンシャルでは, 図 8 に示すように $\omega \rightarrow 0$ と $\omega \rightarrow \infty$ の極限で有限の値に収束する. $\omega \rightarrow 0$ でのトンネル確率は, 摂動強度 ϵ に応じた一定値に収束し, その値は瞬間インスタントンから求めたトンネル確率を摂動周期で平均化したものと良く一致する. 一方, $\omega \rightarrow \infty$ のトンネル確率は, 摂動強度に無関係に無摂動系のそれに収束し無摂動系インスタントンで説明可能である. したがって, $\omega \rightarrow 0$ と $\omega \rightarrow \infty$ の極限では, インスタントンのトンネル効果が支配的になるが, 中間領域では, 安定・不安定多様体誘導トンネル効果が支配的になり, その結果トンネル確率が増大する. このような場合の量子力学的な解釈を試みることは, 極めて面白い課題であり, 周期摂動が入った箱形ポテンシャルが良いモデルになると考えられる.

謝辞

この研究は科学研究費補助金基盤研究 (B) No.20340100 の補助のもとで行われた.

参考文献

- [1] S.Tomsovic ed. *Tunneling in Complex Systems* (World Scientific, Singapore 1998).
- [2] J.Ankerhold, *Quantum Tunneling in Complex Systems The Semiclassical Approach*, (Springer-Verlag Berlin Heidelberg 2007).
- [3] S. Keshavamurthy, Int. Rev. Phys. Chem. **26**, 521 (2007).
- [4] S. Keshavamurthy and P. Schlagheck eds. *Dynamical Tunneling: Theory and Experiment* (CRC Press Taylor & Francis Group 2011).

- [5] O. Bohigas, S. Tomsovic, and D. Ullmo, *Phys. Rev. Lett.* **65**, 5 (1990); *Phys.Rep.* **223**, 45 (1993) 45.
- [6] O. Brodier, P. Schlagheck and D. Ullmo, *Phys. Rev. Lett.* **87**, 064101 (2001); O. Brodier, P. Schlagheck and D. Ullmo, *Ann. Phys.(NY)* **300**, 88 (2002); C. Eltschka and P. Schlagheck, *Phys. Rev. Lett.* **94**, 014101 (2005).
- [7] A. Bäcker, R. Ketzmerick, S. Löck and L. Schilling *Phys. Rev. Lett.* **100**, 104101 (2008); A. Bäcker et al. *Phys. Rev. Lett.* **100**, 174103 (2008); A. Bäcker, R. Ketzmerick, S. Löck, J. Wiersig and M. Hentschel *Phys. Rev. A* **79**, 063804 (2009); S. Löck, A. Bäcker, R. Ketzmerick and P.Schlagheck, *Rhys. Rev. Lett.* **104**, 114101 (2010).
- [8] A. Shudo and K. S. Ikeda, *Phys. Rev. Lett.* **74**, 682 (1995); *Physica D***115**, 234 (1998); T.Onishi *el al*, *Phys. Rev. E* **64**, 025201(R) (2001); *Phys. Rev. E* **68**, 056211 (2003).
- [9] A. Shudo, Y. Ishii and K. S. Ikeda, *J.Phys.A* **35**, L225 (2002); *Europhysics Letters* **81**, 50003 (2008); *J.Phys.A* **42**, 265101 (2009); *J.Phys.A* **42**, 265102 (2009).
- [10] K. Takahashi, A.Yoshimoto and K. S. Ikeda, *Phys.Lett.A* **297**, 370 (2002); K.Takahashi and K. S. Ikeda, *J.Phys.A* **36**, 7953 (2003).
- [11] K. Takahashi and K. S. Ikeda, *Europhysics Letters* **71**, 193 (2005); *J.Phys.A.* **41**, 095101 (2008).
- [12] L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, N.Y., 1981).
- [13] W. H. Miller, *J. Chem. Phys.* **53**, 1949 (1970); *Adv. Chem. Phys.* **25**, 69 (1974).
- [14] K. Takahashi and K. S. Ikeda, *Phys.Rev.A* **79**, 052114 (2009).
- [15] K. Takahashi and K. S. Ikeda, *J. Phys. A* **43**, 192001 (2010).
- [16] K.Takahashi and K.S.Ikeda, *Phys. Rev. Lett.* **97**, 240403 (2006).
- [17] K. Takahashi and K. S. Ikeda, *Ann. Phys. (NY)* **283**, 94 (2000).
- [18] M. Büttiker and R. Landauer, *Phys. Rev. Lett.* **49** (1982) 1739; *Physica Scripta* **32**, 429 (1985).
- [19] L.D. Landau and E.M. Lifshitz, *Quantum mechanics Non-relativistic theory* (Butterworth Heinemann, Amsterdam,, 1977).

Vertex couplings in quantum graphs

Ondřej Turek

(Joint work with Pavel Exner and Taksu Cheon)

Laboratory of Physics, Kochi University of Technology
Tosa Yamada, Kochi 782-8502, Japan
email: ondrej.turek@kochi-tech.ac.jp

Abstract

We propose a parametrization of the family of all physically admissible couplings in quantum graph vertices. With the help of this result we solve the longstanding open problem of approximating by regular graphs all singular vertex couplings in quantum graph vertices. The approximating graph is constructed as follows. We begin with a star graph with n arms and a general coupling in the vertex - the *approximated* system. Then we consider a system of n half lines whose endpoints are connected by lines of a length d such that there are δ couplings and δ interactions placed in the endpoints of the half lines and in the centers of the connecting lines; at the same time, the connecting lines are supposed to support vector potentials - the *approximating* system. It turns out that the approximated system can be considered as a limit of the family of approximating systems as $d \rightarrow 0$, where the convergence is in the norm-resolvent sense.

1 Introduction

The concept of quantum mechanics on graphs was firstly used in the fifties of the twentieth century in the paper [RS53] where it had been suggested as a model to study the spectra of aromatic hydrocarbons. The basic idea is that the graph represents the configuration space of the system, in other words, motion of a quantum particle is confined to the graph. However, at that time the model has not come into practical use. This changed a few decades later, when the quantum graph concept turned out to be very convenient for the study of physical properties of nanostructures. It has two natural reasons, namely:

- a graph represents a natural model for a graph-like structure,
- these models are relatively simple from the mathematical point of view.

The mathematical simplicity arises from the fact that the graph may be considered as a quasi-one-dimensional variety, thus its spectral and scattering analysis is reduced to solving ordinary differential equations, whereas in standard two- or three-dimensional problems, partial differential equations have to be solved.

With regard to the expansion of the nanotechnology expected for the near future, the importance and application potential of quantum graphs is going to grow rapidly, and their theory itself will probably widely develop. For now, despite of the significant progress that has been made in the last two decades, the theory is still far from being as complete as the theory of one-dimensional

Schrödinger operators. There remain many open problems, often of a fundamental character, that will have to be solved. In this paper we address two of them: the problem of parametrization of vertex couplings, and above all, the question of physical meaning of quantum graph vertices.

2 Preliminaries

Let us consider a free spinless particle whose motion is confined to a graph Γ having n edges e_1, \dots, e_n of the lengths $\ell(e_1), \dots, \ell(e_n)$. A *wave function* Ψ of the system has then n components, $\Psi = (\psi_1, \psi_2, \dots, \psi_n)^T$, and the corresponding Hilbert space \mathcal{H} is given by $\bigoplus_{j=1}^n L^2(\ell(e_j))$. The Hamiltonian, denoted by H_Γ , acts as a minus second derivative on each wave function component, i.e.

$$H_\Gamma \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} = \begin{pmatrix} -\psi_1'' \\ \vdots \\ -\psi_n'' \end{pmatrix}.$$

In this definition we neglect values of physical constants, i.e. we put $\hbar = 2m = 1$, because they will play no role in our considerations.

In the more general case when there are potentials V_1, \dots, V_n and vector potentials A_1, \dots, A_n imposed on the graph edges, the Hamiltonian action is given by

$$H_\Gamma \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} = \begin{pmatrix} (-i\frac{d}{dx} - A_1)^2 \psi_1 + V_1 \cdot \psi_1 \\ \vdots \\ (-i\frac{d}{dx} - A_n)^2 \psi_n + V_n \cdot \psi_n \end{pmatrix}.$$

The pair (Γ, H_Γ) , where Γ is a graph and H_Γ a Hamiltonian on Γ , is called *quantum graph*.

In order to have the Hamiltonian fully defined, it is necessary to specify also boundary conditions at the vertices. Consider a vertex $v \in V$ of degree n , i.e. there are n edges going from v that may be without loss of generality enumerated as e_1, \dots, e_n . Let us denote for all $j \in \hat{n}$ the wave function component on e_j by ψ_j and suppose that its variable x_j runs over the interval $(0, \ell(e_j))$, where the value 0 corresponds to v and $\ell(e_j)$ to the other endpoint of the edge. By *boundary values* of the wave function in the vertex v we understand the two vectors $\Psi(0)$ and $\Psi'(0)$ defined by

$$\Psi(0) := \begin{pmatrix} \psi_1(0_+) \\ \vdots \\ \psi_n(0_+) \end{pmatrix}, \quad \Psi'(0) := \begin{pmatrix} \psi_1'(0_+) \\ \vdots \\ \psi_n'(0_+) \end{pmatrix},$$

i.e.

- $\Psi(0)$ is a vector from \mathbb{C}^n which contains limits of the values of $\psi_1(x), \dots, \psi_n(x)$ in the vertex v ,
- $\Psi'(0)$ is a vector from \mathbb{C}^n containing limits of the first derivatives of $\psi_1(x), \dots, \psi_n(x)$ taken in the *outgoing* sense.

Since the Hamiltonian is a second-order linear operator, the boundary conditions in the vertex v of degree n have the form

$$A\Psi(0) + B\Psi'(0) = 0 \tag{1}$$

for certain $A, B \in \mathbb{C}^{n,n}$.

The boundary conditions have to be specified in such a way that the Hamiltonian H_Γ is a self-adjoint operator, or in physical terms that the probability currents at all the vertices are conserved.

A standard form of the boundary conditions was derived by Kostykin and Schrader in 1999 [KS99]; they showed that H_Γ is self-adjoint if and only if for every vertex $v \in V$ the corresponding matrices A and B satisfy the following two conditions:

- $\text{rank}(A|B) = \text{deg}(v)$,
 - the product AB^* is a self-adjoint matrix,
- (2)

where the symbol $(A|B)$ denotes the matrix $\text{deg}(v) \times 2 \text{deg}(v)$ with A and B forming first and second $\text{deg}(v)$ columns, respectively.

However, the pair (A, B) from (2) is not uniquely given, because it can be replaced by a pair (CA, CB) for any regular matrix $C \in \mathbb{C}^{\text{deg}(v), \text{deg}(v)}$ – it is obvious that $A\Psi(0) + B\Psi'(0) = 0$ and $(CA)\Psi(0) + (CB)\Psi'(0) = 0$ define the same coupling. To make the matrices A and B unique, one has to impose an additional constraint on them. The first unique (constrained) version of boundary conditions have been found in 2000 independently by Harmer [Ha00] and Kostykin & Schrader [KS00]. They showed that for any vertex coupling there is a unitary matrix $U \in \mathbb{C}^{\text{deg}(v), \text{deg}(v)}$, uniquely given by the formula $U = -(A + iB)^{-1} \cdot (A - iB)$, such that the matrices A and B in Equation (1) may take the form $A = U - I$, $B = i(U + I)$, where I is the identity matrix $\text{deg}(v) \times \text{deg}(v)$. Hence, the boundary conditions can be written as

$$(U - I)\Psi(0) + i(U + I)\Psi'(0) = 0 \quad (3)$$

where U is unambiguously given.

The formulation (3) of boundary conditions can be in fact called a *parametrization* of the family of vertex couplings, since for any boundary conditions in a vertex of degree n there is exactly one matrix $U \in U(n)$ complying with (3). Consequently, the family of vertex couplings in a vertex of degree n has n^2 real parameters, because the same is true for the group $U(n)$. This fact is, however, well known already from the analysis of quantum graphs in terms of self-adjoint extensions [ES89].

3 ST -form of boundary conditions

We have explained in Preliminaries that the matrices A, B in quantum vertex boundary conditions $A\Psi(0) + B\Psi'(0) = 0$ are generally not unique, but can be made unique by requiring them to take the special form $A = U - I$, $B = i(U + I)$. Let us present other special unique forms of A, B , which are in certain applications more convenient.

3.1 ST -form

We demonstrated in [CET10a] that the boundary conditions $A\Psi(0) + B\Psi'(0) = 0$ can always be rewritten in the form

$$\left(\begin{array}{c|c} I^{(m)} & T \\ \hline 0 & 0 \end{array} \right) P\Psi'(0) = \left(\begin{array}{c|c} S & 0 \\ \hline -T^* & I^{(n-m)} \end{array} \right) P\Psi(0), \quad (4)$$

where

- $m = \text{rank}(B)$,
- S is a self-adjoint matrix $m \times m$,

- T is a general complex matrix $m \times (n - m)$,
- $I^{(m)}, I^{(n-m)}$ are identity matrices $m \times m, (n - m) \times (n - m)$,
- P is a permutation matrix.

Note that the presence of the permutation matrix P means nothing but a renumbering of the edges. Equation (4) thus implies that the matrices A, B from (1) can be converted to

$$A = - \left(\begin{array}{c|c} S & 0 \\ \hline -T^* & I^{(n-m)} \end{array} \right) \quad \text{and} \quad B = \left(\begin{array}{c|c} I^{(m)} & T \\ \hline 0 & 0 \end{array} \right). \quad (5)$$

Compared to the constraint $A = U - I, B = i(U + I)$, (5) has the following properties.

- Advantages:
 - The elements of S and T represent the coupling parameters in a straightforward way,
 - the matrices involved are often sparse, i.e. many of the matrix elements vanish.
- Disadvantages:
 - The structure of (4) is complicated and depends on $\text{rank}(B)$,
 - (4) does not allow arbitrary numbering of the edges, i.e., sometimes the numbering has to be “corrected” by a permutation matrix.

From now on we will call the form of boundary conditions (4) the *ST-form*.

3.2 PQRS-form

The matrices A, B constrained according to (5) are structured according to the rank of B . One can symmetrize them, so that they depend on both $\text{rank}(A), \text{rank}(B)$, see our paper [CET10b], but this is done at the price of their increased complexity. The symmetrized boundary conditions take the so-called *PQRS-form*

$$\left(\begin{array}{c|c|c} I^{(p+q-n)} & 0 & P \\ \hline R & I^{(n-p)} & Q + RP \\ \hline 0 & 0 & 0 \end{array} \right) \Psi'(0) = \left(\begin{array}{c|c|c} S & -SR^* & 0 \\ \hline 0 & 0 & 0 \\ \hline -P^* & -Q^* & I^{(n-q)} \end{array} \right) \Psi(0),$$

where

- $p, q \in \mathbb{N}_0, p + q \leq n$,
- S is a self-adjoint matrix $p \times p$,
- $I^{(p+q-n)}, I^{(n-p)}, I^{(n-q)}$ are identity matrices of the given sizes,
- P, Q, R are general complex matrices of the corresponding sizes.

3.3 Kuchment's projector form of boundary conditions

Another formulation of boundary conditions, quite different from the “standard” form $A\Psi(0) + B\Psi'(0) = 0$, has been derived by P. Kuchment in [Ku04]. He has proved the existence of an orthogonal projector \mathcal{P} in \mathbb{C}^n with the complementary projector $\mathcal{Q} = I - \mathcal{P}$ and a self-adjoint operator L in \mathcal{QC}^n such that the boundary conditions may be written as

$$\begin{aligned} \mathcal{P}\Psi(0) &= 0 \\ \mathcal{Q}\Psi'(0) + L\mathcal{Q}\Psi(0) &= 0. \end{aligned} \tag{6}$$

Note that (6) yield $2n$ equations, some of them are linearly dependent, whereas the standard form of boundary conditions $A\Psi(0) + B\Psi'(0) = 0$ contains just n linearly independent equations.

It can be shown that the ST -form is related to the Kuchment's formulation. It holds:

- \mathcal{P} is a projector onto the linear span of the columns of the matrix $\begin{pmatrix} -T \\ I^{(n-m)} \end{pmatrix}$,
- \mathcal{Q} is a projector onto the linear span of the columns of the matrix $\begin{pmatrix} I^{(m)} \\ T^* \end{pmatrix}$,
- the operator $L\mathcal{Q}$ acts as

$$L\mathcal{Q} = - \begin{pmatrix} I^{(m)} \\ T^* \end{pmatrix} (I^{(m)} + TT^*)^{-1} S (I^{(m)} + TT^*)^{-1} \begin{pmatrix} I^{(m)} & T \end{pmatrix}.$$

For the proof see [Tu09].

4 The δ -coupling

As we have seen in Preliminaries, the family of vertex couplings is plentiful: if $\deg(v) = n$, it has n^2 real parameters. However, there is a couplings that plays a prominent role among all of them, namely the δ -coupling.

4.1 δ -interaction

Let us consider at first a “vertex of degree 2”, i.e., a line on which we choose a point x and call it a vertex. We say that there is a δ -interaction (or equivalently δ -potential) in the point x if the wave function obeys the following boundary conditions:

$$\psi(x_+) = \psi(x_-) =: \psi(x), \quad \psi'(x_+) - \psi'(x_-) = \alpha\psi(x), \tag{7}$$

where $\alpha \in \mathbb{R} \cup \{+\infty\}$. We remark that

- the special case $\alpha = 0$ corresponds to a free motion and is sometimes called Kirchhoff boundary conditions,
- the case $\alpha = +\infty$ leads formally to the Dirichlet boundary conditions, i.e. $\psi(x_+) = \psi(x_-) = 0$, the line may be then regarded as a pair of two independent half lines with Dirichlet endpoints.

We observe from Equation (7) that the δ -interaction with parameter α is characterized by continuity of the wave function and discontinuity of its derivative such that the discontinuity is proportional to the wave function value with the factor of α .

4.2 δ -coupling

The δ -coupling is a generalization of the δ -interaction for vertices of degree $n > 2$. It is described by the relations

$$\psi_j(0) = \psi_k(0) =: \psi(0), \quad j, k \in \hat{n}, \quad \sum_{j=1}^n \psi'_j(0) = \alpha\psi(0) \quad (8)$$

for $\alpha \in \mathbb{R} \cup \{+\infty\}$; the derivatives are taken in the outgoing sense. Similarly as in the case of the δ -interaction, $\alpha = 0$ corresponds to a free motion and $\alpha = +\infty$ to Dirichlet boundary conditions, i.e. the vertex may be regarded as n independent vertices of degree 1 with Dirichlet endpoints. We stress that it is the continuity of the wave function in the vertex which characterizes the δ -coupling and the δ -interaction and make them important. Both of them, the δ -interaction and the δ -coupling, are sometimes called *regular* in order to be distinguished from *singular* couplings that do not have this property of continuity.

4.3 Explicit boundary conditions expressing the δ -coupling

To demonstrate the use of the various forms of boundary conditions, let us show how (8) can be expressed in the form (1) in the three unique forms that we have introduced above, i.e. in (3), (4) and (6).

1. The form $(U - I)\Psi(0) + (U + I)\Psi'(0) = 0$

The unitary matrix U needed for expressing the δ -coupling in the form (3) has the form $\frac{2}{n+i\alpha}J - I$, where J denotes the $n \times n$ matrix whose all entries are equal to one. Hence $A = U - I = \frac{2}{n+i\alpha}J - 2I$, $B = i(U + I) = \frac{2i}{n+i\alpha}J$, i.e.,

$$\left(\frac{2}{n+i\alpha}J - 2I \right) \Psi(0) + \frac{2i}{n+i\alpha}J\Psi'(0) = 0;$$

note that none of the elements of A, B vanishes.

2. The ST -form

The matrices S, T in the ST -form of δ -coupling with parameter α is characterized by

$$m = 1, \quad S = (\alpha) \in \mathbb{C}^{1,1}, \quad T = (1 \ 1 \ \cdots \ 1) \in \mathbb{C}^{1,n-1}.$$

Therefore, the boundary conditions are written as

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \Psi'(0) = \begin{pmatrix} \alpha & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix} \Psi(0).$$

3. The projector form

It is easy to show that to express the δ -coupling with parameter α in the projector form (6), one has to set \mathcal{P} as the orthogonal projector onto $(1 \ 1 \ \cdots \ 1)^\perp$, i.e., $\mathcal{P} = \frac{1}{\sqrt{n(n-1)}}(J - nI)$, \mathcal{Q}

as the orthogonal projector onto $(1 \ 1 \ \cdots \ 1)$, i.e., $Q = \frac{1}{\sqrt{n}}J$, and $LQ = -\frac{\alpha}{\sqrt{n}}I$. When we express the boundary conditions explicitly, we obtain the system of $2n$ equations

$$\begin{aligned} \frac{1}{\sqrt{n(n-1)}}(J - nI)\Psi(0) &= 0, \\ \frac{1}{\sqrt{n}}J\Psi'(0) - \frac{\alpha}{\sqrt{n}}I\Psi(0) &= 0. \end{aligned}$$

5 Approximations of singular vertex couplings in quantum graphs

The wave function coupling in the vertices is an essential component of quantum graph models. The most significant type, the δ -coupling together with its special case of free matching, is defined by the continuity of the wave function in the vertex. Let us take note of an important fact, namely that the δ -coupling has a simple physical meaning:

Theorem 5.1 (Ex96b). *Let $H_\alpha(V)$ denote the Hamiltonian of a particle on a star graph with n infinite edges supporting potentials V_1, \dots, V_n and with the δ -coupling with parameter α in the center. Suppose that $V_j \in L^1_{\text{loc}}(\mathbb{R}^+)$ are below bounded and $W_j \in L^1(\mathbb{R}^+)$ for $j = 1, \dots, n$. Let us define the scaled potentials*

$$W_{\epsilon,j}(x) := \frac{1}{\epsilon}W_j\left(\frac{x}{\epsilon}\right), \quad j = 1, \dots, n.$$

Then

$$H_0(V + W_\epsilon) \rightarrow H_\alpha(V) \quad \text{as } \epsilon \rightarrow 0_+$$

in the norm-resolvent sense, where $\alpha = \sum_{j=1}^n \int_0^{+\infty} W_j(x)dx$.

In other words, the δ -coupling may be considered as a limit case of properly scaled regular potentials in the norm-resolvent sense, similarly as the Dirac delta function is a limit case of properly scaled regular functions in the sense of distributions. For this reason we call the δ -coupling a regular vertex coupling.

However, the subfamily of δ -couplings, parametrized by the value of its parameter, represents only a tiny subset of the whole family of vertex couplings. To a major part of the couplings, although they have been mathematically well defined for a decade, no simple explanation has been given for a long time, or even no explanation at all. The question whether and in what sense one can understand and/or approximate all the singular couplings by regular ones depending on suitable parameters had become a longstanding open problem. In this section we give the answer.

The key idea how we attempt the problem comes from a paper published by Cheon and Shigehara eleven years ago [CS98], in which the authors showed that a combination of regular point interactions on a line approaching each other with the coupling parameters scaled in a particular way with respect to the interaction distance can produce a singular point interaction.

5.1 The approximation arrangement

Our aim is to explain the meaning of vertex couplings using suitable approximations. We express the matching conditions in the vertex in the ST -form,

$$\begin{pmatrix} I^{(m)} & T \\ 0 & 0 \end{pmatrix} \Psi'(0) = \begin{pmatrix} S & 0 \\ -T^* & I^{(n-m)} \end{pmatrix} \Psi(0); \quad (9)$$

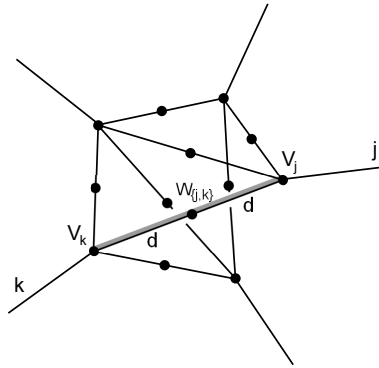


Figure 1: The scheme of the approximation. All inner links are of length $2d$. Some connection links may be missing if the conditions given in the text are not satisfied. The quantities corresponding to the index pair $\{j, k\}$ are marked, and the grey line symbolizes the vector potential $A_{(j,k)}(d)$.

note that the sensitivity of this formulation to the edge numbering mentioned in Section 3.1 plays no role, since one may rename the edges if necessary. As we will see below, the ST -form will help us significantly to simplify the design and tuning of the approximation arrangement.

We are going to show that the star graph with the singular vertex coupling given by the boundary conditions (9) may be understood as a limit case of certain family of graphs constructed only from edges connected by δ -couplings, δ -interactions, and supporting constant vector potentials. The detailed description of our approximating model follows, cf. Fig. 5.1.

- We take n halflines, each parametrized by $x \in [0, +\infty)$, with the endpoints denoted as V_j , and put a δ -coupling (to the edges specified below) with the parameter $v_j(d)$ at the point V_j for all $j \in \hat{n}$. The exact values of $v_j(d)$ are determined by relatively simple functions of the matrix entries of S and T in the boundary conditions (9).
- Certain pairs V_j, V_k of halfline endpoints will be joined by edges of the length $2d$, and the center of each such joining segment will be denoted as $W_{\{j,k\}}$. The pairs $\{j, k\}$ for which the points V_j and V_k , $j \neq k$, are joined, are determined by the positions of zeros in the matrices S and T .
- At each point $W_{\{j,k\}}$ we place a δ -interaction with a parameter $w_{\{j,k\}}(d)$ which is again given as a relatively simple function of the entries of S and T from (9). For convenience we regard the connecting edges of the length $2d$ as being composed of two line segments of the length d , on each of them the variable runs from 0 (corresponding to the point $W_{\{j,k\}}$) to d (corresponding to the point V_j or V_k).
- On each connecting segment described above we put a vector potential $A_{(j,k)}(d)$ which is constant on the whole line between the points V_j and V_k and determined mainly by phases of the entries of S and T .

The expressions for $v_j(d)$, $w_{\{j,k\}}(d)$ and $A_{(j,k)}(d)$ and the dependence on the parameter d are calculated in [CET10a] by expanding the wave function components into Taylor polynomials.

5.2 The norm-resolvent convergence

According to the previous sections, any vertex coupling in the center of a star graph may be regarded as a limit of a certain family of graphs supporting nothing but δ -couplings, δ -interactions and constant vector potentials. Let us take a more detailed look at the question of the convergence of the approximating scheme.

Let us denote the Hamiltonian of the star graph Γ with the coupling (9) at the vertex as H^{Ad} (referring to the *approximated* system), and H_d^{Ag} will stand for the *approximating* family of graphs that has been constructed above.

Our aim is to establish a convergence of H_d^{Ag} to H^{Ad} for $d \rightarrow 0_+$. Both H^{Ad} and H_d^{Ag} are unbounded operators, therefore we will examine and compare their resolvents. Let the symbols $R^{\text{Ad}}(k^2)$ and $R_d^{\text{Ag}}(k^2)$ denote the resolvents of H^{Ad} and H_d^{Ag} at the points k^2 from the resolvent set. Needless to say, the operators act on different spaces: $R^{\text{Ad}}(k^2)$ on $L^2(G)$, where $G = (\mathbb{R}^+)^n$ corresponds to the star graph Γ , and $R_d^{\text{Ag}}(k^2)$ on $L^2(G_d)$, where

$$G_d = (\mathbb{R}^+)^n \oplus (0, d)^{\sum_{j=1}^n N_j},$$

where the symbol N_j stands for the subset of $\{1, 2, \dots, n\}$ containing indices of all the edges that are joined to the j -th one by a connecting segment. In order to be able to compare the resolvents $R^{\text{Ad}}(k^2)$ and $R_d^{\text{Ag}}(k^2)$, we need to identify $R^{\text{Ad}}(k^2)$ with the orthogonal sum

$$R_d^{\text{Ad}}(k^2) = R^{\text{Ad}}(k^2) \oplus 0,$$

where 0 is a zero operator acting on the space $L^2\left((0, d)^{\sum_{j=1}^n N_j}\right)$ which is removed in the limit. Then both the operators $R_d^{\text{Ad}}(k^2)$ and $R_d^{\text{Ag}}(k^2)$ are defined as acting on functions from $L^2(G_d)$ which are vector functions with $n + \sum_{j=1}^n N_j$ components; we will index the components by the set

$$\mathcal{I} = \hat{n} \cup \{(l, h) \mid l \in \hat{n}, h \in N_l\}.$$

Let us now use this setting to state the main theorem of the paper.

Theorem 5.2. *The family H_d^{Ag} converges to H^{Ad} in the norm-resolvent sense as $d \rightarrow 0_+$, i.e.*

$$\lim_{d \rightarrow 0_+} \left\| R_d^{\text{Ag}}(k^2) - R_d^{\text{Ad}}(k^2) \right\| = 0,$$

where $\|\cdot\|$ is the L^2 -norm in G_d .

The main tool used in the proof is the Krein's formula. For details see [CET10a].

As the Hamiltonian of the approximating system converges to the Hamiltonian of the approximated system *in the norm-resolvent sense*, it has the natural consequences for the convergence of eigenvalues, eigenfunctions, etc. Since for some applications it may be useful to know how quickly the term $\left\| R_d^{\text{Ag}}(k^2) - R_d^{\text{Ad}}(k^2) \right\|$ approaches zero if $d \rightarrow 0_+$, we add the following remark.

Remark 5.3. *(The rate of convergence.)*

If the assumptions of Theorem 5.2 are satisfied, there is a number $K > 0$ independent of d such that

$$\left\| R_d^{\text{Ag}}(k^2) - R_d^{\text{Ad}}(k^2) \right\| \leq K \cdot \sqrt{d}.$$

References

- [AGHH05] S. Albeverio, F. Gesztesy, R. Høegh-Krohn and H. Holden: *Solvable Models in Quantum Mechanics*, 2nd edition, AMS Chelsea (2005).
- [CE04] T. Cheon and P. Exner: An approximation to delta' couplings on graphs, *J. Phys. A: Math. Gen.* **37** (2004) L329–335.
- [CET10a] T. Cheon, P. Exner, O. Turek: Approximation of a general singular vertex coupling in quantum graphs, *Ann. Phys. (NY)* **325** (2010) 548–578.
- [CET10b] T. Cheon, P. Exner and O. Turek: Tripartite connection condition for a quantum graph vertex, *Phys. Lett. A* **375** (2010) 113–118.
- [CS98] T. Cheon and T. Shigehara: Realizing discontinuous wave functions with renormalized short-range potentials, *Phys. Lett.* **A243** (1998) 111–116.
- [ENZ01] P. Exner, H. Neidhardt and V.A. Zagrebnov: Potential approximations to δ' : an inverse Klaunder phenomenon with norm-resolvent convergence, *Commun. Math. Phys.* **224** (2001) 593–612.
- [EŠ89] P. Exner and P. Šeba: Free quantum motion on a branching graph, *Rep. Math. Phys.* **28** (1989) 7–26.
- [ET06] P. Exner, O. Turek: Approximations of permutation-symmetric vertex couplings in quantum graphs, in the *Proc. of the NSF Research Conference “Quantum Graphs and Their Applications”*, Snowbird (2005), AMS “Contemporary Mathematics” Series, vol. 415, Providence, R.I. (2006) 109–120.
- [ET07] P. Exner, O. Turek: Approximations of singular vertex couplings in quantum graphs, *Rev. Math. Phys.* **19** (2007), 571–606.
- [Ha00] M. Harmer: Hermitian symplectic geometry and extension theory, *J. Phys. A: Math. Gen.* **33** (2000) 9193–9203.
- [KS99] V. Kostykin, R. Schrader: Kirchhoff's rule for quantum wires, *J. Phys. A: Math. Gen.* **32** (1999) 595–630.
- [KS00] V. Kostykin, R. Schrader: Kirchhoff's rule for quantum wires. II: The Inverse Problem with Possible Applications to Quantum Computers, *Fortschr. Phys.* **48** (2000) 703–716.
- [Ku04] P. Kuchment: Quantum graphs I. Some basic structures, *Waves Random Media* **14** (2004) S107–S128.
- [RS53] K. Ruedenberg and C.W. Scherr: Free-electron network model for conjugated systems, I. Theory, *J. Chem. Phys.* **21** (1953) 1565–1581.
- [SMMC99] T. Shigehara, H. Mizoguchi, T. Mishima, T. Cheon: Realization of a four parameter family of generalized one-dimensional contact interactions by three nearby delta potentials with renormalized strengths, *IEICE Trans. Fund. Elec. Comm. Comp. Sci.* **E82-A** (1999) 1708–1713.
- [Tu09] O. Turek: *Schrödinger Operators on Metric Graphs*, PhD thesis, Czech Technical University in Prague (2009).

量子推定理論による測定誤差の不確定性関係の 定式化とその量子情報幾何における意味付け

渡辺 優, 沙川 貴大, 上田 正仁
(東京大学 大学院理学系研究科 物理学専攻)

2011年7月12日

1 はじめに

1927年に Heisenberg [1] はガンマ線顕微鏡による粒子の位置の測定の思考実験を考察することにより、位置 \hat{x} の測定誤差 $\varepsilon(\hat{x})$ と測定による運動量 \hat{p} への擾乱 $\eta(\hat{p})$ の間にトレードオフ関係

$$\varepsilon(\hat{x})\eta(\hat{p}) \gtrsim h \quad (1)$$

を見出した。この不確定性関係から、位置と運動量の同時測定における測定誤差についても同様に

$$\varepsilon(\hat{x})\varepsilon(\hat{p}) \gtrsim h \quad (2)$$

が成り立つことが期待される。当時、量子測定についての理論はまだ完成していなかったため、同時期に発見された Robertson [3] および Kennard [2] らによる量子揺らぎについてのトレードオフ関係

$$\Delta X_1 \Delta X_2 \geq \frac{1}{2} |\langle [\hat{X}_1, \hat{X}_2] \rangle| \quad (3)$$

が Heisenberg の不確定性関係の数学的定式化であると誤って認識された。ただし、ここで $\langle \hat{X} \rangle := \text{Tr}[\hat{\rho}\hat{X}]$ は量子状態 $\hat{\rho}$ における \hat{X} の期待値、 $(\Delta X)^2 := \langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2$ は \hat{X} の量子揺らぎ、 $[\hat{X}_1, \hat{X}_2] := \hat{X}_1\hat{X}_2 - \hat{X}_2\hat{X}_1$ は交換関係を表す。しかし、量子揺らぎそのものは測定過程に依存しない量であり、したがって Robertson-Kennard の不等式は量子測定の誤差についてのトレードオフ関係ではない。Heisenberg に代表される測定過程についての不確定性関係は相補性 (complementarity) と呼ばれる [4]。一方、Robertson-Kennard の不等式に代表される、

測定過程に依存しない、対象となる量子系そのものの性質のみで表される不確定性関係は非決定性 (indeterminacy) と呼ばれる [1]。

Heisenberg 以降、様々な測定モデルについて不確定性関係が示されたが、どれも等号を達成できない緩いトレードオフ関係であったり [5, 6, 7, 8, 9]、もしくは一般の量子測定については成立しない不等式であったりした [10, 11, 12]。近年の量子制御技術の発展によって、量子系に対する精密な測定が実現できるようになってきた [13, 14, 15, 16]。したがって、非可換な物理量の測定誤差についての達成可能な限界はどこなのかを明らかにする必要がある。

量子測定の誤差はこれまで測定値の分散によって議論されてきた。しかし、測定値の分散が小さくても、測定によって得られた情報量が多いとは一般に言えず、測定値の分散それ自身は測定の精度を表す指標にはならない。量子測定の相補性は、測定値の分散についてのトレードオフ関係ではなく、測定によって得られる互いに非可換な物理量についての情報量の間トレードオフ関係が存在することを示している。1970 年頃、量子測定理論が Kraus [17] や Davies [18] らによって構築される一方、Helstrom [19] や Holevo [20] らによって量子推定理論の基礎が研究された。量子推定理論は、例えば、未知の量子状態が与えられた場合、どのような測定を行えば量子状態を効率よく推定できるかを明らかにする。推定精度が高いほど量子測定によって得られる情報量は多く、逆に推定精度が低い場合には測定によって得られる情報量は少ない。したがって、測定によって得られる情報量は、測定値の分散ではなく、推定値の精度によって与えられる。

古典情報理論では情報幾何 [21] と呼ばれる幾何学的手法によって様々な最適性が統一的に議論されている。情報幾何の量子論への拡張として量子情報幾何があり、量子情報理論の幾何学を用いた統一的な理解につながると考えられている。量子情報理論では測定プロセスが欠くことのできない要素である。量子測定の相補性によるトレードオフ関係を明かにすることによって、量子情報幾何における量子測定のもたらす性質が明らかになると考えられる。

本論文では、第 2 節で測定誤差を推定理論で用いられる Fisher 情報量によって定式化する。第 3 節では、文献 [22] で示された、測定誤差の間に成立するトレードオフ関係を紹介する。そして、第 4 節では、第 3 節で示されたトレードオフ関係の情報幾何的な意味について議論する。

2 測定誤差の定式化

有限次元 Hilbert 空間 \mathcal{H} 上の、未知の同一な量子状態 $\hat{\rho}$ が n 個与えられたとする。それぞれの量子状態に対して同じ一般の量子測定 (POVM 測定) $\mathbf{M} = \{\hat{M}_i\}_{i=1}^m$ を行い、その測定結果から物理量 \hat{X} の期待値 (\hat{X}) を推定するとする。もし、量子測定が \hat{X} のスペクトル分解

$\hat{X} = \sum_i \alpha_i \hat{P}_i$ に対応した射影測定 $\mathbf{P} = \{\hat{P}_i\}$ であれば、 $\langle \hat{X} \rangle$ の推定値 X^{est} を

$$X^{\text{est}}(n_1, \dots, n_d) = \sum_i \alpha_i \frac{n_i}{n} \quad (4)$$

で求めることが最適である。ただし、 α_i は \hat{X} の固有値であり、 n_i は測定結果 i が得られた回数である ($\sum_i n_i = n$)。しかし、射影測定でない測定を行った場合に、どのようにして測定結果から推定値を計算すればよいかは非自明である。測定結果 $\{n_i\}_{i=1}^m$ から推定値への関数 X^{est} は推定量と呼ばれる。推定量はどのような関数でもよいわけではなく、例えば、どのような測定結果に対しても常に同じ値を返すものは好ましくない。推定量が満たすべき条件としてここでは漸近的一致性を要求する。

$$\lim_{n \rightarrow \infty} \text{Prob}(|X^{\text{est}} - \langle \hat{X} \rangle| > \delta) = 0, \quad \text{for all } \hat{\rho}, \text{ and } \delta > 0 \quad (5)$$

これは、サンプル数 n が無限の極限で必ず正確な推定が行えることを意味する。射影測定の場合の推定量 (4) はこの条件を満たしている。測定 \mathbf{M} において $\langle \hat{X} \rangle$ についての漸近的一致推定量が一つでも存在する場合、一般に漸近的一致推定量は無数に存在する。推定量の推定精度は推定値の分散 $\text{Var}[X^{\text{est}}]$

$$\mathbb{E}[X^{\text{est}}] := \sum_{\{n_i\}} p(n_1, \dots, n_m) X^{\text{est}}(n_1, \dots, n_m) \quad (6)$$

$$\text{Var}[X^{\text{est}}] := \mathbb{E}[(X^{\text{est}})^2] - \mathbb{E}[X^{\text{est}}]^2 \quad (7)$$

で量られる。ここで、合計は全ての $\sum_i n_i = n$ を満たす $\{n_i\}$ について取り、 $p(n_1, \dots, n_m)$ は測定結果 i をそれぞれ n_i 回得る確率、

$$p(n_1, \dots, n_m) := n! \prod_i \frac{p_i^{n_i}}{n_i!}, \quad (8)$$

$p_i := \text{Tr}[\hat{\rho} \hat{M}_i]$ は測定結果 i を得る確率である。推定量の分散 $\text{Var}[X^{\text{est}}]$ は量子揺らぎ、測定誤差、推定誤差の3つによって構成される。したがって、推定量の分散から量子揺らぎと推定誤差に起因する部分を取り除くことで定量化される。推定誤差は測定結果から推定値を求める際のアロリズムの非効率さによって引き起こされる。例えば、 n 個の測定結果のうち半分のデータからのみ推定を行う場合には当然推定値の精度は下がり、分散は大きくなる。推定誤差は、最尤推定量のように、以下に述べる Cramér-Rao 不等式の等号を達成する推定量を用いることで取り除くことが出来る。

式 (5) より、漸近的一致推定量の分散は漸近的にゼロに収束する。

$$\lim_{n \rightarrow \infty} \text{Var}[X^{\text{est}}] = 0 \quad (9)$$

したがって、推定量の最適性はこの収束のスピードによって決定される。一般に推定量の分散は n^{-1} で減少するため、収束スピードは $n \text{Var}[X^{\text{est}}]$ で評価される。この収束スピードの上限を与える不等式として、以下の Cramér-Rao 不等式が知られている [23]。

$$\lim_{n \rightarrow \infty} n \text{Var}[X^{\text{est}}] \geq \begin{cases} \mathbf{x} \cdot J(\mathbf{M})^{-1} \mathbf{x} & (\mathbf{x} \in \text{Supp}[J(\mathbf{M})]) \\ +\infty & (\text{otherwise}) \end{cases} \quad (10)$$

ここで、 $J(\mathbf{M})$ は Fisher 情報行列と呼ばれ、その ij 成分は、量子状態を記述するパラメタ $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{d^2-1})^T$ を用いて ($d = \dim \mathcal{H}$)

$$[J(\mathbf{M})]_{ij} := \sum_k p_k (\partial_i \log p_k) (\partial_j \log p_k) \quad (11)$$

と定義される。ただし、 $\partial_i := \partial / \partial \theta_i$ である。また、実ベクトル $\mathbf{x} \in \mathbb{R}^{d^2-1}$ は

$$x_i = \partial_i \langle \hat{X} \rangle \quad (12)$$

で定義される。Fisher 情報行列は準正定値行列であり、ゼロ固有値を持つことがある。Fisher 情報行列がゼロ固有値を持つ場合、 $J(\mathbf{M})^{-1}$ は一般逆行列 (Moore-Penrose pseudoinverse) で定義される。一般逆行列は

$$A^{-1}A = A^T(A^T)^{-1} = P_{\text{Supp}[A]}, \quad AA^{-1} = (A^T)^{-1}A^T = P_{\text{Im}[A]} \quad (13)$$

を満たす。ただし、 $P_{\text{Supp}[A]}$ および $P_{\text{Im}[A]}$ はそれぞれ A のサポートへの射影、 A の像への射影を表す。以降ゼロ固有値を含む行列や非正方行列の逆行列はこの一般逆行列で定義する。Cramér-Rao 不等式の右辺が発散するとは、漸近的一致推定量が存在しないことを表す。例えば、物理量 \hat{X} と非可換な物理量のスペクトル分解に対応する射影測定を行った場合、無限のサンプルがあったとしてもその測定結果からは $\langle \hat{X} \rangle$ の値を正確に推定することはできない。実際、このような測定では \mathbf{x} は $J(\mathbf{M})$ のサポートに含まれていない。一方、Cramér-Rao 不等式の右辺が有限の場合、最尤推定量などのように Cramér-Rao 不等式の等号を達成する推定量が一般に存在し、それらを、推定誤差を含まない、最適な推定量とする。したがって、式 (10) の右辺は推定誤差を含まず、量子揺らぎと測定誤差の合計である。

式 (10) の右辺は $\boldsymbol{\theta}$ のとり方に依らないので、以下では量子状態を Lie 代数 $\mathfrak{su}(d)$ の生成子 $\hat{\boldsymbol{\lambda}} = \{\hat{\lambda}_1, \dots, \hat{\lambda}_{d^2-1}\}$ を用いて

$$\hat{\rho} = d^{-1} \hat{I} + \sum_i \theta_i \hat{\lambda}_i = d^{-1} \hat{I} + \boldsymbol{\theta} \cdot \hat{\boldsymbol{\lambda}} \quad (14)$$

と展開する。ここで、 \hat{I} は恒等演算子であり、生成子 $\hat{\boldsymbol{\lambda}}$ は

$$\hat{\lambda}_i^\dagger = \hat{\lambda}_i, \quad \text{Tr}[\hat{\lambda}_i] = 0, \quad \text{Tr}[\hat{\lambda}_i \hat{\lambda}_j] = \delta_{ij} \quad (15)$$

を満たす。物理量 \hat{X} や POVM 測定 \mathbf{M} の要素も同じ基底で

$$\hat{X} = x_0 \hat{I} + \mathbf{x} \cdot \hat{\boldsymbol{\lambda}}, \quad \hat{M}_i = r_i \hat{I} + \mathbf{v}_i \cdot \hat{\boldsymbol{\lambda}} \quad (16)$$

と展開すれば、期待値 $\langle \hat{X} \rangle$ および確率 p_i は

$$\langle \hat{X} \rangle = x_0 + \mathbf{x} \cdot \boldsymbol{\theta}, \quad p_i = r_i + \mathbf{v}_i \cdot \boldsymbol{\theta} \quad (17)$$

と表される。これらを用いると、Fisher 情報行列およびその一般逆行列は

$$J(\mathbf{M}) = \sum_i p_i^{-1} \mathbf{v}_i \mathbf{v}_i^T = V P^{-1} V^T, \quad (18)$$

$$J(\mathbf{M})^{-1} = (V^T)^{-1} [P - P A (A^{-1} P A)^{-1} A^{-1} P] V^{-1} \quad (19)$$

と求められる。ここで T は転置を表し、行列 V および P はそれぞれ

$$V = (\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_m), \quad P = \text{diag}(p_1, \dots, p_m), \quad (20)$$

で定義される $(d^2 - 1) \times m$ および $m \times m$ 行列であり、また、行列 A は

$$V^{-1} V = I - A A^{-1}, \quad A^{-1} A = I \quad (21)$$

を満たす行列である。つまり、行列 A は $V \mathbf{a} = \mathbf{0}$ を満たすベクトル \mathbf{a} のうち線形独立なもの全てを縦にして並べたものである。

Fisher 情報行列は POVM 測定を変化させることで変化するが、その測定に依存しない上限が知られている。その上限を与える不等式は量子 Cramér-Rao 不等式と呼ばれる [24]。

$$J(\mathbf{M}) \leq J_Q \quad (22)$$

$$\Rightarrow \mathbf{x} \cdot J(\mathbf{M})^{-1} \mathbf{x} \geq \mathbf{x} \cdot J_Q^{-1} \mathbf{x} \quad (23)$$

この行列の不等式は $J_Q - J(\mathbf{M})$ が準正定値行列であることを意味する。行列 J_Q は量子 Fisher 情報行列と呼ばれる。量子 Fisher 情報行列にはいくつもの種類があるが [25]、その中でも最も厳しい上限を与える対称対数微分 (symmetric logarithmic derivative, SLD) Fisher 情報行列をここでは用いることにする [26]。SLD Fisher 情報行列は、

$$[J_Q]_{ij} := \frac{1}{2} \text{Tr}[\hat{\rho} \{\hat{L}_i, \hat{L}_j\}] \quad (24)$$

で定義される。ここで $\{\cdot, \cdot\}$ は反交換関係を表し、 \hat{L}_i は SLD 作用素と呼ばれるエルミート作用素であり、以下の作用素についての線形方程式の解として定義される。

$$\partial_i \hat{\rho} = \frac{1}{2} \{\hat{\rho}, \hat{L}_i\} \quad (25)$$

SLD 作用素を $\hat{L}_i = a_i \hat{I} + \mathbf{b}_i \cdot \hat{\boldsymbol{\lambda}}$ と展開すれば、(25) より、

$$a_i = -\boldsymbol{\theta} \cdot \mathbf{b}_i, \quad \mathbf{b}_i = (G_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^T)^{-1} \mathbf{e}_i \quad (26)$$

が得られる。ただし、 \mathbf{e}_i は i 番目の成分のみが 1 である単位ベクトルであり、 $G_{\boldsymbol{\theta}}$ は

$$[G_{\boldsymbol{\theta}}]_{ij} := \frac{1}{2} \text{Tr}[\hat{\rho}\{\hat{\lambda}_i, \hat{\lambda}_j\}] \quad (27)$$

と定義される。したがって、SLD Fisher 情報行列は

$$J_Q = (G_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^T)^{-1} \quad (28)$$

と求められ、その逆行列は

$$\mathbf{x} \cdot J_Q^{-1} \mathbf{x} = (\Delta X)^2 \quad (29)$$

$$\mathbf{x} \cdot J_Q^{-1} \mathbf{y} = \frac{1}{2} \langle \{\hat{X}, \hat{Y}\} \rangle - \langle \hat{X} \rangle \langle \hat{Y} \rangle =: C_s(\hat{X}, \hat{Y}) \quad (30)$$

を満たす。ここで、 C_s は対称化された相関関数である。したがって、(23) は

$$\mathbf{x} \cdot J(\mathbf{M})^{-1} \mathbf{x} \geq (\Delta X)^2 \quad (31)$$

と表され、これは、どのような測定や推定量を用いても、 $\langle \hat{X} \rangle$ の推定の精度は量子揺らぎの逆数より良くなるならないということを意味する。なお、(31) の等号は \hat{X} のスペクトル分解に対応した射影測定によって達成される。つまり、 $\langle \hat{X} \rangle$ を知るためには射影測定を行い、その測定結果から (4) によって推定する、という手法が最適であることが示された。式 (31) の左辺は測定誤差と量子揺らぎに起因したので、測定誤差を以下で定義する。

$$\varepsilon(\hat{X}; \mathbf{M}) := \mathbf{x} \cdot J(\mathbf{M})^{-1} \mathbf{x} - (\Delta X)^2 \quad (32)$$

つまり、測定誤差を測定によって得られる Fisher 情報量 (の逆数) の最適な場合からの差で定義した。

3 測定誤差に関する不確定性関係

ここでは、前節で定義された測定誤差に対してどのような不確定性関係が成り立つかについて述べる。

定理 1. 任意の量子状態 $\hat{\rho}$, 物理量 \hat{X}_1, \hat{X}_2 , POVM 測定 \mathbf{M} について以下の不等式が成り立つ [22]。

$$\varepsilon(\hat{X}_1; \mathbf{M})\varepsilon(\hat{X}_2; \mathbf{M}) \geq \frac{1}{4} |\langle [\hat{X}_1, \hat{X}_2] \rangle|^2 \quad (33)$$

定理 1 より、どのような測定を行っても測定誤差が Heisenberg の不確定性関係を破ることは出来ない。これは前節で用いた SLD Fisher 情報行列ではなく RLD Fisher 情報行列を用いることで初めて導出された。しかし、文献 [22] でも指摘されているように、この不等式の等号は一般には達成されない。したがって、より厳しい不等式が測定誤差の積について課されることが予想される。

以下で述べる定理のために、次の記号を導入しておく。POVM 測定の全体を \mathcal{M}_{all} 、射影測定の全体を \mathcal{P} とする。また、2 つの物理量を確率的に射影測定する測定の全体を $\mathcal{M}_{\text{random}}$ 、それにノイズがのったものを $\mathcal{M}_{\text{noisy}}$ とする。

$$\mathcal{M}_{\text{random}} := \{q_1 \mathbf{P}_1 + q_2 \mathbf{P}_2 \mid \mathbf{P}_1, \mathbf{P}_2 \in \mathcal{P}, 0 \leq q_1, q_2 \leq 1, q_1 + q_2 = 1\} \quad (34)$$

$$\mathcal{M}_{\text{noisy}} := \{F\mathbf{M} \mid \mathbf{M} \in \mathcal{M}_{\text{random}}, F_{ij} \geq 0, \sum_j F_{ij} = 1\} \quad (35)$$

ただし、

$$q_1 \mathbf{P}_1 + q_2 \mathbf{P}_2 = \{q_1 \hat{P}_1^{(1)}, \dots, q_1 \hat{P}_d^{(1)}, q_2 \hat{P}_1^{(2)}, \dots, q_2 \hat{P}_d^{(2)}\} \quad (36)$$

$$F\mathbf{M} = \left\{ \sum_j F_{ij} \hat{M}_j \right\} \quad (37)$$

を意味する。行列 F は情報処理行列と呼ばれ、測定過程におけるノイズを表す。これら測定の集合の間に

$$\mathcal{P} \subset \mathcal{M}_{\text{random}} \subset \mathcal{M}_{\text{noisy}} \subset \mathcal{M}_{\text{all}} \quad (38)$$

が成り立つ。

次に、物理量の「本質的な」量子揺らぎおよび量子相関を次のように導入する。まず、 \hat{X}_1 および \hat{X}_2 の既約な同時不変部分空間を $\mathcal{H}_1, \mathcal{H}_2, \dots$ とする ($\oplus_a \mathcal{H}_a = \mathcal{H}$)。そして、量子状態 $\hat{\rho}$ のそれぞれの部分空間への射影を

$$\hat{\rho}_a := p_a^{-1} \hat{P}_a \hat{\rho} \hat{P}_a \quad (39)$$

とする。ただし、 \hat{P}_a は \mathcal{H}_a への射影、 $p_a = \text{Tr}[\hat{\rho} \hat{P}_a]$ とした。このとき、本質的量子揺らぎおよび量子相関を以下のように定義する。

$$(\Delta_Q \hat{X}_\mu) := \sum_a \left\{ \text{Tr}[\hat{\rho}_a \hat{X}_\mu^2] - \text{Tr}[\hat{\rho}_a \hat{X}_\mu]^2 \right\} \quad (\mu = 1, 2) \quad (40)$$

$$C_Q(\hat{X}_1, \hat{X}_2) := \sum_a \left\{ \frac{1}{2} \text{Tr}[\hat{\rho}_a \{\hat{X}_1, \hat{X}_2\}] - \text{Tr}[\hat{\rho}_a \hat{X}_1] \text{Tr}[\hat{\rho}_a \hat{X}_2] \right\} \quad (41)$$

仮に、1 つの物理量のみ、もしくは互いに可換な物理量のみに興味があるとき、それらを対角化する基底で量子状態を記述した際の非可換成分は物理に影響せず、その系は確率的に振る舞

う古典的な系と区別をつける必要はない。上で導入した「本質的な」量子揺らぎおよび物理量は、興味の対象となっている 2 つの物理量を同時ブロック対角化し、量子状態の非ブロック対角成分を無視したときの量子揺らぎおよび量子相関となっており、古典的な系では表れない、真に量子的な効果による量となっている。なお、 \hat{X}_1, \hat{X}_2 が同時ブロック対角化できない場合には、

$$\Delta_Q \hat{X}_\mu = \Delta \hat{X}_\mu, \quad C_Q(\hat{X}_1, \hat{X}_2) = C_s(\hat{X}_1, \hat{X}_2) \quad (42)$$

となり、また、 \hat{X}_2, \hat{X}_2 が可換な場合には

$$\Delta_Q \hat{X}_\mu = 0, \quad C_Q(\hat{X}_1, \hat{X}_2) = 0 \quad (43)$$

を満たす。

次の定理は、任意の測定 $\mathbf{M} \in \mathcal{M}_{\text{noisy}}$ について測定誤差の積が達成する限界を与える。

定理 2. 任意の量子状態 $\hat{\rho}$, 物理量 \hat{X}_1, \hat{X}_2 , POVM 測定 $\mathbf{M} \in \mathcal{M}_{\text{noisy}}$ について以下の不等式が成り立つ [22]。

$$\varepsilon(\hat{X}_1; \mathbf{M})\varepsilon(\hat{X}_2; \mathbf{M}) \geq (\Delta_Q X_1)^2 (\Delta_Q X_2)^2 - C_Q(\hat{X}_1, \hat{X}_2)^2 \quad (44)$$

また、この不等式の等号を達成する測定が存在する。

Proof. 任意の POVM 測定 \mathbf{M} および情報処理行列 F について、Fisher 情報行列の単調性 [21] より、

$$J(F\mathbf{M}) \leq J(\mathbf{M}) \quad \Rightarrow \quad \varepsilon(\hat{X}; F\mathbf{M}) \geq \varepsilon(\hat{X}; \mathbf{M}) \quad (45)$$

が成り立つため、任意の $\mathbf{M} \in \mathcal{M}_{\text{random}}$ について (44) を示せばよい。

いま、 $\mathbf{M} \in \mathcal{M}_{\text{random}}$ を、2 つの物理量 \hat{Y}_1, \hat{Y}_2 のスペクトル分解に対応した射影測定 $\mathbf{P}_1, \mathbf{P}_2$ をそれぞれ確率 q_1, q_2 で行う POVM 測定とする。このとき、(19) より、 $J(\mathbf{M})$ の逆行列は

$$\mathbf{y}_\nu \cdot J(\mathbf{M})^{-1} \mathbf{y}_\nu = (\Delta Y_\nu)^2 + \frac{1 - q_\nu}{q_\nu} (\Delta_Q Y_\nu)^2 \quad (46)$$

$$\mathbf{y}_1 \cdot J(\mathbf{M})^{-1} \mathbf{y}_2 = C_s(\hat{Y}_1, \hat{Y}_2) - C_Q(\hat{Y}_1, \hat{Y}_2) \quad (47)$$

と求められる。したがって、 $(d^2 - 1) \times 2$ 行列 $R := \begin{pmatrix} \mathbf{y}_1 & \mathbf{y}_2 \end{pmatrix}$ を用いて

$$\tilde{\varepsilon}(\mathbf{M}) := R^T (J(\mathbf{M})^{-1} - J_Q^{-1}) R \quad (48)$$

とすれば、

$$\tilde{\varepsilon}(\mathbf{M}) = \begin{pmatrix} \frac{q_2}{q_1} (\Delta_Q Y_1)^2 & -C_Q(\hat{Y}_1, \hat{Y}_2) \\ -C_Q(\hat{Y}_1, \hat{Y}_2) & \frac{q_1}{q_2} (\Delta_Q Y_2)^2 \end{pmatrix} \quad (49)$$

と表される。いま、測定したい物理量 \hat{X}_1, \hat{X}_2 を

$$\hat{X}_\mu = x_{\mu,0} \hat{I} + \mathbf{x}_\mu \cdot \hat{\boldsymbol{\lambda}} \quad (50)$$

とし、それらが

$$\hat{X}_\mu = \sum_{\nu=1,2} a_{\mu\nu} \hat{Y}_\nu \quad \Rightarrow \quad \mathbf{x}_\mu = \sum_{\nu=1,2} a_{\mu\nu} \mathbf{y}_\nu \quad (51)$$

と表されたとすると、測定誤差は、

$$\varepsilon(\hat{X}_\mu; \mathbf{M}) = [A\tilde{\varepsilon}(\mathbf{M})A^T]_{\mu\mu} \quad (52)$$

と求められる。ただし、 $A = (a_{\mu\nu})$ とした。したがって、

$$\begin{aligned} \varepsilon(\hat{X}_1; \mathbf{M})\varepsilon(\hat{X}_2; \mathbf{M}) &\geq \det[A\tilde{\varepsilon}(\mathbf{M})A^T] \\ &= \det \left[A \begin{pmatrix} (\Delta_Q Y_1)^2 & \mathcal{C}_Q(\hat{Y}_1, \hat{Y}_2) \\ \mathcal{C}_Q(\hat{Y}_1, \hat{Y}_2) & (\Delta_Q Y_2)^2 \end{pmatrix} A^T \right] \\ &= \det \begin{pmatrix} (\Delta_Q X_1)^2 & \mathcal{C}_Q(\hat{X}_1, \hat{X}_2) \\ \mathcal{C}_Q(\hat{X}_1, \hat{X}_2) & (\Delta_Q X_2)^2 \end{pmatrix} \\ &= (\Delta_Q X_1)^2 (\Delta_Q X_2)^2 - \mathcal{C}_Q(\hat{X}_1, \hat{X}_2)^2 \end{aligned} \quad (53)$$

となり、(44) が示された。また、等号成立条件は、 $\mathbf{a}_\mu = (a_{\mu 1}, a_{\mu 2})^T$ として、

$$\mathbf{a}_1 \cdot \tilde{\varepsilon}(\mathbf{M}) \mathbf{a}_2 = 0 \quad (54)$$

と表され、これを満たす物理量 \hat{Y}_1, \hat{Y}_2 および確率 q_1, q_2 は任意の \hat{X}_1, \hat{X}_2 および $\hat{\rho}$ について存在する。 ■

不等式 (44) は $\mathbf{M} \in \mathcal{M}_{\text{noisy}}$ に対してしか証明されておらず、一般の \mathcal{M}_{all} について成り立つかどうかは明らかになっていない。しかし、任意の測定が (44) を満たすことが数値的に示されている [22]。以下の予想に対して証明を与えることは今後の課題である。

Conjecture 3. 任意の POVM 測定 $\mathbf{M} \in \mathcal{M}_{\text{all}}$ に対して (44) が成り立つ。

この予想が正しければ、定理 2 において等号が達成可能であることが示されているので、条件式 (54) を満たす測定が 2 つの物理量 \hat{X}_1, \hat{X}_2 の測定誤差の積を最小にする測定である。文献 [22] では、2 つの異なる方向のスピンについて、最適な測定の冷却原子を用いた実装が提案されている。

4 量子情報幾何との関連性

情報幾何はパラメタ θ にしたがう確率分布の集合に対して、座標を導入し、その幾何学的な性質から情報理論や確率分布のダイナミクスを議論する。確率分布空間の計量は Fisher 情報行列で与えられ、これを Fisher 計量と呼ぶ。Fisher 計量は任意のマルコフ写像によって単調減少する [21]。

一方、量子情報幾何はパラメタ θ にしたがう量子状態の集合についての幾何学的な性質を議論する。量子状態空間の計量は量子 Fisher 情報行列によって与えられ、これを量子 Fisher 計量と呼ぶ。量子 Fisher 計量は任意の量子操作によって単調減少する [25]。

量子情報理論を量子情報幾何的にとらえるためには、量子測定による状態空間から確率分布空間への写像の性質を調べるのが重要である。量子測定による写像

$$M : \hat{\rho} \mapsto \{p_i = \text{Tr}[\hat{\rho} \hat{M}_i]\} \quad (55)$$

も量子操作であるので、マップされた確率分布空間の Fisher 計量は元の量子状態空間の量子 Fisher 計量と比べて減少する。これが量子 Cramér-Rao 不等式 (22) の幾何的な表現である。しかし、不確定性関係により量子 Cramér-Rao 不等式は一般に等号を達成できず、より強い制限がかかるはずである。我々の定義した測定誤差は

$$\varepsilon(\hat{X}; M) = \mathbf{x} \cdot [J(M)^{-1} - J_Q^{-1}] \mathbf{x} \quad (56)$$

であり、量子状態空間と確率分布空間の Fisher 計量のある種の差になっている。したがって、不確定性関係 (33) や (44) は、量子 Cramér-Rao 不等式よりも強い、量子状態空間と確率分布空間の Fisher 計量が満たさなければならない条件式になっていることがわかる。

量子推定理論を用いて不確定性関係を議論したことにより、各空間の計量についての関係が導かれた。今後、量子検定理論など他の問題設定における不確定性関係を議論することで、各空間のローカルではない大局的な性質についての関係も導けるのではないかと予想される。

参考文献

- [1] W. Heisenberg, *Zeitschrift für Physik* **43**, 172 (1927), English translation: J. A. Wheeler and H. Zurek, *Quantum Theory and Measurement* (Princeton Univ. Press, New Jersey, 1983), p. 62.
- [2] E. H. Kennard, *Z. Phys.* **44**, 326 (1927).
- [3] H. P. Robertson, *Phys. Rev.* **34**, 163 (1929).
- [4] N. Bohr, *Phys. Rev.* **48**, 696 (1935).

- [5] E. Arthurs and J. L. Kelly, *Bell Syst. Tech. J.* **44**, 725 (1965);
- [6] E. Arthurs and M. S. Goodman, *Phys. Rev. Lett.* **60**, 2447 (1988).
- [7] M. Ozawa, *Phys. Lett. A* **320**, 367 (2004).
- [8] H. P. Yuen and M. Lax, *IEEE Trans. Inf. Theory* **19**, 740 (1973).
- [9] H. Nagaoka, *Trans. Jpn. Soc. Indust. Appl. Math.* **1**, 305 (1991) (in Japanese), English translation: M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference: selected papers* (World Scientific, Singapore, 2005) p. 133.
- [10] T. Sagawa and M. Ueda, *Phys. Rev. A* **77**, 012313 (2008);
- [11] B.-G. Englert, *Phys. Rev. Lett.* **77**, 2154 (1996).
- [12] P. Busch and C. Shilladay, *Phys. Rep.* **435**, 1 (2006);
- [13] A. Kuzmich *et al.*, *Phys. Rev. A* **60**, 2346 (1999).
- [14] Y. Takahashi *et al.*, *Phys. Rev. A* **60**, 4974 (1999).
- [15] A. Kuzmich *et al.*, *Phys. Rev. Lett.* **85**, 1594 (2000).
- [16] T. Takano *et al.*, *Phys. Rev. Lett.* **102**, 033601 (2009).
- [17] K. Kraus, *Annals of Physics* **64**, 311 (1971).
- [18] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
- [19] C. W. Helstrom, *J. Stat. Phys.* **1**, 231 (1969).
- [20] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, (North-Holland Publishing Company, Amsterdam, 1982).
- [21] S. Amari, H. Nagaoka, *Methods of Information Geometry* (AMS, 2007).
- [22] Y. Watanabe, T. Sagawa, and M. Ueda, arXiv:1010.3571 (2010).
- [23] H. Cramér, *Mathematical Methods of Statistics* (Princeton University, Princeton, NJ, 1946).
- [24] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [25] D. Petz, *Linear Algebra Appl.* **244**, 81 (1996).
- [26] C. W. Helstrom, *Phys. Lett. A* **25**, 101 (1967).