

博士論文

論文題目：特異値間隔分布によるセルオートマトンの
分類

氏名：金子 勇治

目次

0	全体の序文	4
0.1	歴史的経緯	4
0.2	研究目的	5
0.3	提案手法	5
0.4	数値計算結果	6
0.5	ECA の特異値間隔分布のまとめ	7
0.6	ECA 以外の CA への適用	7
0.7	主結果	8
1	はじめに	9
1.1	ウルフラムによる分類	9
1.2	数値的な接近法と解析的な接近法	11
2	研究目的と方針	12
2.1	用語の定義	12
2.2	研究目的	13
2.3	研究方針	14
3	分類手法	16
3.1	特異値分解	16
3.2	ランダム行列	20
3.3	ランダム行列の性質（固有値分布）	20
3.4	ランダム行列の固有値間隔分布	22
3.5	固有値間隔分布を求める（数値計算の）際の注意点	22
3.6	相対周期パターン	24
4	数値実験	25
4.1	分類対象	25
4.2	実験結果	25
4.3	ECA の実験結果のまとめ	29
5	結果の分析	30
5.1	統計的検定	30
5.2	興味深い ECA1：ルール 18,146 の隠れた規則性	32
5.3	興味深い ECA2：ルール 150 の相対周期	33
5.4	興味深い ECA3：ルール 90 とルール 150 の関係	38
5.5	興味深い ECA4：ルール 170 の固有値分布	44
5.6	興味深い ECA5：ルール 73 の特異値間隔分布	51
5.7	ECA の特異値間隔分布のまとめ	51

6	ECA 以外の CA に於けるクラス 4 の探索	52
6.1	クラス 4 と 5 近傍 CA	52
6.2	クラス 4 と箱玉系	52
6.3	pBBS に対する数値計算方法	53
6.4	実験結果	54
6.5	実験結果のまとめ	56
6.6	結果に対する考察	56
7	アンフォールドされた特異値間隔の分布の利用	60
7.1	累積寄与率基準の欠点	60
7.2	スクリープロット基準	60
7.3	CST に於ける『飛び』の定量化	62
7.4	補足	63
8	本研究のまとめ	64
9	今後の課題	64
10	謝辞	65
付録 A	サンプルパターン	66
A.1	クラス 3 (ルール 18,22,30,45,60,90)	66
A.2	クラス 3 (ルール 105,106,150,122,129,146)	67
A.3	クラス 2 (ルール 154,170,73), クラス 4 ([54],110)	68
A.4	その他	69
付録 B	ウィグナー分布の導出	70
付録 C	ランダム行列理論と量子カオス論の歴史的な背景	71
C.1	原子核物理とランダム行列	71
C.2	量子可積分系とポアソン分布	71
C.3	量子カオスとビリャード系	71

Abstract

We investigate some of the Elementary Cellular Automaton (ECA) and the periodic Box Ball System (pBBS). The first part concerns the ECA. We review the Wolframs classification of ECA and quantization method of that classification introduced in [4] and give a new method based on the singular value decomposition. We first characterize class 3 which has chaotic behavior by comparing ones output patterns with random patterns (random matrix). We then give the criteria to compare them, which is based on spacing distribution of singular values. In the random matrix theory it is known that this distribution is approximated by the Wigner distribution and it has universality in the sense that it does not depend on details of the random matrix. We can use this property. And distinguish class 3 from class 4 which has complex behavior with localized structures like a soliton. This complexity is interested in computation theory. In the second part we apply the method proposed in the first part to the pBBS which is a soliton cellular automaton. We expect that the pBBS's spacing distribution has similar characters to those of class 4. By numerical simulation this expectation was confirmed and we found the Poisson-Wigner transition.

概要

我々はエレメンタリーセルオートマトン (ECA) と周期箱玉系を調べた。前半は ECA に関する部分である。我々は [4] に於ける ECA のウルフラムの分類とその定量化手法を概観し特異値分解に基づく新しい分類手法を導入する。まず、カオス的な挙動を示すクラス 3 をその出力パターンとランダムパターン (ランダム行列) との比較によって特徴付ける。それらを比較する基準は特異値間隔分布に基づくものとする。ランダム行列理論に於いて、この分布はウィグナー分布により近似され、ランダム行列の詳細によらない普遍性を持つことが知られている。我々はこの性質を用いてソリトンに類似した局所的構造を持ち複雑な挙動を生じるクラス 4 をクラス 3 と区別する。この複雑さは計算理論の観点から興味を持たれている。後半は前半の手法をソリトンセルオートマトンである周期箱玉形 (pBBS) に適用する。我々は pBBS の間隔分布がクラス 4 の間隔分布の特徴に類似すると予想する。我々は数値シミュレーションによってこの予想を確認した。そして、我々は pBBS に於ける Poisson-Wigner 遷移を発見した。

0 全体の序文

0.1 歴史的経緯

セルオートマトン (CA) とは、独立変数 (時間変数・空間変数) と従属変数 (状態変数) の両方が離散値で、かつ、従属変数の値域が有限集合の系である。フォン・ノイマンが生物の自己複製を数学的に定式化する際に用いたことで有名である。現代では交通流のモデリング等にも利用されている。

ウルフラムにより考案されたエレメンタリーセルオートマトン (ECA) は特に単純な CA で、ECA には 256 種類のルール (時間発展則) が存在する (各ルールは 0~255 の番号で呼ばれる)。その単純さにも拘わらず、極めて多様な時間発展パターンが得られる。ECA の分類で最も有名なものは、ウルフラムが数値実験を元に主張した次の分類 (ウルフラムの分類) であろう [1]。

Class1 Evolves to homogeneous state.

Class2 Evolves to simple separated periodic structure.

Class3 Yields chaotic aperiodic patterns.

Class4 Yields complex pattern of localized structure.

本分類が後の研究に与えた影響は非常に大きいが問題点も存在する。根本的な問題の 1 つはその恣意性であり、クラス 2 の simple, クラス 3 の chaotic 及び、クラス 4 の localized structure は厳密に定義されたものではない。

このような分類基準の恣意性を指摘し、定量的な分類を試みる研究は少なくない [10],[11],[12]。

CA では時間発展による数値誤差を考慮しなくてもよく、その点では信頼できるデータを比較的容易に得ることができる。そこで、これらのデータに対して 4 つのクラスの違いを反映できる統計量を構成する事で分類を定量化する。次に述べる解析的なアプローチと異なり、データがありさえすればどのような CA に対しても適用できる反面、時間発展をどこで打ち切るか等の恣意性が問題になる。例えば、ルール 90 はクラス 3 に属すが、その時間発展はセル数 $N = 2^m (m \geq 2)$ のとき $t = N/2$ で全てのセルの値が 0 になる。これはクラス 1 に他ならない。勿論、このようなシステムサイズ依存性が解析的に証明できる場合は特殊なシステムサイズを予め例外として扱うことも可能であろう。しかし、数値計算を行う背景には、そもそも解析的な扱いが非常に難しいから取敢えず出力パターンを観察して、それが解析的に扱うに足るルールか判断したいという動機もある。そのような状況で打ち切り時刻を予め指定すると興味深い挙動を見逃す恐れがある。現在のところ分類を達成する完全な統計量は発見されていない。

数値実験によらない解析的な研究としては、有限体を用いた代数的な手法と、統計力学等で発展した大自由度系の解析手法の応用が主に試みられている。有限体を用いた方法では、上に述べたルール 90 の例のように周期の長さの最大値等を具体的に求める事ができる場合がある [1]。しかし、線形ルールを除くと有限体で扱えるルールは少なく、分類の指標作りという観点からは致命的である。熱力学 (統計力学) 的な扱いについては、CA の挙動はセル数に強く依存する為、単純に熱力学的極限を取る事はできず、セル数無限大への飛ばし方を議論する必要がある。さらに、極限はセル数に対する大自由度極限の他に時間に対する長時間極限も考えられる。巨大パターンを扱うときに、どちらの極限を先に取りなのか、或いは別の極限の取り方をするのかといったことを考えなければならない。

0.2 研究目的

本研究ではウルフラムの分類が持つ恣意性の中でも、時間発展の打ち切り時刻の設定に対する恣意性と、パターンの分類に於ける目視の恣意性を問題にする。すなわち、本研究が行うべきことは、(数値計算で扱える)十分な時間を明確に定義し、その上で、ルールの挙動を(ウルフラムの基準に沿って)分類できる目視によらない定量的な指標を与える事である。

0.3 提案手法

我々はウルフラムと同じく、ECA の分類とは、その出力パターン U の分類であるという立場に立つ。本研究では特異値分解を用いてパターンを解析する。以下に特異値分解の定義を記す。尚、特異値分解は任意の実行列 X に対して使用可能である。

$${}^tX X \mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad X^t X \mathbf{u}_j = \lambda_j \mathbf{u}_j, \quad \mu_j = \sqrt{\lambda_j}, \quad \forall j \|\mathbf{u}_j\| = \|\mathbf{v}_j\| = 1$$

$$X = \mu_1 \mathbf{u}_1 {}^t\mathbf{v}_1 + \mu_2 \mathbf{u}_2 {}^t\mathbf{v}_2 + \cdots + \mu_r \mathbf{u}_r {}^t\mathbf{v}_r$$

上式右辺を X の特異値分解, $\mu_j (1 \leq j \leq r, r \text{ は } X \text{ の階数})$ を特異値という。特異値分解は基底となる長方形列の線形結合(特異値が重み)で X を表現する方法と考えることもできる。

まず、時間発展の打ち切り時刻 τ について考察する。予め定めた有限の時刻で時間発展を打ち切ることに対する恣意性を問題にするのであれば、時間発展を打ち切らないことはその解決法の1つである。しかし、計算機で無限の時間発展を行う事は不可能であるし、パターンを格納するメモリも有限である。

ここで、数値計算で使用するセル数が有限であることに注意すると時間発展は必ず有限の長さの周期状態(周期パターン U_c)に落ち込む。この周期の長さを T とする。

$$X = U/\sqrt{\tau}$$

とすれば、 $\tau \rightarrow \infty$ の極限で、その特異値は U_c/\sqrt{T} の特異値に一致する。これが我々が周期パターンに着目する理由である(ただし、周期境界条件による並進対称性を考慮して、 U_c そのものではなく相対周期パターン U_r を用いる)。

また、特異値分解の工学的な利用という文脈では、適当な画像データ等を X として特異値分解して、特異値の大きい部分に着目することで元の画像 X の特徴を抽出すること等に利用されている[4]。CA のパターンを分類する際にセンシティブな基準を選択すると、パターンの微妙な差異がクラスの違いとして認識される事がある。そこで、逆に特異値の小さい部分に着目することで、パターンの詳細によらない大まかな分類が可能になるのではないかと考えた。特異値の小さい部分も特異値と、基底の長方形列に分けられるが、本研究では特異値のみを分類に用いる。

本研究では、まず、chaotic と形容されたクラス3のパターンの複雑さをランダムパターンとの比較によって定量化する。ランダムパターンを行列とみなせばランダム行列であり、ランダム行列の特異値の間隔の分布は次のウィグナー分布 $P_{wigner}(s)$ で近似されることが知られている[5]。

$$P_{wigner}(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2} \quad (s \text{ は特異値の間隔})$$

0.4 数値計算結果

クラス 3 はルール 18,60,90,105,122,129,146,150 クラス 4 はルール 110 に対して計算を行った。代表的な結果のみ示す。左図は適当な同じ初期値から時間発展した各ルールのサンプルパターン、右図の実線はウィグナー分布、点が数値計算結果 $P(s)$ である。

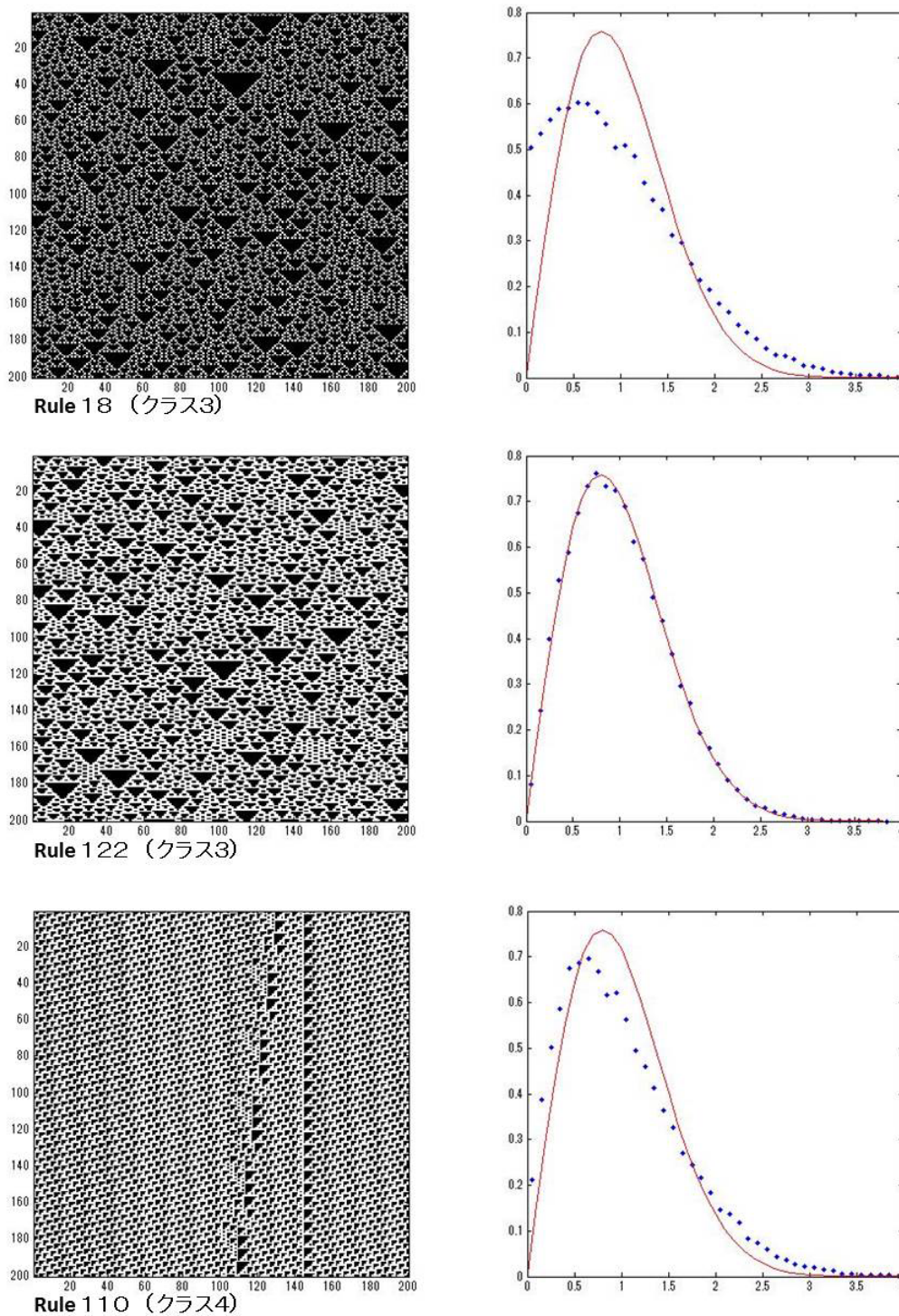


図1 ウィグナー分布（実線）と代表的な ECA の特異値間隔分布（点）の比較

0.5 ECA の特異値間隔分布のまとめ

数値計算を行った範囲で以下の結果を得た。但し、擬ウィグナー分布とはサイズ N のランダム行列の特異値間隔分布 ($N = 2, N \rightarrow \infty$ 以外では数値的に求める) で $N = 2$ のときはウィグナー分布である。

クラス 2 相対周期が非常に短いものが多いため U_r に代えて U_c を用いると、極端に周期が短いものを除けばポアソン分布が観察された。ただし、ルール 73 では domain wall と呼ばれる時間変化しないセルに挟まれた領域に於いて比較的複雑な時間発展が生じるため、 U_r はサンプリングできたが他のクラス 2 と同様に特異値の個数はかなり少ない。

クラス 3 特異値間隔分布は擬ウィグナー分布になる。しかし、ルール 18,146 ではシステムサイズが偶数のとき擬ウィグナー分布とは異なる分布が観察された。この分布の特徴は、特異値の縮退により $P(0) \simeq 0.5$ となっている点である。尚、この 2 つのルールでもセルの個数が奇数の場合は擬ウィグナー分布が観察された。

クラス 4 擬ウィグナー分布と異なり、かつ、ルール 18,146 でシステムサイズが偶数のときに見られた分布とも異なる間隔分布が観察された。原点付近の値は $P(0) \simeq 0$ である。

経験則の範囲で $P(s)$ の原点付近の挙動は tUU の持つ規則性に強く依存することが分かった。

ルール 170 (シフト写像) の周期パターン U_c (左右反転で巡回行列) の tU_cU_c はテープリッツ行列となり、その場合はポアソン分布が現れる。 $P(0) = 1$ である。

ルール 18,146 の相対周期パターン U_r はそれ自体に一見規則性はないように見えるが tU_rU_r の要素には 0 がチェッカーボード状に並ぶ。 $P(0) \simeq 0.5$ である。

ルール 90 の tU_cU_c は $P(0) \simeq 0.5$ である。ただし、この規則性は上の 2 つと異なり、直接 tU_cU_c を見て確認できる程ははっきりしたものではないが、相対周期が周期の 2 倍である。

以上より、相対周期パターンの特異値間隔分布は分類の指標として有効と思われる。ウルフラムが重要視するクラス 4 の探索に於いては、まず、原点付近で $P(0) \simeq 0$ であることを確認し、次に、統計的検定で擬ウィグナー分布ではないものを求めれば良いと思われる。

0.6 ECA 以外の CA への適用

ECA 以外でクラス 4 の特異値間隔分布を観察できる CA を探す試みた。ここで、研究者によってはウルフラムの言う localized structure を (擬) ソリトン [6] と呼ぶことがあることに注意する。ソリトンとは元々は古典可積分系に見られる粒子性を持つ孤立波のことである。

ソリトンの挙動を示す CA は、ソリトンセルオートマトンと呼ばれ、[7] を初めとして様々なものが考案されている。ただし、それらの多くは特定の初期値でソリトンのようなパターンが見られるに過ぎない。これに対して箱玉系と呼ばれる CA では、任意の初期値でソリトンのような時間発展パターンが観測できる [8]。周期境界条件を課した箱玉系を周期箱玉系 (pBBS) という。

この、localized structure とソリトンの類似性から、pBBS の U_r の特異値間隔分布がクラス 4 の特徴を示すのではないかと予想した。数値実験の結果、pBBS の特異値間隔分布は (保存量によるが) クラス 4 の特徴を示すことが分かった。また、保存量を変化させると、ランダム行列に見られるような Poisson-Wigner 遷移 [9] が観察された。

0.7 主結果

(I) セルオートマトン (CA) の相対周期パターンの特異値間隔分布を, CA の分類に用いることを提案し, エレメンタリーセルオートマトン (ECA) の分類を行った. その結果, ウルフラムの基準に沿った ECA の分類を, より詳細に分類する結果を得た.

(II) 特にルール 150 の特異値間隔分布は, サイズ無限大のランダム行列 (ガウス型直交アンサンブル) の特異値間隔分布に非常に良く一致することが分かった.

(III) ルール 150 の周期に関して, セルの個数を N とするとき, $N = 7 \cdot 2^m, m \in \mathbb{Z}_+$ では, すべての状態のうち, 相対周期がちょうど N になるものの割合は $m \rightarrow \infty$ の極限で 1 になることを証明した.

(IV) ECA の中でも特に興味深い性質を持つとされるクラス 4 に該当するルール探索を, ECA 以外の CA に対して行なった. その結果, 周期箱玉系 (pBBS) と呼ばれる CA のあるクラスが, クラス 4 の特異値間隔分布の特徴を持つ分布が観察されることが分かった.

(V) pBBS の保存量をパラメタとみなして, 保存量を変化させながら特異値間隔分布を描くと, ランダム行列理論に於ける Poisson-Wigner 遷移に類似した特異値間隔分布の遷移が見られることが分かった.

参考文献

- [1] S.Wolfram "Cellular Automata and Complexity" Addison-Wesley, Reading (1994)
- [2] 金子勇治 時弘哲治 間田潤 "エレメンタリーセルオートマトンの相互相関行列の解析" 九州大学応用力学研究所 研究集会報告 **24O-S3**, 107-112 (2012)
- [3] 金子勇治 "ランダム行列と周期箱玉系に於ける Poisson-Wigner 遷移" RIMS 講究録別冊 (2013 投稿中)
- [4] 柳井晴夫 竹内啓 "射影行列 一般逆行列 特異値分解" (UP 応用数学選書 10) 東京大学出版会 (1983)
- [5] 長谷川洋 "量子系の準位統計—量子カオス序論" (物理学最前線 28) 共立出版 (1991)
- [6] G.J.Martnez, A.Adamatzky, F.Chen, L.Chua "On Soliton Collision between Localization in Complex Elementary Cellular Automata: Rule 54 and 110 and Beyond" Complex Systems **21**, 117-142 (2012)
- [7] K.Park, K.Steiglitz and P.Thurston "Soliton like behavior in automata" Physica D **19**, 423-432 (1986)
- [8] D.Takahashi, J.Satsuma "A soliton cellular automaton" J.Phys.Soc.Jpn. **59**, 3514 (1991)
- [9] S.M.Nishigaki, M.Giordano, T.G.Kovács, F.Pittler "Critical statistics at the mobility edge of QCD Dirac spectra" arXiv:1312.3286v1 (2013)
- [10] W.Li, N.Packard "The structure of the elementary cellular automata rule space" Complex Systems **4**, 281-297 (1990)
- [11] S.Ninagawa "Power spectral analysis of elementary cellular automata" Complex Systems **17**, 399-411 (2008)
- [12] M.Schule, R.Stoop "A full computation-relevant topological dynamics classification of elementary cellular automata" Chaos **22**, 043143 (2012)

1 はじめに

セルオートマトン (CA) とは、独立変数 (時間変数・空間変数) と従属変数 (状態変数) の全てが離散値でかつ従属変数の値域が有限集合となっている系で、フォン・ノイマンが生物の自己複製を数学的に定式化する際に用いたことで有名である。現在では、交通流のモデリング手法 [1] 等としても利用されている。

CA の特徴は局所的に定義された単純なルールから、極めて多様な時間発展パターンが得られる点にある。CA の分類に関して、最も有名なものは 1.1 章に述べるウルフラムの分類であろう。

ウルフラムの分類がその後の CA 研究に与えた影響は非常に大きい、その分類基準は現象論的なもので、明確な定義に基づくものではない点が問題となっていた。例えば、クラス 3 と呼ばれる CA の時間発展則の集合は、その出力パターンがカオス的であるものとされている。しかし、空間変数と状態変数が有限個ならば必ず周期的な状態に落ち込むため、カオスではありえない。

これに対して本研究では、CA の出力パターンを特異値分解し、そこから得られた特異値の間隔の分布を分類指標として用いる事で ECA の定量的な分類を試みた。その結果、数値実験を行った範囲でウルフラムのクラスに沿った分類ができ、特に、クラス 3 の時間発展則では特異値間隔分布がランダム行列理論で知られるところのウィグナー分布によく一致する事が分かった。

本稿の構成は次の通りである。1 章でウルフラムの分類の説明を行い、2 章で研究目的を述べる。3 章で本研究で用いる分類手法を導入、4 節でその分類指標に基づく数値実験の結果を示す。5 章は結果の分析を、6 章は ECA 以外の CA に対する提案手法の適用結果を記す。7 章で CA の分類以外への本指標の応用について触れ、8 章で全体のまとめを行う。

1.1 ウルフラムによる分類

1.1.1 ウルフラムによる分類の目的

ウルフラムの分類を説明する前に、ウルフラムにとって分類の動機付けとなった問題に関して、ごく簡単に触れる。ライフゲームを例に取ろう。ライフゲームは空間 2 次元の CA で、グライダーと呼ばれるパターンが存在し、そのパターンは、あたかも粒子のように、離散化された時空を伝搬する。このことから、グライダーを用いた情報の伝達を考えることができる。

ところで、現在の計算機は NOT, OR, AND といった論理ゲートを電気回路を用いて構成し、種々の計算を実行しているが、上記のグライダー等のパターンを適当に組み合わせると、これらの論理ゲートを構成でき、現在の計算機が実行できる計算をライフゲームを用いて実行できることが証明されている [2]。

ウルフラムはライフゲームより単純な CA で同様の計算能力を有するものについて考察し、ECA と呼ばれる非常に単純な CA について詳しい数値実験を行なった結果、ECA のルール 110 がそれに該当すると予想した。この予想は、後にクックにより肯定的に証明された [3]。クックの証明において、やはりグライダーが重要な役割を果たす。彼は、ルール 110 で生じるグライダー (Appendix-A3 の右下の図に於いて縦・左下方向に走る黒い線がその例である) から、グライダーシステムと呼ばれるものを構成し、ルール 110 を用いた計算が可能であることを示した。

ここで重要なことは、ウルフラムがどのような基準に従って、ルール 110 が計算能力を有すると予想したかである。これに関しては、次節で ECA の導入を行った後で述べることにする。

1.1.2 ECA とウルフラムの分類方法

まず、ECA の定義を述べる。エレメンタリーセルオートマトン (ECA) とは 1 次元 2 値 3 近傍決定系であり、次の時刻のセルの値が前の時刻の自分自身とその両隣の値で決まる。尚、以下に『パターン』や『周期』といった語が登場するが、それらの定義は 2 章で改めて行う。

時刻 t の 3 つのセルの取り得る状態 8 通りに対し時刻 $t+1$ のセルの値はそれぞれ 2 通り考えられるので、ECA における時間発展則は $2^8 = 256$ 種類存在する。表は 110 番目の時間発展則 (ルール) である。110 というルールの番号は、表の時刻 $t+1$ の値 01101110 を 2 進数とみなし、それを 10 進数表示したものである。

ルール 110

1 1 1	1 1 0	1 0 1	1 0 0	0 1 1	0 1 0	0 0 1	0 0 0
0	1	1	0	1	1	1	0

次に、これらのルールを分類する。ウルフラムの分類方法では各ルールの出力パターンの分類をもって、ルールの分類とする。彼は、ランダムに生成した複数の初期状態から得られるパターンのアンサンブルを見たときに、その傾向が以下の 4 種類に大別できるとした [4]。

- (1) Evolves to homogeneous state.
- (2) Evolves to simple separated periodic structure.
- (3) Yields chaotic aperiodic patterns.
- (4) Yields complex pattern of localized structure.

具体的なパターンの例については Appendix-A を参照されたい。見やすくする為 $1=\square$, $0=\blacksquare$ でパターンを表示している。ただし、クラス 1 に関してはその様子が自明なため掲載していない。Appendix-A1・A2 はクラス 3 の、A3 はクラス 2・4 のパターンである。尚、A1~A3 では、A3 の右下以外は全て、同じ初期値から時間発展させたものである。

- (1)~(4) はクラス 1~4 と呼ばれる。ウルフラムは次のように力学系との対応を述べている [4]。

Continuous dynamical systems provide analogues for the class of behaviour seen in cellular automata. Class 1 cellular automata show limit points, while Class 2 cellular automata may be considered to evolve to limit cycles. Class 3 cellular automata exhibit chaotic behaviour analogous to that found with strange attractors. Class 4 cellular automata effectively have very long transients, and no direct analogue for them has been identified among continuous dynamical systems.

クラス 4 の特徴が CA の計算能力に関連するというウルフラムの直観は、結果的には正しかったのであるが (ルール 110 はクラス 4 に属す)、本分類は実質的に目視という極めて恣意的なもので、ウルフラムが、セルオートマトン研究の 20 の課題を列挙した際の 1 番目の問題は、

Problem 1

What overall classification of cellular automaton behaviour can be given?

である。彼自身、いくつかの定量的な基準を提案している。例えば、微小摂動に対する安定性に注目する方法では、ランダムな初期値 x と、 x と 1 セルのみ異なる初期値 y から得られる時間発展パターンをそれぞれ X, Y として、それらの排他的論理和をとり、初期値の違いがどのように伝搬するかを調べる。

その場合のクラス 1~4 の分類基準は以下の通りである。

- (1) Such differences in pattern die out.
- (2) They may persist, but remain localized.
- (3) They typically expand at an asymptotically constant rate.
- (4) Gives rise to a pattern of differences that typically expands irregularly with time.

しかし、結局は排他的論理和を取ったパターン (pattern of differences) を目視により分類している点では恣意性を除去できておらず、irregularly がどう特異的なのか明確な定義は無い。特に、有限の時間しかない数値計算で観測される漸近挙動が、どの程度系の性質を反映したものと言えるのか不明である。Appendix-A4 の上段に、本判定法によるクラス 3 と 4 の比較の例を載せておく。このような分類基準の恣意性を指摘し、定量的な分類を試みる研究は少なくない [5][39][40][41]。そのアプローチの仕方には大きく分けて以下の 2 つの方法が取られている。

1.2 数値的な接近法と解析的な接近法

CA では偏微分方程式の数値計算等と異なり、時間発展による数値誤差を考慮しなくてもよく、その点では信頼できるデータを比較的容易に得ることができる。そこで、これらのデータに対して 4 つのクラスの違いを反映できる統計量を構成することで分類を定量化する。次に述べる解析的なアプローチと異なり、データがありさえすればどのような CA に対しても適用できるが、時間発展をどこで打ち切るか等の恣意性が問題になる。例えば、ルール 90 はクラス 3 に属すが、その時間発展はセル数が $N = 2^m (m \geq 2)$ のとき $t = N/2$ で全てのセルの値が 0 になることが証明できる。これはクラス 1 に他ならない。勿論、この様なシステムサイズ依存性が解析的に証明できる場合は特殊なシステムサイズを予め例外として扱うことも可能であろう。しかし、数値計算を行う背景には、そもそも解析的な扱いが非常に難しいから取敢えず出力パターンを観察して、それが解析的に扱うに足るルールであるか判断したい、という動機もある。そのような状況で打ち切り時刻を予め指定すると興味深い挙動を見逃す恐れがある。現在のところ、分類を達成する完全な統計量は発見されていないが、代表的なものは [5] にまとめられている。

では、数値実験によらない解析的な研究はどうかというと、有限体を用いた代数的な手法と統計力学等で発展した大自由度系の解析手法の応用が主に試みられている。有限体を用いた方法では、上に述べたルール 90 の例のように周期の長さの最大値等を具体的に求める事ができる場合がある [4]。しかし、線形ルールを除くと有限体で扱えるルールは少なく、分類の指標作りという観点からは致命的である。熱力学 (統計力学) 的な扱いはウルフラムの著書にもこれからの課題として取り上げられており、セルオートマトン研究に於ける 6,17 番目の問題は次の通りである。

Problem 6

How does thermodynamics apply to cellular automata?

Problem 17

What is the nature of the infinite size limit for cellular automata?

興味深い研究として射影演算子を用いた研究 [6] 等があるが、形式的な議論に終始している印象を拭えない。また、CA の挙動はセル数に強く依存する為、単純に熱力学的極限を取る事はできず、セル数無限大への飛ばし方を議論する必要がある。さらに、極限はセル数に対する大自由度極限の他に時間に対する長時間極限も考えられる。巨大パターンを扱うときに、どちらの極限を先にするのか、或いは別の極限の取り方をするのかといったことを考えなければならない。

2 研究目的と方針

本節では初めに（1 節で定義なしに使用していた概念を含め）以降の議論で用いる基本的な用語を定義する．次に，1.2 節と 1.3 節の内容を踏まえ，研究目的を述べ，その目的を達成するため我々が取るべき方針について考察する．具体的な研究手法は第 3 章に記す．

2.1 用語の定義

- ・時間変数（時刻） $t \in \mathbb{Z}_{\geq 0}$ ，空間変数 $n \in \mathbb{N} (1 \leq n \leq N)$ ， t, n でのセルの値 $u_n^t \in \{0, 1\}$ ．
 - ・システムサイズ…セルの総数． N と記す．
 - ・パターン（軌道）… u_n^t を行列要素とする行列 U ．
 - ・停止時刻（観測時間）…時間発展を終える時刻． $t = T_F$ と書く．
 - ・状態… U の各行のこと．時刻 t の状態 $(u_1^t, u_2^t, \dots, u_N^t)$ を \mathbf{u}_t と表す．特に \mathbf{u}_0 を初期状態または初期値， \mathbf{u}_{T_F} を終状態と呼ぶ．また，全状態の集合を状態空間という．
 - ・時間発展則（ルール）…状態 \mathbf{u}_t に対して次の状態 \mathbf{u}_{t+1} を定める関数．
 - ・エレメンタリーセルオートマトン (ECA) … $F : \{0, 1\}^3 \rightarrow \{0, 1\}$ で $u_n^{t+1} = F(u_{n-1}^t, u_n^t, u_{n+1}^t)$ なる時間発展則に従う CA．ECA に属す時間発展則の個数及びその命名法は 1.2.2 に述べた通りである．左右及び 0-1 の反転で一致するルールを同一視すると独立なルールは 88 種類である．
- システムサイズが有限なら，選ばれた初期状態 \mathbf{u}_0 ごとにある時刻 $\tau(\mathbf{u}_0), T(\mathbf{u}_0)$ が存在して (1) 式が成り立つ．（一意ではない）証明はディリクレの箱入れ原理より明らか．

$$\mathbf{u}_{\tau(\mathbf{u}_0)} = \mathbf{u}_{\tau(\mathbf{u}_0)+T(\mathbf{u}_0)} \quad (1)$$

- ・初期状態 \mathbf{u}_0 に対する緩和時間（緩和時刻）…(1) 式が成立する最小の時刻 $\tau(\mathbf{u}_0)$ ．以降この最小の値を以て $\tau(\mathbf{u}_0)$ と表記する． $t = 0$ から $\tau(\mathbf{u}_0)$ までの状態遷移を緩和過程と呼ぶ．
- ・基本周期…初期状態 \mathbf{u}_0 に対する最小の $T(\mathbf{u}_0)$ ．以降この最小の値を $T(\mathbf{u}_0)$ と書く．
- ・基本周期パターン（基本周期軌道）… $U_c = \{\mathbf{u}_{\tau(\mathbf{u}_0)}, \dots, \mathbf{u}_{\tau(\mathbf{u}_0)+T(\mathbf{u}_0)-1}\}$ のこと． U_c に対して $\{\mathbf{u}_{\tau(\mathbf{u}_0)+k}, \dots, \mathbf{u}_{\tau(\mathbf{u}_0)+k+T(\mathbf{u}_0)-1}\}, k \in \mathbb{N}$ を U_c の時間方向の k シフトパターンと呼ぶ．
- ・周期状態…与えられたルールのもとで周期軌道に属す状態の集合のこと． Ω_c と書く．周期軌道を構成する状態は全て有限時間で自分自身に回帰するので， $\mathbf{u}_0 \in \Omega_c$ の場合はこれまで用いていた表記 $T(\mathbf{u}_0)$ を単に T と書くことにする．緩和時間に関しては明らかに $\tau(\mathbf{u}_0) = 0$ である．
- ・緩和状態…与えられたルールのもとで周期軌道に属さない状態の集合のこと． Ω_t と書く．
- ・可逆 CA（可逆ルール）…任意の状態 \mathbf{u} に対して \mathbf{u} に遷移する状態がただ 1 つ存在するルール．
- ・基本相対周期… $\mathbf{u}_0 \in \Omega_c$ とするとき，その初期状態 $\mathbf{u}^0 = (u_1^0, \dots, u_N^0)$ を空間方向にシフトした状態が時刻 $t = T'$ に現れたとする $\mathbf{u}^{T'} = (u_{1+S}^0, \dots, u_{N+S}^0)$ ただし空間変数は $(\text{mod } N)$ で考える．そのとき， T' の最小値を \mathbf{u}^0 の基本相対周期と呼ぶ．
- ・シフト数… $\mathbf{u}^{T'} = (u_{1+S}^0, \dots, u_{N+S}^0)$ の S のこと． $\mathbf{u}^{T'}$ が \mathbf{u}_0 に対して空間方向に何セル分ずれているのかを表す．尚， $T' = T$ の場合， $S = N$ と約束する．先に述べた時間方向のシフトと混同しないようにする． $\mathbf{u}^{T'}$ を \mathbf{u}^0 を空間方向に S シフトした状態と呼ぶ．
- ・基本相対周期パターン（基本相対周期軌道）… U_c の一部 $\{\mathbf{u}_0 \dots \mathbf{u}_{T'-1}\}$ のこと． U_r と書く．尚，以下では特に注意がない限り，基本（相対）周期パターンを（相対）周期パターンと書く．

2.2 研究目的

本研究の目的を述べるに当たって、1 章を踏まえ、次の 3 点に注意する。

・注意 1【有限セルの制限】

ウルフラムの分類は（連続系に対応物を持たない）興味深い性質を持つと予想されるルールの候補を出力パターンから見つける為の目安である．これによって一応の目星をつけ、次の段階として、例えば、CA の計算能力の証明、といったテーマがあるのである．

目安は数値計算等を利用できる簡便なものが望ましい．しかし数値計算はシステムサイズが有限の場合に限られるので、最終的には全て周期軌道に落ち込み、クラス 3(Yields chaotic aperiodic patterns) は存在しないことになる．

・注意 2【数値計算とパターン】

次に、数値計算で見えているパターンに意味があるのか、という根本的な問題がある．

私が興味を持っているのは、CA の巨大出力パターンの解析である．これは N, T_F が共に大きい状況である．パターンのサイズに関する極限に対し、ウルフラムが数値計算を元に主張した有限セルに於けるパターンの特徴は生き残るのだろうか、という疑問が生じる．

尚、数値計算で見えていたパターンが極限を取ったときに消える例として次の現象が知られている．量子物理に於いて、スタジアムビリヤードの固有関数を数値計算すると、その絶対値が大きくなる部分が、対応する古典系に於ける粒子の軌跡に見えるという指摘が注目を集めた事があるが、そのパターンはエネルギーが高くなると ($\hbar \rightarrow 0$ に対応) 消えることが後に証明された [7]．

・注意 3【恣意性の除去】

1 章で問題にした恣意性は 2 つである．1 つは、目視に対する恣意性であり、もう 1 つは、観測時間に対する恣意性である．

前者は、これから行う適切な定量化手法の導入によって解消される．適切さの判定はウルフラムの分類を導入手法が再現可能かどうかにより判断される．例えば、クラス 2 のルールとクラス 3 のルールが異なるクラスのルールであると判定できないような定量化は不適切である．

後者は、 T_F を有限の値で打ち切らねばならないことに起因するが、もし、注意 2 で述べた極限を、うまく数値計算で扱えるなら、この問題は解消されたと言えるかも知れない．

以上の注意 1～3 を踏まえて本研究の目的を述べる．

— 研究目的 —

私の興味の対象は CA の計算能力等ではなく、ウルフラムの分類そのものにある．本研究の目的は注意 2 の最後に述べた通り巨大パターンに対するウルフラムの分類の妥当性『パターンサイズに関する極限を取っても 4 つのクラスに分類できるか』を検証する事である．この検証のために、注意 3 で述べた 2 つの恣意性を除去した定量的な分類基準を導入しなくてはならない．ただし、そこに注意 1 の『パターンサイズ有限（数値計算を使う）という条件下で』という制限がつく．

2.3 研究方針

前節に記した研究目的を達成するための研究方針を述べる．この方針を満たす具体的な接近法は 3 章に記す．

2.3.1 クラスの種類

分類の精密化の過程で，新たなクラスを定義する必要が生じることもあるかも知れないが，極端な話，256 種類のルールに対して 256 種類のクラスを定義することは明らかにナンセンスである．基本になるのはウルフラムによる 4 つのクラスであり，新たなクラスの導入は極力控えたい．

2.3.2 分類指標の種類

先行研究には複数の分類指標を組み合わせて ECA を分類したものがある [5]．しかし，これを許すと，極端な話，256 個のルールを分類するために 256 種類の指標を用いる，という話になってしまう．単一の指標による分類を目指す．

2.3.3 パターンサイズに対する極限

パターンサイズに関しては， N と T_F をどのように大きくするかをまず決めなければならない．ここで，本研究の出発点は数値計算による出力パターンの解析である点に留意すると， $N \rightarrow \infty$ は扱えない．初期状態に無限のセルがあると，次の状態を作る為に無限の時間が必要だからである．また，無限個のセルを用意すること自体不可能である．

しかし， N を有限の値に固定したもとの $T_F \rightarrow \infty$ は，扱える可能性がある．例えば，観測時間 T_F での状態の期待値ベクトルを $\bar{\mathbf{u}} = \sum_{t=0}^{T_F} \mathbf{u}_t / T_F$ とする． T_F が十分大きく，周期軌道に入っているとすると， $T_F = \tau(\mathbf{u}_0) + R'T(\mathbf{u}_0) + \hat{T}(\mathbf{u}_0)$ ．ここで， R' は周期の繰り返し回数， $\hat{T} = (T_F - \tau(\mathbf{u}_0) \bmod T(\mathbf{u}_0))$ である． T_F が非常に大きいとき， $R' \gg 1$ より， $\lim_{T_F \rightarrow \infty} \bar{\mathbf{u}} =$

$$\lim_{R' \rightarrow \infty} \frac{\sum_{t=0}^{\tau(\mathbf{u}_0)-1} \mathbf{u}_t + R' \sum_{t=\tau(\mathbf{u}_0)}^{T(\mathbf{u}_0)-1} \mathbf{u}_t + \sum_{t=\tau(\mathbf{u}_0)}^{\hat{T}} \mathbf{u}_t}{\tau(\mathbf{u}_0) + R'T(\mathbf{u}_0) + \hat{T}(\mathbf{u}_0)} = \frac{1}{T(\mathbf{u}_0)} \sum_{t=\tau(\mathbf{u}_0)}^{T(\mathbf{u}_0)-1} \mathbf{u}_t \quad (2)$$

となり，周期パターンを構成する状態の期待値に一致する．期待値ベクトルを使うかどうかは別として，この様な性質を持つ統計量を使うと，本来無限の時間が必要な時間発展を有限時間の解析に還元することができる．ここで，扱える『可能性がある』と言ったのは，一般に時刻 t で周期軌道に入ったかを判定するためには，時刻 $t-1$ までに生成された状態との比較を行わなければならない，それには初期状態から時刻 t までに生成された状態を保存しておかなくてはならない．緩和時間や周期が非常に長いと，保存しなければならない状態の数が数値計算で扱える行列サイズを超える事がある為である．その点， N が有限の可逆 CA は初期状態への回帰が保障されているので，初期状態との比較だけを行えばよく，数値計算の対象としては性質が良い．

尚，緩和の性質に着目した研究もある [8] が，数値計算という観点からは，ランダムな初期状態から出発しても一旦周期軌道に入ったことが確認できればその軌道の全容が確定するのに対して，緩和状態はランダムに選んだ初期状態が，緩和過程のどこにヒットしているかを知る事が難しく，従って，緩和状態の全容を把握する事も難しい．

2.3.4 目視によらない分類

[4] の後に出版された [9] には、次のように各クラスの特徴がより詳しく書かれている。

(1) the behavior is very simple, and almost all initial conditions lead to exactly the same uniform final state.

(2) there are many different possible final states, but all of them consist just of a certain set of simple structure that either remain the same forever or repeat every few steps.

(3) the behavior is more complicated, and seems in many respects random, although triangles and other small-scale structures are essentially always at some level seen.

(4) class 4 involves a mixture of order and randomness: localized structures are produced which on their own are fairly simple, but these structures move around and interact with each other in very complicated ways.

同一のクラスであっても、(3) の triangles and other small-scale structures のように見た目の異なる様相が含まれていることに注意する。ウルフラムはこれらの見た目の違いは無視して同一のクラスに分類しているため、このような特徴の違いに対して頑健な基準を用いる必要がある。

その際に、参考になるのが主成分分析 (PCA) である [10]。PCA を用いた複数の会社の株価の時系列データの解析では、第 1 主成分は市場全体の株価の動向を表すとされる。一方、寄与率の小さい主成分は、通常、市場のノイズとみなされる。我々が求める細部の特徴が除かれたパターンとは、いわばこのようなノイズに対応するものとなるだろう。尚、本研究は ECA の出力パターンを多変量時系列データとみなし解析した [12] を発展させた [13] に新しい結果を加えたものであり、[12] の解析手法は [11] を参考にしている。主成分や寄与率については 7 章を参照されたい。

2.3.5 初期条件

全ての初期条件に対してその軌道を求めることは計算量的に無理なので、各セル独立に確率 $1/2$ でランダムに初期値を作成して時間発展させる。このようにして数値的に出力パターンのランダムサンプリングを行い、その統計的な性質を解析する。

ただし、初期値を全て等しい重みとするサンプリングでは『幾つかの周期軌道が存在して、そのうちの小数の軌道に殆どの緩和状態が吸い込まれる』様な状況で長時間極限に対する十分なデータが得られない可能性がある事に注意しておく（この例は 3.6.3 節に記す）。

2.3.6 境界条件

代表的な境界条件は、両端のセルを固定したディリクレ境界条件と周期境界条件及び両端のセルの値を各時刻ごとにランダムに決定する境界条件がある。最後に挙げた境界条件は、統計力学と CA の関係を議論する場合に用いられることがあるようである [14]。本研究では、カオス的な要素は初期状態を除いて決定論的に生成された各状態の複雑さに求めるべきであるとする立場を取る。確率的境界条件では、各状態の複雑さの原因がルールなのか境界条件なのか不明になる為、この境界条件は使わない。周期境界の場合、対称性によってディリクレ境界条件より周期が短くなることが期待できる。これは数値計算をする上で都合が良いので周期的境界条件を用いる。因みに、ウルフラムの数値実験に於いても、周期境界条件が多用されている [4]。

3 分類手法

2.3 節に記した研究方針

(i) 単一の指標で

(ii) 出力パターンの長時間極限から

(iii) ウルフラムのクラスに沿った分類ができ

(iv) (iii) を満たすために同一クラス内でのパターンの違いには頑健な基準

を満たすように分類手法を導入する．(i) と (iii) については実際に数値実験を行なった後，これから与える分類手法が妥当であるかどうかが判明する．特に (iii) については，提案手法で明確な違いが見えるなら，(与えられたシステムサイズの元で) 長時間極限を取っても，やはり別々のクラスに分かれていると主張できる．

3.1 特異値分解

上の条件 (ii),(iv) を満たす手法として特異値分解と呼ばれるものがある [15]．

定義 3.1 X は任意の実行列であるとする．

$$\begin{aligned} {}^tXX\mathbf{v}_j &= \lambda_j\mathbf{v}_j, \quad X^tX\mathbf{u}_j = \lambda_j\mathbf{u}_j, \quad \mu_j = \sqrt{\lambda_j}, \quad \forall j \|\mathbf{u}_j\| = \|\mathbf{v}_j\| = 1 \\ X &= \mu_1\mathbf{u}_1{}^t\mathbf{v}_1 + \mu_2\mathbf{u}_2{}^t\mathbf{v}_2 + \cdots + \mu_r\mathbf{u}_r{}^t\mathbf{v}_r \end{aligned} \quad (3)$$

上式の右辺を X の特異値分解， $\mu_j (1 \leq j \leq r, r \text{ は } X \text{ の階数})$ を特異値という．特異値分解は基底となる長方形行列（本稿では基底行列と呼ぶ）の線形結合（特異値が重み）で X を表現する方法と考えることもできる．特異値は降順にソートされている ($\mu_j \geq \mu_{j+1}$) とする．ここで，固有ベクトル $\mathbf{v}_j, \mathbf{u}_j$ のノルムは 1 に規格化されているとする．また， tXX は相関行列という意味を持つ．

3.1.1 条件 (iv) と特異値分解

工学的な応用では，特異値分解して重みの大きい部分だけに着目し X の特徴を抽出すること等に利用されている．適当な基準の元で j' 番目までを使用するとする．このとき， X を元パターン，以下の P, B を主パターン，背景パターンと呼ぶ． $X = P + B$ である．

$$P = \sum_{k=1}^{j'} \mu_k \mathbf{u}_k {}^t\mathbf{v}_k, \quad B = \sum_{k=j'+1}^r \mu_k \mathbf{u}_k {}^t\mathbf{v}_k \quad (4)$$

カットオフの基準としては様々なものが考案されているが，代表的なものに以下の累積寄与率 S'_j がある．

$$S'_{j'} = \left(\sum_{k=1}^{j'} \mu_k / \sum_{k=1}^r \mu_k \right) \times 100 (\%) \quad (5)$$

累積寄与率基準を用いる場合， $S'_{j'}$ が初めて 80 % を超えたところでカットオフするのが一般的なようであるが，80 % という値に特に数学的な理由がある訳ではない．勿論，他のカットオフの基準を考えることもできる．

実際にクラス 3 の出力パターンを特異値分解し, P, B を求めた例を示す.

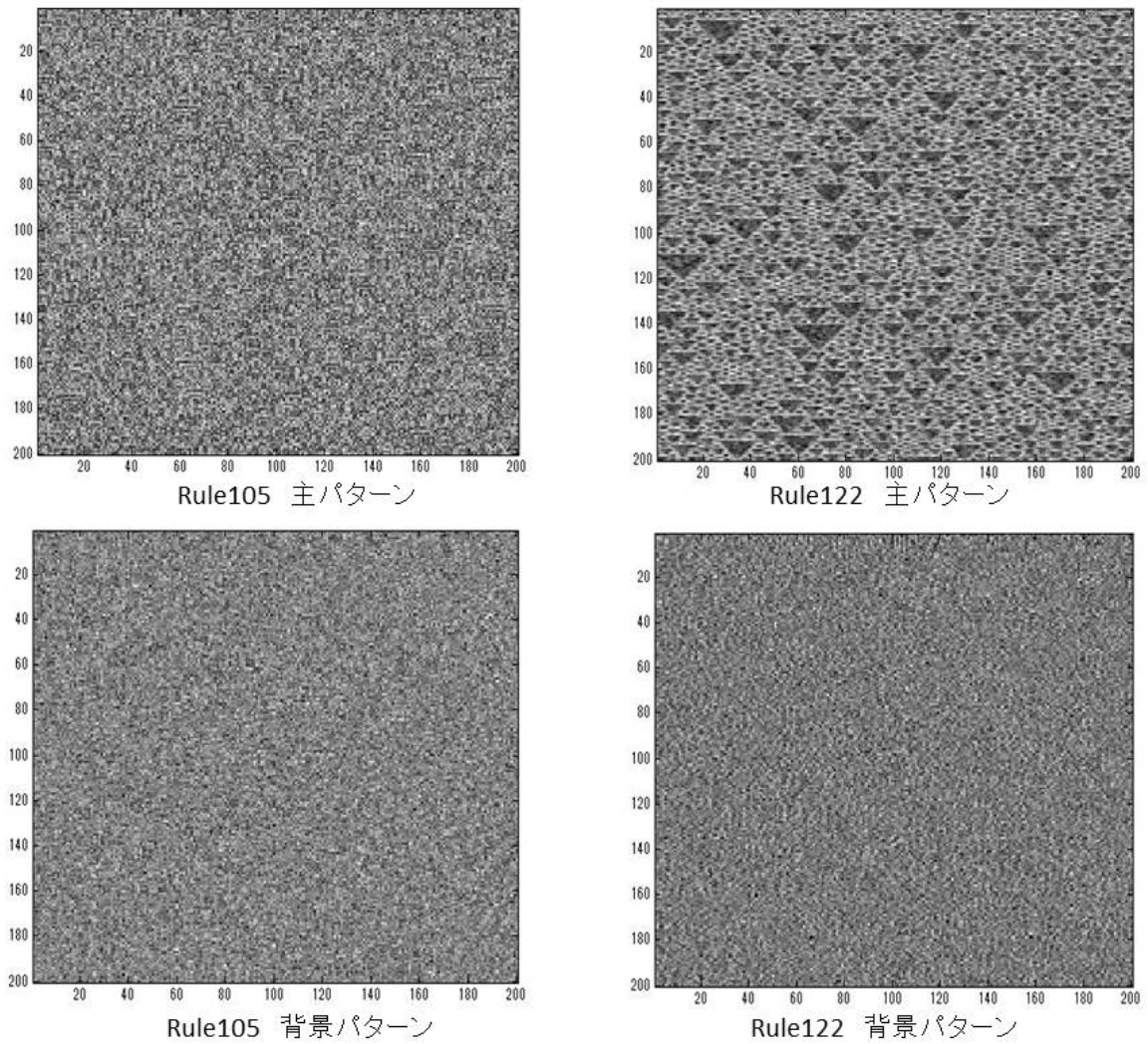


図 2 クラス 3 のルール 105,122 の出力パターンの P, B を比較したもの

図 2 は Appendix-A2 に掲載したルール 105 と 122 のパターンをともに $j' = N/4, N = 200$ でカットオフした際の主パターン P (上段) と背景パターン B (下段) である. ルール 105 と 122 の P の累積寄与率は各々 49.8 %, 55.6 % である. 累積寄与率は 80 % を切っているが, それでも P に関してはどちらのルールから出力されたか目視によって判断できる. 一方, B では双方の様相が非常に似ている点に注意する. パターンの特徴に鋭敏な基準ではクラス 3 と 4 が分類できる反面, 同じクラス 3 に属すパターンであっても異なるものとみなされ, その結果ウルフラムが当初与えたクラスを更に細分化する, といった事態に至ることがあるが, 背景パターン B に着目することでロバストな分類が期待できる.

ところで, B は X と常に異なる様相を呈するとは限らない. 単純なパターンの場合, X と B は似たものになる. 次のページに Appendix-A3 に掲載したクラス 2 のルール 154 とルール 170 の出力パターンを本節と同様に $j' = N/4, N = 200$ で分解した例を示す.

クラス 2 の出力パターンを特異値分解し, P, B を求めた例を示す.

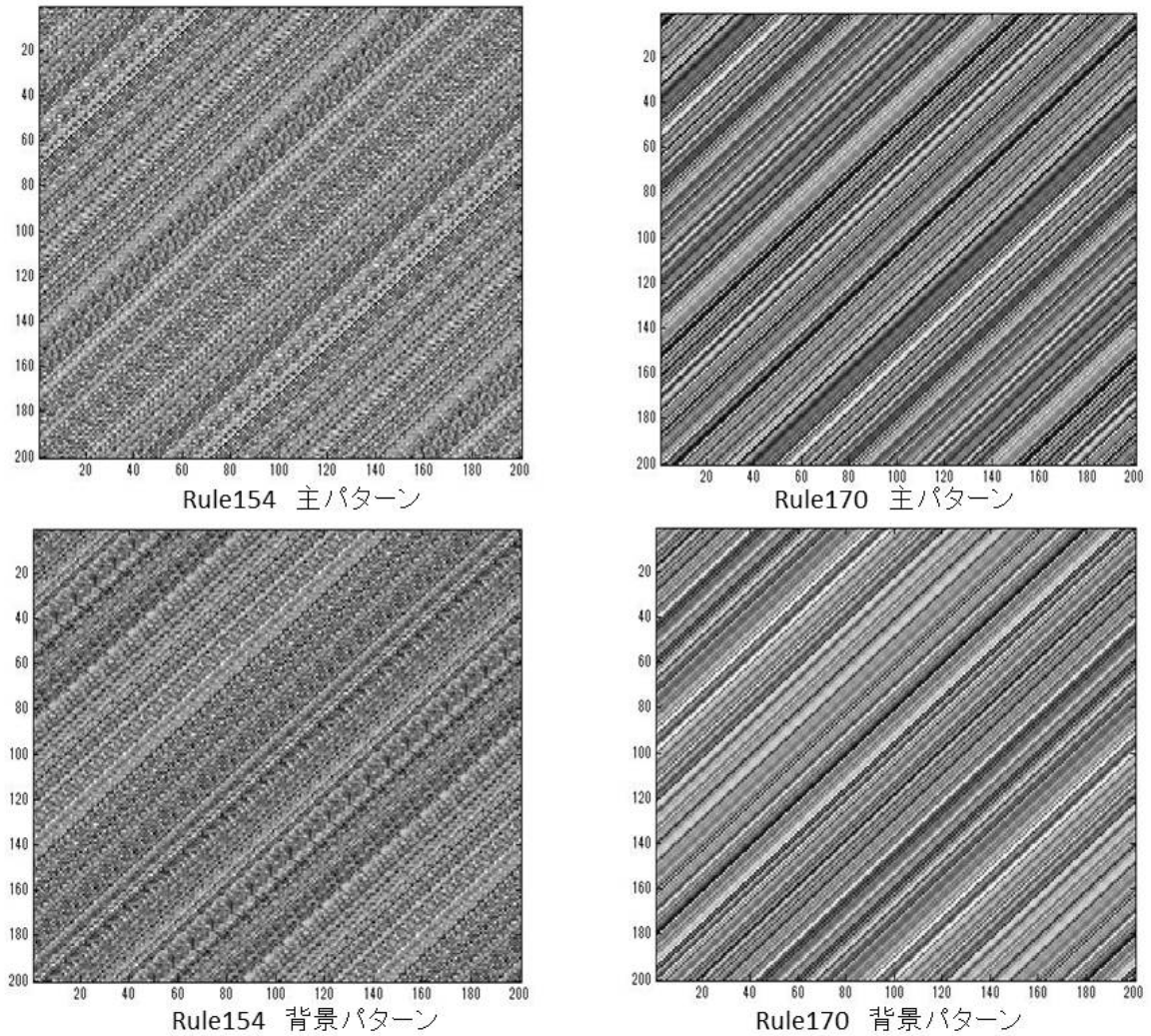


図 3 クラス 2 のルール 154,170 の出力パターンの P, B を比較したもの

ルール 154 と 170 の P の累積寄与率は各々 47.9 %, 45.7 %である. これらの出力パターンに関しては, X が元々似ている事もあり P から 2 つのルールを区別することは難しいが, それでも, B に X の特徴が生き残っており, クラス 2 から出力されたということは分かる.

B の様相は共に斜めずれパターンで, クラス 3 の B とは異なる. しかし, 以上の違いを以て, クラス 2 とクラス 3 が区別できると主張する事はできない. この段階での異なるとの判断は目視によっているからである. 従って以降では, CA の出力パターン背景パターン B に着目して, B のどのような量を定量的な分類の基準として採用するかを議論する.

尚, 本節で例として扱ったルール 154 は [5] に於いて特殊なクラスに属すルールとして分類されている. このルールでは局所的には比較的複雑なパターンが生じ得るが, それは domain wall と呼ばれる相互作用を制限する構造の内部に閉じ込められる. domain wall はルール 73 でも生じるので Appendix-A3 を参照されたい. 但し, ルール 73 は domain wall が縦方向に生じるがルール 154 では斜め方向に生じる. ルール 154 に於ける domain wall の存在は解析的に証明されている [16].

3.1.2 背景パターン B の定量化

背景パターン B に着目を定量的に扱うことを考える．例えば， B の行列要素の期待値なども何らかの指標にはなり得るかも知れない．しかし， B を作るために特異値分解で得られる特異値自体が，次の 2 つの観点から非常に興味深い量である．

1 つ目は，パターンの分析を一旦離れて軌道を構成する状態の集合の性質に着目したとき，時刻 t の状態の何個のセルの値が 1 になっているのかは重要な情報である．さらに詳細な情報として，セル同士の相関（ i 番目のセルに 1 があるとき， j 番目のセルにも 1 があるか否か）の統計性も軌道の特性を反映する．相関関数 tXX は，この相関を定量的に扱う指標の 1 つであり，その固有値は特異値の 2 乗に対応している．

2 つ目は，ランダム行列との対応を考えたときに，ランダム行列理論でランダム行列の特異値がよく研究されている為である．これについての詳しい議論は 3 章で行う．

3.1.3 条件 (ii) と特異値

3.1 節で述べた特異値分解の性質は（一部 X の例として CA の出力パターンを用いたが）一般的なもので X が実行列であれば成り立つ． X として CA の出力パターン U を用いる場合に問題になるのが条件 (ii) である． U に対し $T_F \rightarrow \infty$ の極限を素直に実行すると，その特異値は発散してしまう．そこで，発散を防ぐために U を次式のように規格化したパターンを用いる．

$$U \rightarrow U/\sqrt{T_F} \quad (6)$$

こうしておけば， $T_F \rightarrow \infty$ の極限を取ったとき，求める特異値は，ランダムに生成された初期状態から時間発展して到達する周期パターンを，その周期の平方根で割った行列 U_c/\sqrt{T} の特異値に一致する．この証明は (2) 式と全く同様にできるので省略する．

上の結果によって，ウルフラムが言及したクラス 4 の次の特徴

Class 4 cellular automata effectively have very long transients

は，直接的には本研究の解析の対象とならない．もし，本研究によって緩和状態を無視しても尚，クラス 4 と他のクラスの特異値の性質に明確な差異が見出されたなら，それは，クラスの特徴をより厳密に特定したことになる．

ここで、『直接的には』と断わっている理由は，緩和状態の個数が周期軌道の個数や周期に対し制限を加えるからである．また，(6) 式の変換により，解析対象となる特異値は元の周期パターン U_c の特異値を定数倍したものとなる．特異値の何を解析するか（平均値なのか，分散なのか，他の量なのか）についてはこれから議論するが，例えば特異値の平均値に着目するのであれば， U_c/\sqrt{T} と U_c の特異値の平均値とは異なる．従って，現段階では，長時間極限を取ったパターンの解析が周期パターンの解析に帰着されると素直に主張できない点に注意しておく必要がある．

3.1.4 比較対象

最後に CA の出力パターン（の特異値）を特徴づける為の基準について考える．クラス 3 の出力パターンはどれも複雑だが三角波の有無等は異なりルール毎の特徴がある．しかし，背景パターンではその特徴が消え，完全にランダムなパターンである ランダム行列 に似た性質を持つ事が期待される．そこでまず我々はクラス 3 とランダム行列の特異値の性質を比較する．

3.2 ランダム行列

ランダム行列は 1920 年代に数理統計学の分野で考案されたものであるが、原子核物理への転用を始めとして、量子カオスや数理生態学、金融工学等の幅広い分野に応用されている。原子核物理と量子カオスに関するランダム行列の歴史的な背景は Appendix-C に簡単にまとめておく。

ランダム行列とは行列要素が確率分布関数から出力された行列である。行列の構成の仕方の違いなどにより、様々な種類があるが、以降の議論に関連する 2 種類のランダム行列を以下に例として挙げる。

定義 3.2 ガウス型直交アンサンブル GOE:Gaussian Orthogonal Ensemble

これは、 $N \times N$ 実対称行列で、その独立な $N(N+1)/2$ 個の行列要素がそれぞれ独立にガウス分布に従う行列の集合である。 $M_1 \in \text{GOE}$ とするとき、行列 M_1 の従う確率分布関数は、

$$P(M_1)dg(M_1) \propto e^{-\frac{1}{2}\text{Tr}M_1^2}dg(M_1) \quad (7)$$

である。測度 $dg(M_1)$ は、下式で与えられる。

$$dg(M_1) = \prod_{j=1}^N dM_{1jj} \prod_{j<l}^N dM_{1jl}$$

尚、GOE の Orthogonal は確率分布関数の直交変換に対する不変性に由来する。

定義 3.3 カイラルガウス型直交アンサンブル chGOE:chiral Gaussian Orthogonal Ensemble

各行列要素が独立に同一のガウス分布に従う $L \times N$ 長方形行列を A とする ($L \leq N$)。このとき、下式で定義される行列 M_2 の集合のことである。

$$M_2 = \left(\begin{array}{c|c} 0 & {}^t A \\ \hline A & 0 \end{array} \right) \quad (8)$$

${}^t A A$ (これは Wishart 行列と呼ばれる) の固有値 λ と M_2 の非零固有値 λ' には次の関係がある。

$$\lambda' = \pm\sqrt{\lambda} \quad (9)$$

以降で、chGOE の固有値と CA から求めた特異値の性質を比較する事は明らかであろう。

3.3 ランダム行列の性質 (固有値分布)

chGOE の固有値と CA から求めた特異値の性質を比較するに当たって、どのような性質に着目すべきなのだろうか。ランダム行列理論では固有値の統計的な性質に関して、固有値分布、固有値間隔分布、数分散、といった種々の量が調べられている。

本節では、CA の出力パターン (ルール) の分類指標として統計量が満たすべき条件を議論する。その際に重要な指針となるのが 3 章に記した条件 (i) である。これは分類指標がロバスト性を持つ事を要請している。例えば、chGOE の固有値分布を比較対象に選んだとする。このとき、クラス 3 に属すルールの出力パターンの特異値分布がルールによらず chGOE の固有値分布に似たものになれば、それはロバストな基準であると言える。

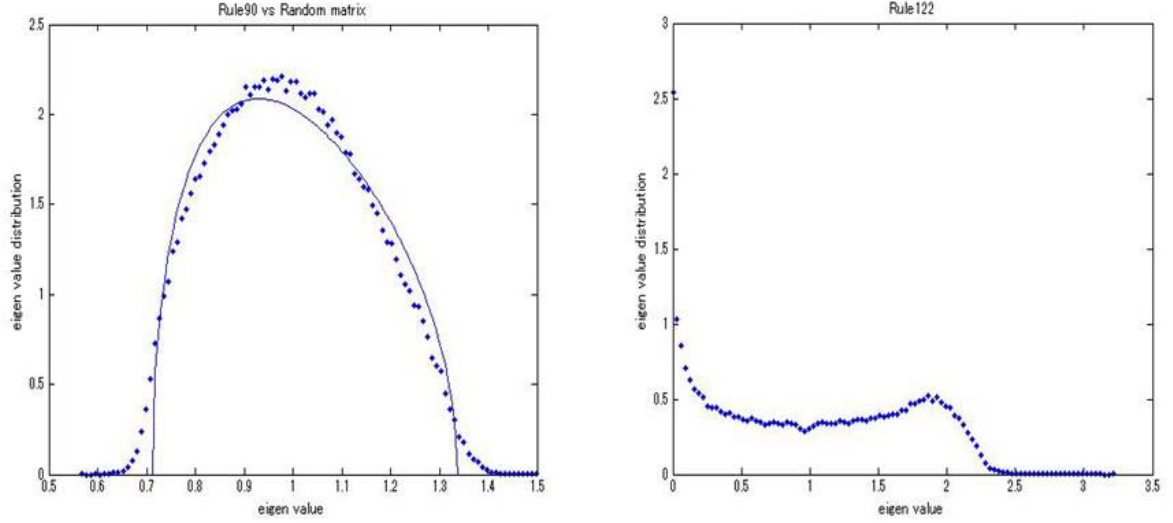


図4 固有値分布

ランダムパターンとの比較という目的を一旦忘れて ${}^tU_c U_c / T$ と ${}^tA A / L$ の固有値の比較を考えてみる． A の行列要素を出力するガウス分布の平均を 0，分散を 1 とし比率 $Q = L/N$ を固定したまま $N \rightarrow \infty, L \rightarrow \infty$ とするとき，その固有値分布 $\rho_{rm}(\lambda)$ は下式で表される [11]．

$$\rho_{rm}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \quad (10)$$

$\rho_{rm}(\lambda)d\lambda$ は微小区間に存在する固有値の割合である．ここで，

$$\lambda_{\pm} = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}} \quad (11)$$

である．本研究では有限サイズのパターンを用いるが，システムサイズをある程度大きく取れば，近似的には近い分布が観測されるだろう．

ただし，直接 ${}^tU_c U_c / T$ と比較すると平均と分散の条件が異なるので， U_c の列ごとに行列要素の平均値を $\bar{u}_n = \frac{1}{T} \sum_{t=1}^T u_{tn}$ 分散を $\sigma_n^2 = \frac{1}{T} \sum_{t=1}^T (u_{tn} - \bar{u}_n)^2$ とし， $u'_{tn} = (u_{tn} - \bar{u}_n)/\sigma_n$ を要素とする行列 U'_c から ${}^tU'_c U'_c / T$ を求め，比較を行う．尚，この方法は経済物理学に於ける株価の相関解析の手法 [11] と基本的に同じものである．

図4はクラス3に属すルール90と122の固有値間隔分布である．システムサイズはともに $N = 200$ ，サンプル数を500として計100000個の固有値を使用した．図中の点はECAから計算した結果である．同じクラス3のルール90と122で明らかに分布形が異なる．しかし，これから直ちに固有値分布が分類指標として適切ではないと結論することはできない．パラメタ Q の値を変えれば，どちらの分布にもフィッティングできる可能性があるからである．因みに，ルール90では得られた500サンプルすべてに関して $T = 8184$ であった．左図の実線は， $\rho_{rm}(\lambda)$ に於いて $Q = 8184/200$ としたものである．ルール90に関しては点が実線に比較的よく乗っていることが分かる．しかし，ルール122では複数の周期が観測された．これは大変本質的な問題である．なぜなら，比較対象である $\rho_{rm}(\lambda)$ の Q の値を固定できないからである．ルール122に限らず，ECAの周期パターンの周期は一般に複数存在する．

複数の初期状態をランダムに生成し，それぞれを時間発展させて複数の周期軌道のサンプリングを行い，そこから求めた特異値の統計的な性質を議論したいのであれば，比較対象となる chGOE の性質の方に $L(T)$ や N に鋭敏な差を示さない量を選ぶべきである．これを要請 (v) とする．

3.4 ランダム行列の固有値間隔分布

本節では要請 (v) を満足する統計量として固有値間隔分布に着目し、その性質について簡単にまとめる。本題に入る前に、GOE の平均 0 分散 1, $N \rightarrow \infty$ の固有値分布を以下に記す。

$$\rho(\lambda) = \frac{1}{2\pi} \sqrt{4 - \lambda^2}, \quad |\lambda| \leq 2 \quad (12)$$

これは Wigner の半円則と呼ばれるもので、 $-2 \leq \lambda \leq 2$ の外側で固有値密度が 0 になる。同様に (10) 式でも $\lambda = \lambda_{\pm}$ の外側の固有値密度は 0 である。このような分布の端を edge と呼ぶ。(10) 式と (12) 式では分布形が異なる事に注意する。次に、固有値間隔分布について説明する。

定義 3.4 固有値間隔分布

これはランダム行列の固有値を大きさの順にソートしたときの隣り合う固有値の間隔の分布で、固有値の揺らぎを定量化した指標の 1 つである。GOE の固有値間隔分布は、次の **Wigner 分布** によって極めてよく近似される事が知られている [17]。s は固有値の間隔である。

$$P_{wigner}(s) = \frac{\pi}{2} s e^{-\frac{\pi}{4} s^2} \quad (13)$$

Wigner 分布は 2×2 サイズの GOE の固有値間隔の厳密な分布であり、その導出は Appendix-B を参照されたい。 $N \rightarrow \infty$ の場合の厳密な分布は、JMMS 方程式を数値的に解いて得られ [18]、Wigner 分布とは微妙にずれたものとなるが、多くの場合に、表式が簡便な (13) 式が用いられる。

特に重要な性質として、chGOE の固有値間隔分布も Wigner 分布で近似されることが知られている [11]。これらの行列アンサンブルは全て実対称行列である。固有値間隔分布は行列の対称性に依存し、実対称行列、エルミート行列、シンプレクティック行列で形が異なるが、本研究の対象は全て実対称行列である。以上から、Wigner 分布は条件 (v) を満たし、分類基準となり得る。

4 章で (i) の条件が満たされるかを数値計算と統計的検定によって確認する。CA のパターンを議論する文脈では特異値間隔分布、chGOE の固有値について述べる場合には固有値間隔分布という表現を用いるが混乱は生じないだろう。

3.5 固有値間隔分布を求める（数値計算の）際の注意点

固有値間隔分布を求めるに当たって、大きさの順にソートした固有値全部の間隔を求め、それを全て使ってヒストグラムを作成することは適切ではない。間隔を求めるために全部の固有値が使用できる訳ではなく、更に、使用できる固有値に対してもアンフォールディングという操作を施したものの間隔を用いなくてはならない。

【使用する固有値】

edge 近傍の固有値の性質は他の部分とは異なることが知られており、この部分の固有値を予め取り除いておく。3.1.2 節に述べた通り、本研究で解析の対象となるのは背景パターン B の特異値であるから、その時点で最大固有値付近は解析から除外される。具体的には、上位 $1/4$ 個の特異値を大きい特異値としてカットする。カットオフの位置を固定することで全てのパターンで使用する特異値の個数を統一できる。同様に下位 $1/4$ 個の特異値も除外する。

定義 3.5 アンフォールディング

アンフォールディングとは、局所的なスケーリングにより固有値間隔の期待値を規格化する操作のことである。

本操作の必要性は (12) 式を思い出すとイメージしやすい。【使用する固有値】の内容に従って edge 付近の固有値を削除した残りの固有値に着目する。このとき、 $\lambda = 0$ に近づくほど、固有値の平均間隔は密になる。即ち、ナイーブに計算した固有値間隔は固有値が存在する領域に依存する。

Wigner の半円の原点の近傍で、平均間隔程度の領域内を観測する場合、次の変換により固有値間隔の固有値密度依存性を消去する。

$$\lambda \mapsto \rho(\lambda)\lambda \quad (14)$$

平均間隔 $1/\rho(\lambda)$ で λ を割ることで間隔の期待値が 1 になるようにしている ($s = 1$ からのずれが、領域によらない固有値の配置の揺らぎを表す)。

ここで、3.1.3 節の最後に $U_c \rightarrow U_c/\sqrt{T}$ によって生じる特異値の『定数倍の違い』が分類に影響するかについて議論する必要があると記していた事を思い出す。結局、特異値と平均間隔がともに定数倍された状況ではアンフォールディングするとそれらが相殺して、パターンを \sqrt{T} で割っても変化は生じない。これによって、特異値間隔分布は出力パターンを直接用いて計算できる。但し、比較対象となるバイナリランダム行列は通常のランダム行列とは異なり、平均は 0 ではない。この点に関しては、計算機で適当なサイズのバイナリランダムパターンを作成し、その特異値間隔分布を求め、それが数値計算で求めた同じサイズの chGOE の固有値間隔分布によく一致する事を確認した。この普遍性もアンフォールディングによると考えられる。

以上を踏まえ、今回用いた特異値間隔分布の数値計算方法を以下に記す。アンフォールディングの操作は **3~5** に対応する。上述のものとは一見異なるが、アンフォールディングを数值的に行う方法は複数提案されており、代表的なものは [19] に簡潔にまとめられている。

- 1 ランダムに生成した初期値から時間発展させ、相対周期パターン U_r の特異値を求める。
- 2 1 で得た特異値を小さい順にソートした特異値列 $\{\mu_1, \dots, \mu_N\}$ に対してそれぞれの端から $1/4$ (計半分) の特異値を捨てる。
- 3 2 の特異値 μ_j の位置で 1 だけ上がる階段関数を作る。
- 4 3 の各階段の中点をとり適当な次数の多項式 (今回は次数を 6 とした) $f(\mu)$ で最小 2 乗フィッティングする。
- 5 4 で得られた $f(\mu)$ から $y_j = f(\mu_j)$ を求めその間隔 $s_j = y_j - y_{j-1}$ を求める。
- 6 1~5 を色々な初期値について繰り返す。 $N = 200$, サンプル数 M ($M = 500$ でランダムサンプリングを行う), i 番目の行列の j 番目の間隔を $s_j^{(i)}$ として集合 $S = \{s_j^{(i)} | j = 1, 2, \dots, L'(i); i = 1, 2, \dots, M\}$ を作成し、これをビンの間隔 0.1 で度数分布にして (横軸 s が特異値の間隔の広さ、縦軸は度数) それを規格化して間隔分布 $P(s)$ を得る。

U_c ではなく U_r を用いる理由は次節で改めて説明する。また、6 のサンプリングに関して、場合によっては使用できる特異値間隔が減少することがあり、使用可能な特異値の間隔を $L'(i) \leq [N/2] - 1$ としている。ここで $[\dots]$ はガウス記号である。例えば、 $T' < N$ ではランク落ちにより、自明な特異値の減少が生じる。極端に特異値が減少すると 4 の平滑化が不安定になるため、今回は予め $T' \geq N$ の U_r を M 個サンプリングする。勿論、 $T' \geq N$ でも特異値が極端に減ることはあり得る。その例がルール 73 である。

3.6 相対周期パターン

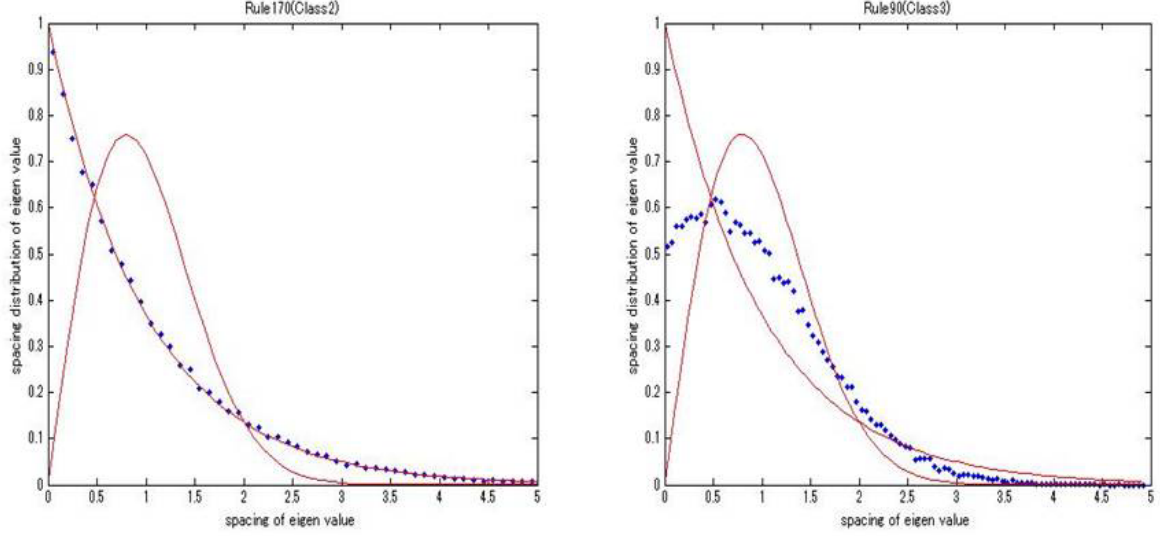


図 5 周期パターンの特異値間隔分布

解析の対象を周期パターン U_c から相対周期パターン U_r に変更する点について説明する．まず，周期パターンを用いた場合に生じる不都合を見ておこう．

3.6.1 ルール 170 (クラス 2)

左図はクラス 2 に属すルール 170 (シフト写像) から求めた特異値間隔分布である．ウィグナー分布でない実線はポアソン分布 $P_{\text{poisson}}(s) = e^{-s}$ であり，ルール 170 に限らず U_c に斜めずれが生じるクラス 2 のルールはポアソン分布によく一致する．ルール 170 に周期境界条件を課した場合，並進対称性から ${}^tU_c U_c$ は対称テプリッツ行列となる．

対称ランダムテプリッツ行列に関する先行研究 [20] に，その固有値間隔分布を数値計算するとポアソン分布が見られたという記述がある（ただし，[20] の対称ランダム行列は本研究とは違い，行列の積から作ったものではない）．証明は得られていないようだが， ${}^tU_c U_c$ の持つ実対称性以外の規則性が間隔分布に強く影響すると予想できる．

3.6.2 ルール 90 (クラス 3)

一方，その複雑さ故にウィグナー分布の出現が予想されたクラス 3 では，ルール 105 と 150 に関しては予想通りであったが，それ以外のルールでは右図のルール 90 に見られるような中間的な分布となっていた．クラス 3 に属すルールは次章（分類対象）にまとめておく．

この原因について調べたところ，ルール 90 では 500 サンプル全てで $T' = T/2$ となっている事が分かった．すなわち，一見複雑でも $T/2 \leq t < T$ のパターンは $0 \leq t < T/2 - 1$ のパターンを空間方向に半分シフトさせたものになっており，これが ${}^tU_c U_c$ に（完全な対称テプリッツ行列となるには足りない）規則性を生じさせた結果，特異値間隔分布の形が中間的なものになると考えられる（ルール 105 と 150 では $T' = T$ であった）．この推測が正しければ，間隔分布はパターンの複雑さに関係なく T' と N の関係で決まることになるが，我々が知りたいのはパターンの複雑さの方であるから，解析対象を U_r とする．

3.6.3 ルール 45（クラス 3）

ルール 45 はクラス 3 に属すルールの中でも特に長い周期を持ち [9], かつシステムサイズが奇数のとき可逆であることが知られている [16]. 手持ちの計算機で最大周期を求めることができた最大のシステムサイズは $N = 23$ で, その値は $T = 3459591$ であった. 下表は, サンプル数 $M = 500$ の実験から得られた周期の長さ（上段）と, その個数（下段）をまとめたものである.

ルール 45（システムサイズ 23）の周期の長さ（上段）とその個数（下段）

138	276	522	4117	11914	48484	123280	1358104	1608206	1630792	3459591
1	3	1	1	1	2	8	90	95	97	201

全状態数は $2^{23} = 8388608$ であるから, ランダムに生成した初期状態は 40 % 以上の確率で最大周期に属しており, それらの初期状態は同じ特異値列を出力する. その結果, 出現し難い軌道から得られる特異値間隔の相対頻度が極端に小さくなり, ウィグナー分布との比較に耐え得る様な滑らかな分布を得る事ができない. これを回避するための方法の 1 つは重複して現れる軌道を削除する事であるが, これは軌道に対する等しい重みでサンプリングであり, 一般に初期値に対する等しい重みでのサンプリングとは異なる. 尚, 最大周期の軌道の相対周期は $T' = 150417$ であった.

4 数値実験

4.1 分類対象

1 章で述べた独立な 88 個のルールのうち, 数値計算を試みたのは以下の 80 個である. クラス等に応じてそれらを表にまとめた. クラス 1 に属すルールは定義が明確なため対象から除外した.

クラス	ルール番号
クラス 2	1,2,3,4,5,6,7,9,10,11,12,13,14,15,19,23,24,25,26,27,28,29,33,34,35,36,37,38,41,42,43,44,46,50,51,56,57,58,62,72,73,74,76,77,78,94,104,108,130,132,134,138,140,142,152,154,156,162,164,170,172,178,184,200,204,232
クラス 3	18,(22),(30),(45),60,90,105,(106),122,126,146,150
クラス 4	54 ,110

ルール 54 が太字になっているのは, クラス 4 か否かを巡り現在も議論が続いている為である [21]. 又, 括弧が付きの 4 つのルールは相対周期または緩和の長さが非常に長く手持ちの計算機では U_r が得られなかったルールである. これらは今回は扱わない. ウルフラムの次の指摘 In a given Class 4 cellular automaton, the simplest self-reproducing structure may have a size of say 100 sites.[4] を考慮し, システムサイズは 100 以上の適当な数として $N = 200$ とした.

4.2 実験結果

まず, クラス 3 及び 4 の結果を記す. グラフの点が数値計算結果, 実線がウィグナー分布, 右図は左図の縦軸を対数プロットしたものである.

4.2.1 クラス 3 (ルール 18,60,90)

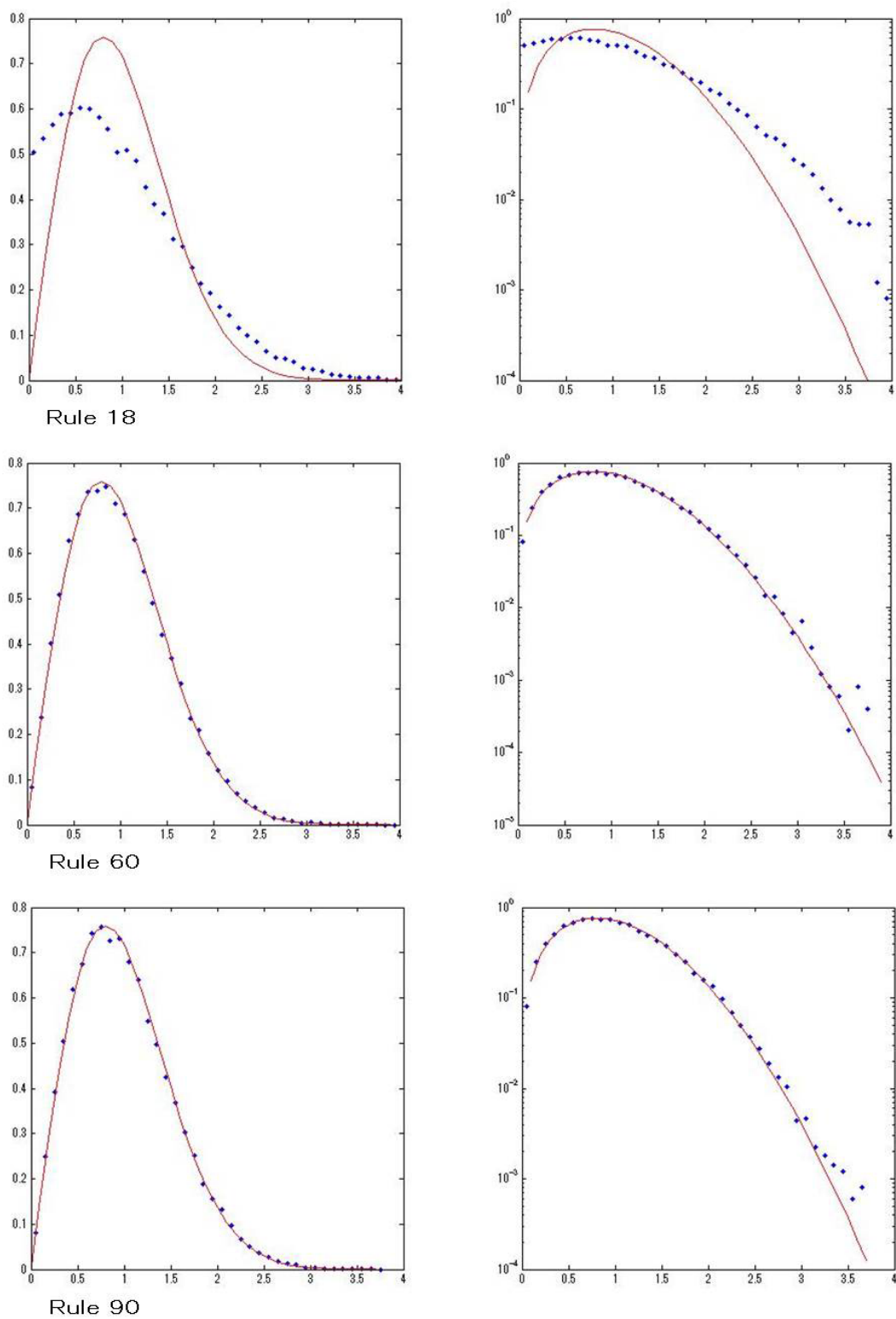
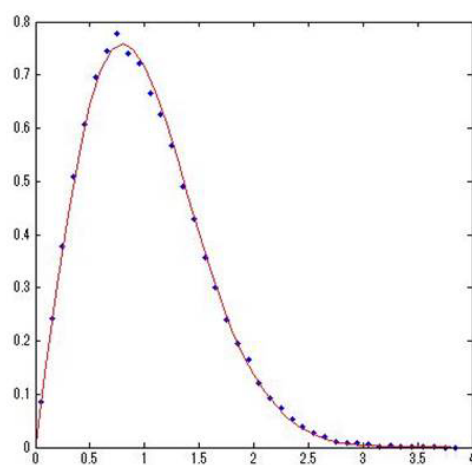
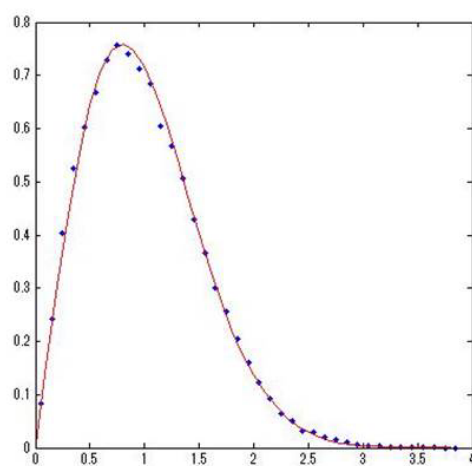
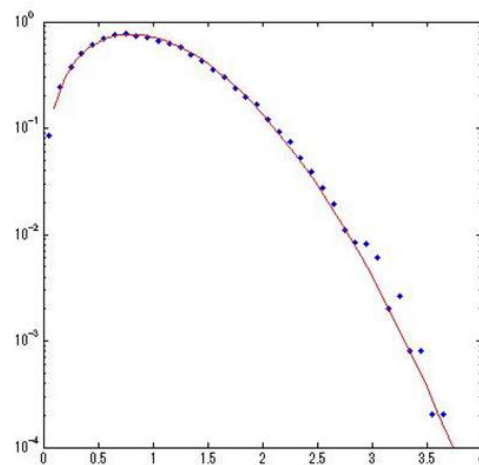


図 6 クラス 3 (ルール 18,60,90) の特異値間隔分布

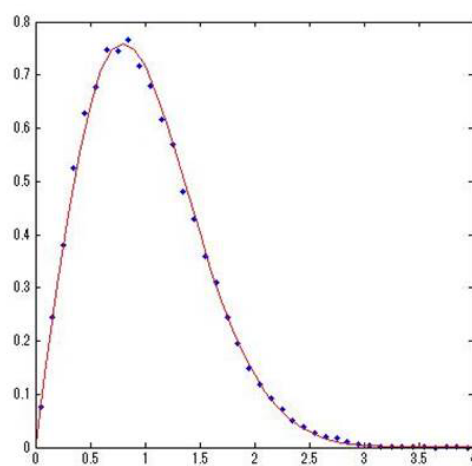
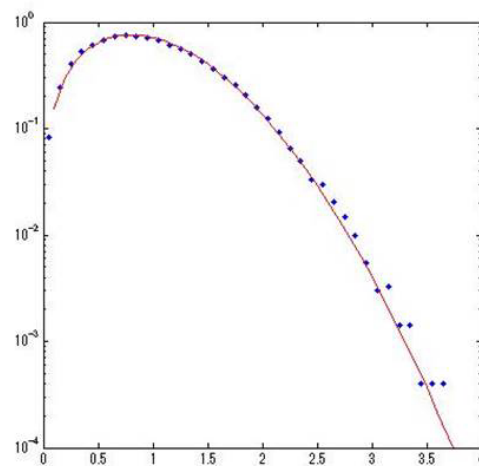
4.2.2 クラス 3 (ルール 105,122,126)



Rule 105



Rule 122



Rule 129

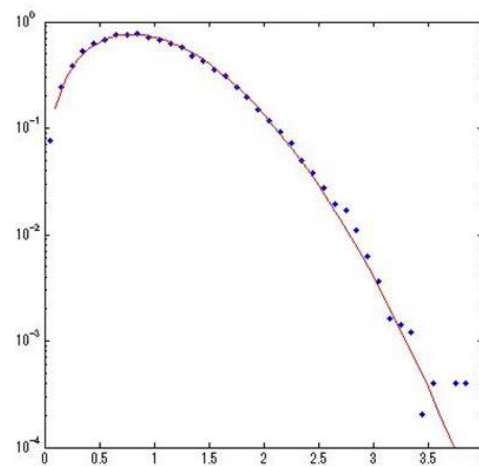


図7 クラス 3 (ルール 105,122,126) の特異値間隔分布

4.2.3 クラス 3 (ルール 146,150), クラス 4 (ルール 110)

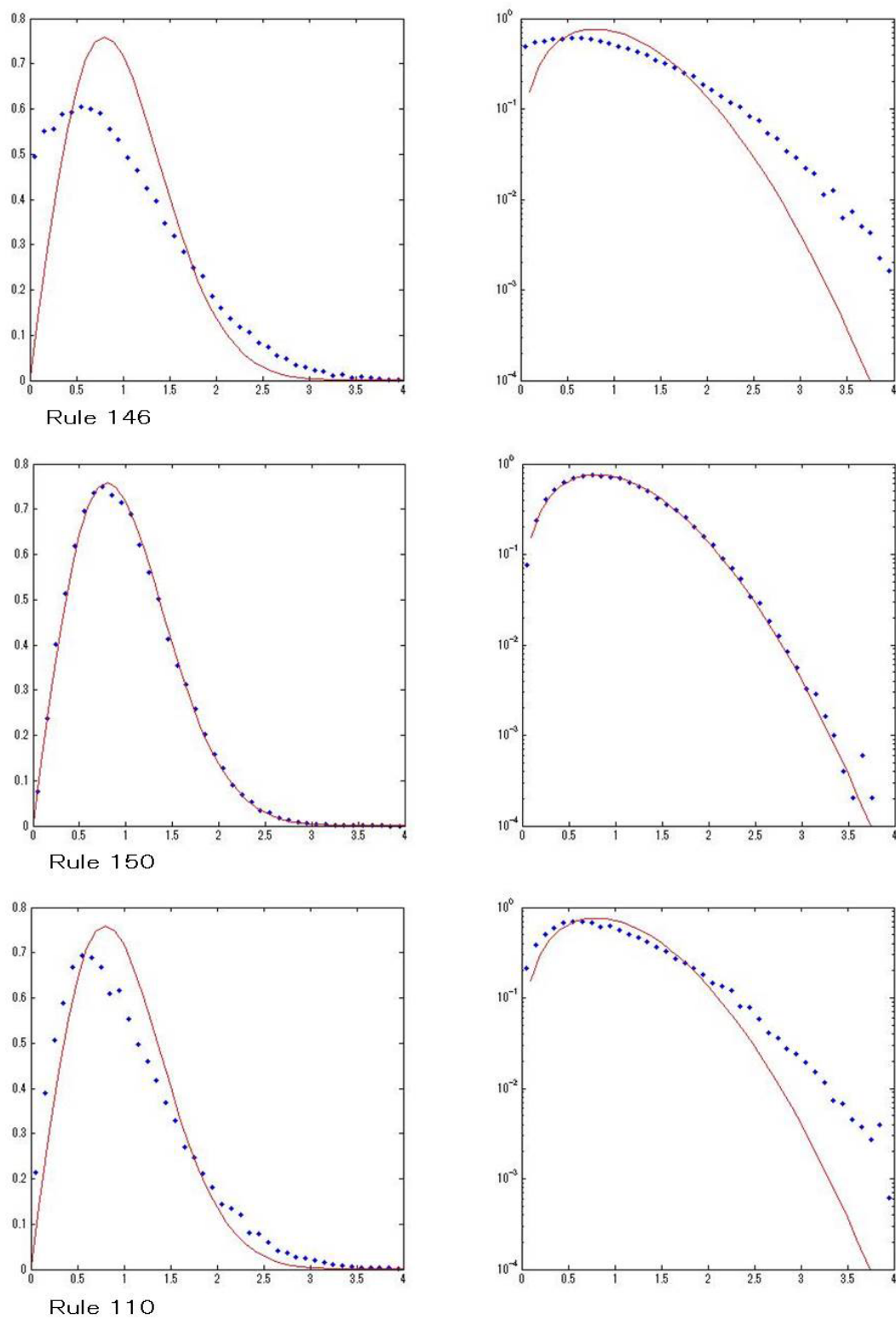


図 8 クラス 3 (ルール 146,150), クラス 4 (ルール 110) の特異値間隔分布

4.2.4 クラス不明とされているもの（ルール 54）

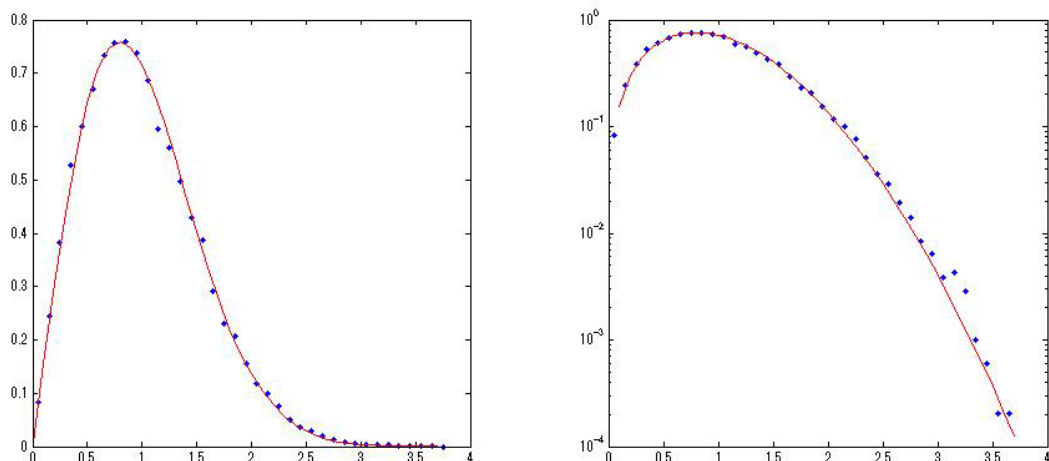


図 9 ルール 54（クラス不明とされる）の特異値間隔分布

4.3 ECA の実験結果のまとめ

次章ではルールごとの詳しい議論を行うが、その前に数値実験の結果を簡単にまとめておく。

(I) クラス 4 か否かを巡って現在も議論が続いているルール 54 について、その特異値間隔分布は上図の通りクラス 3 を特徴付けるウィグナー分布とよく一致することが分かった。

(II) ほとんどのクラス 2 では相対周期がシステムサイズ以上の U_r がサンプリングできない。例えば、比較的複雑なパターンを生成するルール 154 であっても、 $N = 2000$ で $T' = 16 \sim 32$ 程度である。ルール 73 では domain wall と呼ばれる時間変化しないセルに挟まれた領域に於いて複雑な時間発展が生じる為 U_r はサンプリングできたが、他のクラス 2 と同様、特異値の個数は少ない。尚、クラス 2 では U_r の代わりに U_c を用いるとポアソン分布が観察されることが多い。

(III) 実験を行った全てのクラス 3 のルールでウィグナー分布によく一致する結果が得られた訳ではなく、ルール 18 とルール 146 に関しては相対周期パターンを用いても尚、ルール 90 で周期パターンを用いた際に見られたような中間的な分布が観測された。ルール 18 と 146 に於ける中間的な分布が何に起因して生じるかは次章で考察する。

(IV) クラス 4 の特異値間隔分布はウィグナー分布からのずれが大きいものとなった。クラス 3 とクラス 4 を区別できる定量的な指標の作成という観点からは良い結果と言える。しかし、(III) の中間的な分布との違いが問題になる。結論を先に述べるとクラス 4 の分布はクラス 3 の全ての分布と異なることが次章の統計的検定で示された。

(V) ルールごとの使用可能な特異値の個数については、特異値が減少したサンプルも存在するものの、グラフを描いた 10 個のルールすべてに於いて 500 サンプルの特異値の個数の平均値は 199 以上であった。

5 結果の分析

5.1 統計的検定

4 章の結果はウィグナー分布によく一致している 7 つのルール（ルール 54,60,90,105,122,129,150）と、そうではない 3 つのルール（ルール 18,146,110）に大別できそうである。（クラス 3 の全ルールがウィグナー分布になっている訳ではない点は後で議論するとして）現時点では目視によって一致・不一致を判断している点に於いて、恣意性を除去できていない。本節では統計的検定を行うことでこの問題点を解消する。

5.1.1 クラス 3 の特異値間隔分布が Wigner 分布か否か

初めに判断したいのは、 U_r の特異値間隔がウィグナー分布に従っているか否かである。これを確かめる為に 1 標本コルモゴロフ-スミルノフ検定（1 標本 KS 検定）を用いる。

定義 5.1 1 標本コルモゴロフ-スミルノフ検定 [22]

n 個のデータ d_1, d_2, \dots, d_n の経験累積分布関数を $F_n(d)$ とする。

$$F_n(d) = \frac{1}{n} \{d_i \leq d \text{ を満たすデータの個数} \} \quad (15)$$

データが想定している分布を $F(d)$ とする。1 標本 KS 検定では $F(d)$ と $F_n(d)$ の離れ具合を表す量として次の量を用いる。データが $F(d)$ からのランダムサンプルなら D_n は小さいであろう。

$$\begin{aligned} D_n &= \max(D_n^+, D_n^-) \\ D_n^+ &= \max_{-\infty < d < +\infty} (F_n(d) - F(d)) \\ D_n^- &= \max_{-\infty < d < +\infty} (F(d) - F_n(d)) \end{aligned} \quad (16)$$

帰無仮説は『得られた標本は $F(d)$ から抽出された』であり、対立仮説は『得られた標本は $F(d)$ から抽出されていない』である。

尚、本研究の数値計算には MATLAB 7.12.0(R2011a) を用いているが、MATLAB の `kstest` というコマンドで 1 標本 KS 検定を行う事ができる。ウィグナー分布の累積分布関数に関しても、種々の確率密度関数が予め用意されており、`cdf` コマンドを用いるとそれらの累積分布関数を得ることができる。具体的には、下式のレイリー確率密度関数を `rayl` で呼び出し、

$$P(x|b) = \frac{x}{b^2} e^{-\frac{x^2}{2b^2}} \quad (17)$$

パラメタ値を $b = \sqrt{2/\pi}$ として `cdf` に渡せばよい。また、データの個数 n は [22] によれば $n \geq 1000$ 程度で十分に大きいとされているが、本稿では 1 サンプル約 99 個 \times 100 サンプル = 約 9900 個を用いた。4.2 節のグラフを描く為にサンプリングした U_r は 500 個なので、これらを全て用いればデータ数をさらに約 5 倍まで増やすことができる。しかし、データ数を上げれば必ず良い結果が得られるとは限らない。これについての説明は便宜上、次節で行う。

有意水準 5 % で検定を行ったところ、上記の 7 つのルールに関しては帰無仮説を棄却できず、ウィグナー分布に従っている可能性を否定できない事が分かった。また、残り 3 つに関しては帰無仮説が棄却された。

5.1.2 ルール 18,146,110 の特異値間隔分布が同じか否か

次に判断したいのは、ウィグナー分布に従っていない事が判明した 3 つのルールのデータが同じ確率分布から出力されたか否かである。1 標本の場合と異なり、分布が仮定できない状況では標本同士を比較する 2 標本コルモゴロフ-スミルノフ検定 (2 標本 KS 検定) を用いる。

定義 5.2 2 標本コルモゴロフ-スミルノフ検定

あるサンプルから得られた m 個のデータ $d_1^1, d_2^1, \dots, d_m^1$ の経験累積分布関数を $F_m^1(d)$ とし、別のサンプルから得られた n 個のデータ $d_1^2, d_2^2, \dots, d_n^2$ の経験累積分布関数を $F_n^2(d)$ とする。

$$F_m^1(d) = \frac{1}{m} \{d_i^1 \leq d \text{ を満たすデータの個数} \}, F_n^2(d) = \frac{1}{n} \{d_j^2 \leq d \text{ を満たすデータの個数} \} \quad (18)$$

2 標本 KS 検定では $F_m^1(d)$ と $F_n^2(d)$ の離れ具合を表す量として次の量を用いる。

$$\begin{aligned} D_{m,n} &= \max(D_{m,n}^+, D_{m,n}^-) \\ D_{m,n}^+ &= \max_{-\infty < d < +\infty} (F_m^1(d) - F_n^2(d)) \\ D_{m,n}^- &= \max_{-\infty < d < +\infty} (F_n^2(d) - F_m^1(d)) \end{aligned} \quad (19)$$

帰無仮説は『得られた標本同士は同一の分布から抽出された』であり、対立仮説は『得られた標本同士は同一の分布から抽出されていない』である。

尚、MATLAB では `kstest2` というコマンドで 2 標本 KS 検定を行う事ができる。ともに 100 標本で、有意水準 5 % として検定を行うとルール 18 と 146 のペアでは帰無仮説が棄却できず、それ以外のペアでは帰無仮説が棄却された。3 種類の分布が登場したので、ウィグナー分布に似た分布以外の分布を分布 A、分布 B とする。

ここで、前節で述べたデータ数について 2 標本 KS 検定を踏まえて説明する。ウィグナー分布との 1 標本 KS 検定を行う際に、500 サンプル全てを用いると 5 章で扱った全ての分布について帰無仮説は棄却された。一方、数値的に $N = 200$ のバイナリランダム行列を 500 個及び 10000 個作り 3.5 節と同じ手順で求めた特異値間隔のデータと ECA から得られたデータで 2 標本 KS 検定を行ったところ、いずれの場合も帰無仮説は棄却されなかった。

以上の結果から、1 標本 KS 検定が棄却される理由は、定義 3.4 に記した通りウィグナー分布が近似分布である事によると思われる。サンプル数が非常に大きいと、 $N = 200$ の真の分布と近似分布の差異が分布の違いとして検出されてしまうのである。この真の分布を解析的に求めることは難しいので、数値的にバイナリランダム行列のデータを作成し 2 標本 KS 検定を行ったのである。

一般の有限サイズ N の chGOE の固有値間隔分布が存在するとして、それを (サイズ N の) 擬ウィグナー分布 $P_N(s)$ と呼ぶことにする。 $P_2(s) = P_{\text{wigner}}(s)$ である。下表は本節で行った KS 検定の結果をまとめたものである。

ルール番号	特異値間隔分布
(クラス 3) 54,60,90,105,122,129,150	(サイズ 200 の) 擬ウィグナー分布
(クラス 3) 18,146	分布 A
(クラス 4) 110	分布 B

5.2 興味深い ECA1：ルール 18,146 の隠れた規則性

1041	0	519	0	537	0	384	0	517	0
0	1017	0	489	0	551	0	535	0	499
519	0	1014	0	543	0	373	0	539	0
0	489	0	975	0	494	0	477	0	486
537	0	543	0	1046	0	383	0	654	0
0	551	0	494	0	1059	0	659	0	530
384	0	373	0	383	0	779	0	391	0
0	535	0	477	0	659	0	1038	0	518
517	0	539	0	654	0	391	0	1041	0
0	499	0	486	0	530	0	518	0	1039

図 10 ルール 18 のサンプルパターンから求めた ${}^tU_r U_r$ の一部

図 10 は $N = 200, M = 500$ で観測されたルール 18 の ${}^tU_r U_r$ の一部で 0 が市松模様状に並んでいる．500 サンプル中 479 サンプルで同様の現象が確認された．これらのサンプルでは U_r の偶数番目と奇数番目のセルのペアに於いて同時刻に 1 が出力されていないことが分かる．

この性質を手掛かりに ${}^tU_r U_r$ が市松模様となる U_r をより詳しく調べたところ，偶数番目のセルの値が全て 0 になっている状態が見つかった．ルール 18 では 3 近傍のうち 001, 100 の 2 つを除き次の時刻の状態は 0 になるので，一旦ある時刻で上述した状態が出現すると次の時刻では奇数番目のセルの値が全て 0 になる．この繰り返しにより， U_r は 0 が市松模様状に配置された構造を持つ．ただし，1 つおきに配置された 0 に挟まれたセルの値が 1 になるとは限らないため U_r の空間方向には 0 が連続して並ぶことができる．これが U_r 自体の市松模様構造を見え難くしている．

尚，ルール 146 の時間発展則は $111 \rightarrow 1$ となる点のみルール 18 と異なる．ルール 146 の ${}^tU_r U_r$ は 500 サンプル中 476 サンプルで市松模様になっていた．

${}^tU_r U_r$ が市松模様になる場合，行（列）の入れ替えにより ${}^tU_r U_r$ を偶数セル同士の相互作用と奇数セル同士の相互作用の直和の形にすることができる．そこで，分離された相互作用から別々に特異値間隔分布を求めると，ウィグナー分布によく一致する結果が得られた．逆にランダム行列を 2 つ用意してそれらの特異値を混ぜ合わせた後でソートし間隔分布を求めればルール 18 や 146 と同様の分布が観察できる．従って，分布 A は本来避けるべき異なるパリティに属する特異値の混在によって生じた無意味な分布であり，直和分解して得られたウィグナー分布を，ルール 18(146) の特異値間隔分布とする方が適切に思えるが，次の理由からそう言い切ることもできない．

ECA の時間発展則にとってバイナリ値の値そのものは本質的ではない．例えば，状態変数を $\{-1, 1\}$ としても同じ分布が観測されるべきである．勿論，分布 A は $\{0, 1\} \rightarrow \{-1, 1\}$ としても観測される事を確認している．しかし， $\{-1, 1\}$ では，一般に ${}^tU_r U_r$ を直和の形にすることができない．以上より，本稿では分布 A をルール 18(146) の間隔分布として掲載した．

本研究で用いた境界条件が周期境界条件であることに注意すると， N が奇数ならばセル番号の偶奇性は意味を持たなくなるので，この規則性は生じない事が予想される． $M = 500, N = 99, 100, 101, 102, 103, 104, 199, 200$ で予想を確認したところ， $N =$ 偶数で分布 A と同様の分布に，システムサイズが奇数のとき擬ウィグナー分布となっていた．

また，ルール 18 で $N = 199$ のときは上述した通り擬ウィグナー分布となっていたが，クラス 4 のルール 110 では同じシステムサイズで擬ウィグナー分布ではないことを KS 検定で確認した．

本稿では，分布 B に限らず，擬ウィグナー分布でなく，かつ， $P(0) \simeq 0$ なる分布をクラス 4 の特異値間隔分布とする．

5.3 興味深い ECA2 : ルール 150 の相対周期

ルール 150 に対して, より詳しい数値実験 $N = 1792, M = 20000$ で得られた結果を示す.

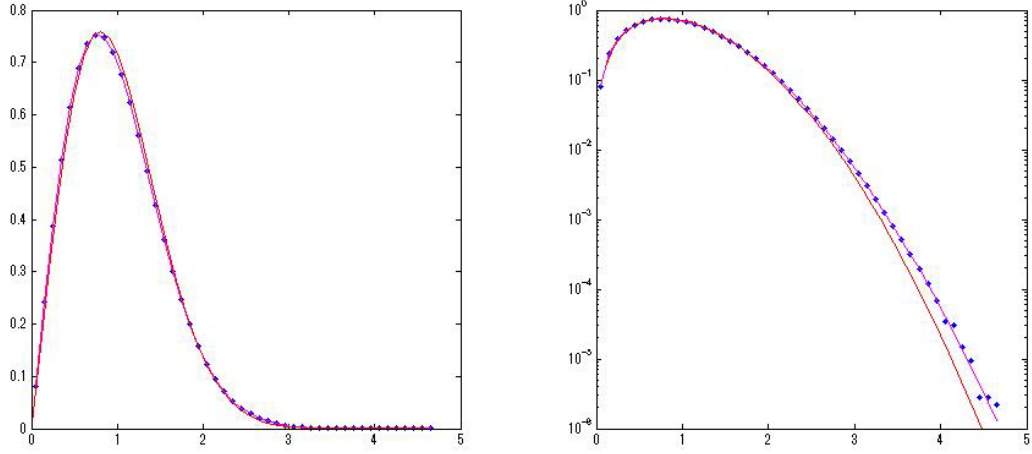


図 11 ルール 150, $N = 1792, M = 20000$ の特異値間隔分布 (点) vs $P_2(s)$ vs $P_\infty(s)$

左図には 2 本の実線が描かれている. 1 本はウィグナー分布であり, もう 1 本は 3.4 節でふれた GOE の $N \rightarrow \infty$ の固有値間隔分布の理論曲線である. 実線は, ほぼ重なって見える. 縦軸を対数プロットした右図に於いて分布裾の値の小さい方がウィグナー分布である. ECA の間隔分布 (点) は極めて正確に理論曲線に一致している. 尚, ランダム行列との一致が高い精度で確認されている例として, リーマンゼータ関数の実部 $1/2$ の零点の虚数部の間隔に関するモンゴメリ・オドリズコの法則が知られている [23].

勿論, 5.1.2 節の最後に記した通り, ECA の分類指標として特異値間隔分布を用いるという文脈では, ここまで精密な数値実験は必要ないだろう. しかし, CA は決定論的な時間発展則に従ってパターンを生成していることに注意すると, そのパターンを行列と見做したときの性質がランダム行列に非常に良く一致することは非自明な結果であり, クラス 3 の特異値間隔分布の理論的な導出は興味深い課題である.

(14) 式のようにアンフォールディングを解析的に行うには, 特異値分布を解析的に求める必要がある. その為には, U_r を 1 つ固定した条件付き特異値分布に対して (数値計算との対応を考えると初期状態を等しい重みでランダムサンプリングしたときの) U_r の出現確率に対する平均を実行しなければならない. しかし, 基本周期や緩和過程の長さ及び種類を解析的に求める事は難しく, 従って U_r の出現確率を求める事も難しい.

ところが, ルール 150 では特別なシステムサイズ $N = 7 \cdot 2^m$ を選ぶと, $m \rightarrow \infty$ で殆ど全ての状態が $T' = N$ なる最大基本相対周期軌道に属することが証明できる.

これは初期状態のランダムサンプリングが, 最大相対周期軌道に対するランダムサンプリングになっていることを意味し, 将来的に先の平均計算を行う上で非常に性質が良いと予想できる.

また, $T' \geq N$ では特異値の最大個数は T' によらず N 個であり, 相対周期は短い方が数値計算にかかる時間を節約できる. ただし, $T' < N$ では自明な特異値の減少が生じてしまう. この事情に留意すると $T' = N$ は最も効率的に特異値のサンプリングを行う事ができる特別な状況である点を強調しておきたい.

定義 5.3 150 は線形ルールであり，その時間発展は下式で与えられる．

$$\text{Rule 150} \quad u_n^{t+1} = u_{n-1}^t + u_n^t + u_{n+1}^t \pmod{2} \quad (20)$$

ルール 150 の状態遷移を表す遷移行列を A_{150} とする．周期境界条件の元， A_{150} はシフト行列 Λ を用いて書く事ができる． E_m は $m \times m$ 単位行列であり， $m = N$ の場合，添え字を省略する．

$$\Lambda = \left(\begin{array}{c|c} 0 & 1 \\ \hline E_{N-1} & 0 \end{array} \right) \quad (21)$$

$$A_{150} = E + \Lambda + \Lambda^{-1} \quad (22)$$

尚，これまでは状態を横ベクトル $\mathbf{u}_t = (u_1^t, u_2^t, \dots, u_N^t)$ で表していたが，本節では縦ベクトル $\mathbf{u}^t = {}^t(u_1^t, u_2^t, \dots, u_N^t)$ で表す．それに伴いベクトルの時刻の添え字を上付きに変更している．

(20) 式は (22) の行列を用いて以下の形に書ける．

$$\mathbf{u}^{t+1} = A_{150} \mathbf{u}^t \pmod{2} \quad (23)$$

定理 5.1 $N = T = 2^m(2^k - 1)$ かつ 3 の倍数ではないとすると， $(E + \Lambda + \Lambda^{-1})^T = E$

証明 5.1 一般に，次式が成り立つ．

$$(E + \Lambda + \Lambda^{-1})^{2^m} = E + \Lambda^{2^m} + \Lambda^{-2^m} \quad (24)$$

$$\begin{aligned} & \text{したがって，} T + 2^m = 2^{k+m} \text{ より，} (E + \Lambda + \Lambda^{-1})^{T+2^m} = (E + \Lambda^{2^m} + \Lambda^{-2^m})^{2^k} \\ & = E + \Lambda^{2^{m+k}} + \Lambda^{-2^{m+k}} = E + \Lambda^{2^m} + \Lambda^{-2^m} = (E + \Lambda + \Lambda^{-1})^{2^m} \end{aligned}$$

ここで， N が 3 の倍数でないとき $E + \Lambda + \Lambda^{-1}$ には逆行列が存在する [16] ので題意が成立．■

以下では特に $N = 7 \cdot 2^m$ の場合を議論する．

定理 5.2 全状態数 $7 \cdot 2^m$ のうち最大基本周期以外の軌道に属す状態の総数は $2^{4 \cdot 2^m}$ 個

証明 5.2 T よりも短い周期軌道は，周期 2^m ，または周期 $7 \cdot 2^{m-1}$ のどちらか（あるいは両方）に属す．ただし，本節でいう「周期」とは，基本周期ではなく，基本周期の倍数も許すことにする．

1. 周期 2^m に属す状態の総数

次の方程式の $\mathbf{u} \in (\mathbb{Z}_2)^T$ に於ける解の個数を求めればよい．

$$(E + \Lambda + \Lambda^{-1})^{2^m} \mathbf{u} = \mathbf{u} \quad (25)$$

(24) 式から $(E + \Lambda^{2^m} + \Lambda^{-2^m}) \mathbf{u} = \mathbf{u}$ より，周期 2^m を持つものでは， $r := 2^m$ として，任意の n に対して $u_{n-r} + u_{n+r} \equiv 0 \pmod{2}$ が成り立つ (u の添え字は \pmod{T} で数えるものとする)．

$$\forall n, u_n \equiv u_{n+2r}$$

一方， $\text{GCM}(2r, T) = r$ であるので，周期境界条件より，求める解は次の条件を満たす．

$$\forall n, u_n = u_{n+r} = u_{n+2r} = \dots = u_{n+6r}$$

このような状態の数は 2^{2^m} 個である．

2. 周期 $7 \cdot 2^{m-1}$ に属す状態の総数

$$(E + \Lambda + \Lambda^{-1})^{7 \cdot 2^{m-1}} \mathbf{u} = \mathbf{u} \quad (26)$$

(24) 式から $(E + \Lambda^{2^{m-1}} + \Lambda^{-2^{m-1}})^7 \mathbf{u} = \mathbf{u}$ より, $s := 2^{m-1}, \Gamma := \Lambda^s$ として計算を行うと,

$$(\Gamma^{-6} + \Gamma^{-4} + \Gamma^{-3} + \Gamma^{-1} + E + \Gamma^1 + \Gamma^3 + \Gamma^4 + \Gamma^6) \mathbf{u} = \mathbf{u} \quad (27)$$

これから, 下式の解の個数を求めればよい.

$$(\Gamma^{-6} + \Gamma^{-4} + \Gamma^{-3} + \Gamma^{-1} + \Gamma^1 + \Gamma^3 + \Gamma^4 + \Gamma^6) \mathbf{u} = \mathbf{0}$$

具体的には,

$$\forall n, u_{n-6s} + u_{n-4s} + u_{n-3s} + u_{n-s} + u_{n+s} + u_{n+3s} + u_{n+4s} + u_{n+6s} \equiv 0$$

この式は 14 個の変数 $\{u_{n+ks}\}_{k=-6}^7$ で閉じているので,

$$\mathbf{u}' := {}^t(u_{n-6s}, u_{n-5s}, u_{n-4s}, \dots, u_{n+5s}, u_{n+6s}, u_{n+7s})$$

を用いて, 14 元連立方程式

$$(\Lambda^{-6} + \Lambda^{-4} + \Lambda^{-3} + \Lambda^{-1} + \Lambda^1 + \Lambda^3 + \Lambda^4 + \Lambda^6) \mathbf{u}' = \mathbf{0} \quad (28)$$

を考える. その解の個数を K とすると, もとの方程式の解の個数は K^s で与えられる. 上式の左辺の行列の階数は 6 より, $K = 2^{14-6} = 2^8$. よって, 求める状態の個数は $(2^8)^s = 2^{4 \cdot 2^m}$

3. 周期 2^{m-1} に属す状態の総数

周期 2^m のものと, 周期 $7 \cdot 2^{m-1}$ のものの共通部分は, 周期 s である. それらは,

$$(E + \Lambda + \Lambda^{-1})^s \mathbf{u} = (E + \Lambda^s + \Lambda^{-s}) \mathbf{u} = \mathbf{u}$$

を満たすので, $\forall n, u_{n-s} + u_{n+s} \equiv 0 \pmod{2}$. したがって, $r = 2^n$ として下式が成立.

$$\forall n, u_n = u_{n+r} = u_{n+2r} = \dots = u_{n+6r}$$

このような状態の数は 2^{2^m} 個である. 以上より, 全状態数 $7 \cdot 2^m$ のうち最大基本周期以外の軌道に属す状態の総数は, $2^{4 \cdot 2^m} + 2^{2^m} - 2^{2^m} = 2^{4 \cdot 2^m}$ となる. ■

ここまでの議論から $N = 7 \cdot 2^m, m \rightarrow \infty$ のとき, ルール 150 の状態は殆どの場合, 最大基本周期の軌道に属することが分かった. ここからは相対周期に対する議論を行い, 次の定理を得る.

定理 5.3 ルール 150 で $N = 7 \cdot 2^m$ とする. $m \rightarrow \infty$ のとき, ルール 150 の殆ど全ての状態は, 最大基本相対周期 $T' = 7 * 2^m$ の軌道に属す.

証明 5.3 T よりも短い相対周期軌道は, $T' = 2^m, S_1 = 2^m, S_2 = 2 \cdot 2^m, S_3 = 3 \cdot 2^m$ または $T' = 7 \cdot 2^{m-1}, S_4 = N/2$ のどちらか (あるいは両方) に属す. ただし, 本節でいう「相対周期」には基本相対周期の倍数も含まれることとする.

1. $T' = 2^m, S_1 = 2^m$ に属す状態の総数

$$(E + \Lambda + \Lambda^{-1})^{2^m} \mathbf{u} = \Lambda^{2^m} \mathbf{u}$$

より, $\Lambda^{2^m} \mathbf{u} = \mathbf{u}$ を満たす解を求めれば良い. $r = 2^n$ として下式が成立する.

$$\forall n, u_n = u_{n+r} = u_{n+2r} = \cdots = u_{n+6r} \quad (29)$$

このような状態の数は 2^{2^m} 個である. ここで, (29) 式の表す空間方向に対して周期 2^m を持つ状態は, 周期境界条件を課した $N = 2^m$ の系と見なすことができる. その場合, $\Lambda^{N/2} = \Lambda^{-N/2}$ より, $A_{150}^{N/2} = E$ より $T = 2^{m-1}$ となる. したがって, 今回考えた状態は $T = 2^m$ の議論で既にカウントされていたことになる.

2. $T' = 2^m, S_1 = 2 \cdot 2^m$ に属す状態の総数

$$(E + \Lambda + \Lambda^{-1})^{2^m} \mathbf{u} = \Lambda^{2 \cdot 2^m} \mathbf{u} \quad (30)$$

上式に加えて, $t = 2 \cdot 2^m$ に対して下式が成立する.

$$(E + \Lambda + \Lambda^{-1})^{2 \cdot 2^m} \mathbf{u} = \Lambda^{4 \cdot 2^m} \mathbf{u} \quad (31)$$

(30),(31) 式を展開して整理すると下式が得られる.

$$(\Lambda^{4 \cdot 2^m} + \Lambda^{-4 \cdot 2^m}) \mathbf{u} = (\Lambda^{2 \cdot 2^m} + \Lambda^{-2 \cdot 2^m}) \mathbf{u} \quad (32)$$

(32) 式の両辺に \mathbf{u} を加えると,

$$A_{150}^{4 \cdot 2^m} \mathbf{u} = A_{150}^{2 \cdot 2^m} \mathbf{u} \quad (33)$$

を得る. これは $T = 2^m$ の状態であり, 先程と同様に既にカウントされている.

3. $T' = 2^m, S_1 = 3 \cdot 2^m$ に属す状態の総数

上と同様にして, $t = 2^m, 2 \cdot 2^m, 3 \cdot 2^m$ に対して次の 3 つの式が成立する.

$$(E + \Lambda + \Lambda^{-1})^{2 \cdot 2^m} \mathbf{u} = \Lambda^{6 \cdot 2^m} \mathbf{u} \quad (34)$$

$$(E + \Lambda + \Lambda^{-1})^{5 \cdot 2^m} \mathbf{u} = \Lambda^{2^m} \mathbf{u} \quad (35)$$

$$(E + \Lambda + \Lambda^{-1})^{6 \cdot 2^m} \mathbf{u} = \Lambda^{4 \cdot 2^m} \mathbf{u} \quad (36)$$

(34),(35),(36) 式を適当に整理して次式を得る.

$$A_{150}^{2^m} \mathbf{u} = \Lambda^{4 \cdot 2^m} \mathbf{u} \quad (37)$$

(37) 式と $A_{150}^{2^m} \mathbf{u} = \Lambda^{3 \cdot 2^m} \mathbf{u}$ から空間方向に周期 2^m を持つ解が得られる. $T' = 2^m, S_1 = 2^m$ の場合と同様に, これらの解の周期は $T = 2^{m-1}$ となり, 既にカウントされている.

結局, $T' = 2^m$ となる状態は全て $T = 2^m$ の状態として既にカウントされていたことになる.

4. $T' = 7 \cdot 2^{m-1}, S_4 = N/2$ に属す状態の総数

$$(E + \Lambda + \Lambda^{-1})^{7 \cdot 2^{m-1}} \mathbf{u} = \Lambda^{7 \cdot 2^{m-1}} \mathbf{u} \quad (38)$$

に対して (27) 式と同じ処理を施すと, (28) 式に対応する次の式を得る.

$$(\Lambda^{-6} + \Lambda^{-4} + \Lambda^{-3} + \Lambda^{-1} + E + \Lambda^1 + \Lambda^3 + \Lambda^4 + \Lambda^6 + \Lambda^7)\mathbf{u}' = \mathbf{0} \quad (39)$$

上式の左辺の行列の階数は 7 より, 求める状態の個数は $2^{7 \cdot 2^{m-1}}$ 個.

議論の短縮のため, この $2^{7 \cdot 2^{m-1}}$ の状態とこれまでに登場した相対周期が 2^m の状態の共通部分の個数は求めず, $T' = N$ 未満の状態の個数を上から評価する. 周期が N 未満の状態の相対周期は明らかに N 未満なのでそれらの総数は $2^{4 \cdot 2^m} + 2^{7 \cdot 2^{m-1}}$ 以下. よって最大基本相対周期が N 未満の状態の割合は次式以下となる.

$$\frac{2^{4 \cdot 2^m} + 2^{7 \cdot 2^{m-1}}}{2^{7 \cdot 2^m}} \quad (40)$$

$m \rightarrow \infty$ の極限で, 殆ど全ての状態は $T' = N$ の最大基本相対周期軌道に属している. ■

以上の結果から, ルール 150 で $N = 7 \cdot 2^m$ の場合, 効率的に基本相対周期軌道がサンプリングできることが分かった.

本節の締めくくりとして, 上記の性質に着目し特異値間隔分布以外の比較を行ってみよう. ランダム行列理論では最大固有値分布や最小固有値分布も詳しく調べられているが, 背景パターンの特徴付けとしては chGOE の最小固有値分布との比較が興味深い (間隔分布の計算では特異値サンプルの端に位置する最小特異値は捨てられていることに注意する).

ここで, 最小特異値は 1 サンプルにつき 1 個である. したがって, その統計性を議論する為には (ランク落ちが無いと仮定したときの間隔の個数が $\mathcal{O}(N)$ なのに対して) サンプル数を十分に取らなければならない. この要請に対して効率的なサンプリングが可能な点は好都合なのである.

尚, $Q = 1$ の chGOE のアンフォールドされた最小固有値の分布は (41) 式で与えられる [43]. 下図はルール 150 で $N = 112, M = 50000$ としたときの比較である. 良い一致が見られる.

ただし, 最小固有値分布の形は行列の縦横差に依存する [44] ため, CA の分類基準として用いるには本稿で分類基準に対して要請した条件 (v) を満たしていない.

$$p_{smallest}(x) = -\frac{d}{dx} \exp\left(-\frac{\pi}{2}x\right) \exp\left(-\frac{\pi^2}{8}x^2\right) \quad (41)$$

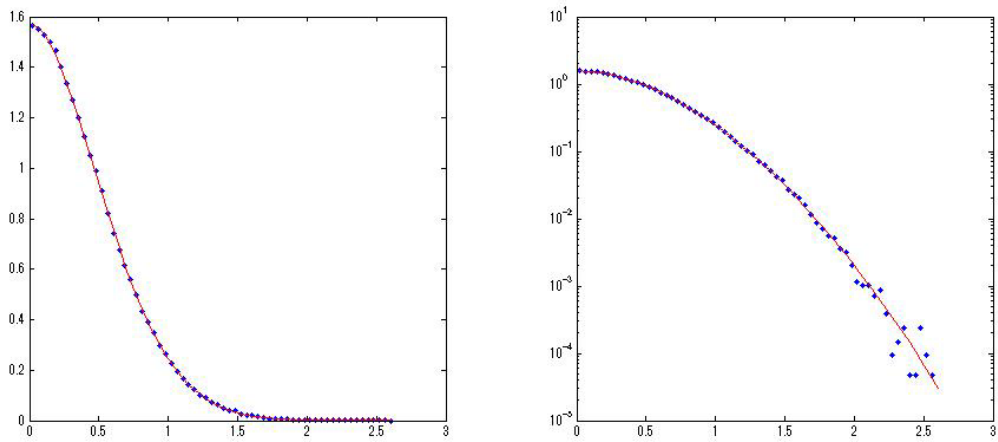


図 12 $N = 112, M = 50000$ のルール 150 と chGOE の最小特異値分布の比較

5.4 興味深い ECA3 : ルール 90 とルール 150 の関係

ルール 90 とルール 150 はともに線形ルールで代数的に扱いやすいため、他のルールに比べると、かなり研究されている [4].

前節ではルール 150 の特別なシステムサイズ $N = 7 \cdot 2^m$ の性質を議論した. ルール 150 はシステムサイズが 3 の倍数でないとき可逆であり, 初期状態への回帰が保障される為, 緩和過程に留意する必要がなく数値計算を行う上で非常に都合が良い.

一方, ルール 90 は特異値の効率的なサンプリングという目的に対してルール 150 ほど便利な性質が現時点では発見できていない. しかし, ルール 90 は二項展開との対応やフラクタル性, max-plus 代数を用いて時間発展を簡潔に書ける点 [42] など, 本研究とは別の観点から盛んに研究されており, これらのルールを結ぶ研究は, 今後, 互いの知見を利用する上でも有益と思われる.

乱雑な初期値から時間発展したルール 90, 150 の出力パターンは Appendix-A に掲載しており, ともに三角波がカオス的に生じるクラス 3 の特徴が観察できる. しかし, 当然のことながら, その時間発展パターンは同じではない. ところが, $N = 7 \cdot 2^m$ の場合, これらのパターンは全く無関係ではなく, 対応関係見られるというのが以下の内容である. 参考までに, ルール 90 とルール 150 の $N = 200$ で 101 番目のセルのみ 1 とした初期値から生成したパターンを示しておく.

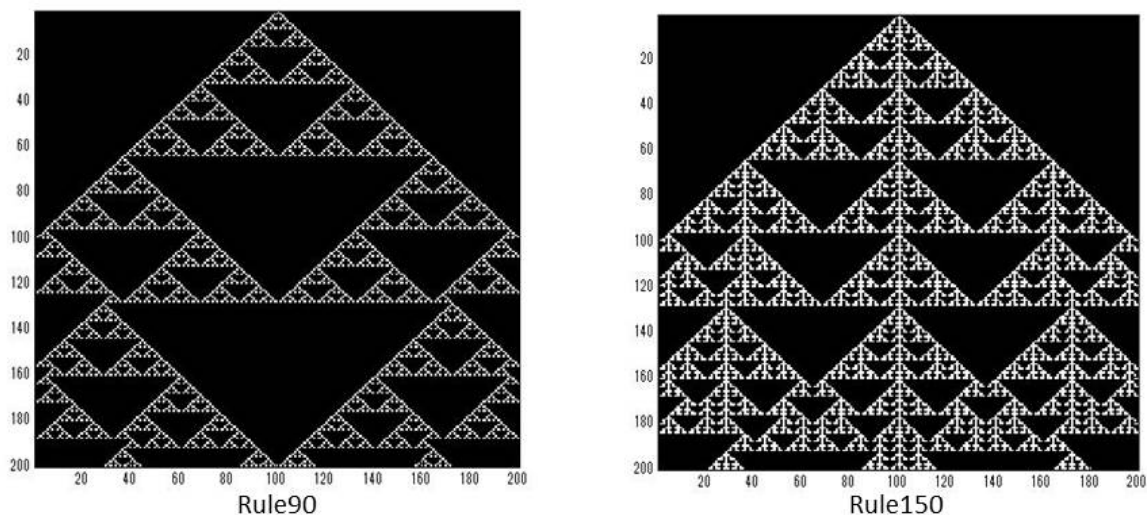


図 13 ルール 90 とルール 150 の時間発展の比較

定義 5.4 ルール 90 は線形ルールであり, その時間発展は下式で与えられる.

$$u_n^{t+1} = u_{n-1}^t + u_{n+1}^t \pmod{2} \quad (42)$$

ここで, 周期的境界条件を課したルール 90 の遷移行列を A_{90} とすると, $A_{90} = \Lambda + \Lambda^{-1}$ なので, この 2 つのルールの遷移行列間には以下の関係が成り立つ.

$$A_{150} = E + A_{90} \quad (43)$$

定理 5.1 の内容と被る部分もあるが, 具体的に時間発展を書いてみる.

5.4.1 $N = 7$ の場合

まず，具体的なシステムサイズでシフト行列による遷移行列の表示を使い最大周期を計算する．

$$\begin{aligned}
A_{90} &= \Lambda + \Lambda^{-1} \\
A_{90}^2 &= \Lambda^2 + \Lambda^{-2} \\
A_{90}^3 &= \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} \\
A_{90}^4 &= \Lambda^4 + \Lambda^{-4} = \Lambda^3 + \Lambda^{-3} \quad \star \\
A_{90}^5 &= \Lambda^2 + \Lambda^{-2} + \Lambda^3 + \Lambda^{-3} \\
A_{90}^6 &= \Lambda + \Lambda^{-1} + \Lambda^2 + \Lambda^{-2} \\
A_{90}^7 &= \Lambda + \Lambda^{-1} + \Lambda^2 + \Lambda^{-2} + \Lambda^3 + \Lambda^{-3} \\
A_{90}^8 &= \Lambda^8 + \Lambda^{-8} = \Lambda + \Lambda^{-1} = A_{90}
\end{aligned}$$

上式から $N = 7$ のときルール 90 の最大周期は 7 で最大緩和長は 1 であることが分かる．計算機で全探索を行うと，全ての成分が 0 または 1 の初期状態の周期は 1 でそれ以外は 7 であった． \star 印以降ではシフト行列の性質 $\Lambda^k = \Lambda^{k-N}$ を用いた．

$$\begin{aligned}
A_{150} &= E + \Lambda + \Lambda^{-1} \\
A_{150}^2 &= E + \Lambda^2 + \Lambda^{-2} \\
A_{150}^3 &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^3 + \Lambda^{-3} \\
A_{150}^4 &= E + \Lambda^3 + \Lambda^{-3} \\
A_{150}^5 &= E + \Lambda + \Lambda^{-1} + \Lambda^2 + \Lambda^{-2} \\
A_{150}^6 &= E + \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} \\
A_{150}^7 &= E
\end{aligned}$$

$N = 7$ のとき可逆で，ルール 90 と同様に最大周期は 7 であることが分かる．

定理 5.4 ルール 150 の軌道の状態を各々 A_{90} で遷移させた状態は再びルール 150 の軌道を作る．

証明 5.4 $\mathbf{u}^{t+1} = A_{150}\mathbf{u}^t$ とすると，

$$\mathbf{u}^{t+1} + \mathbf{u}^t = (A_{150} + E)\mathbf{u}^t = A_{90}\mathbf{u}^t \quad (44)$$

$\mathbf{u}^{t+2} = A_{150}\mathbf{u}^{t+1}$ から， $\mathbf{u}^{t+2} + \mathbf{u}^{t+1} = A_{90}\mathbf{u}^{t+1}$ なので， $A_{90}\mathbf{u}^t$ から $A_{90}\mathbf{u}^{t+1}$ への遷移行列を B とすれば， $A_{90}\mathbf{u}^{t+1} = BA_{90}\mathbf{u}^t$ より， $A_{90}A_{150}\mathbf{u}^t = BA_{90}\mathbf{u}^t$ ここで，単位行列及びシフト行列は可換なので，それらを用いて書かれる A_{90} と A_{150} も可換であり，

$$A_{150}A_{90}\mathbf{u}^t = BA_{90}\mathbf{u}^t \quad (45)$$

よって， $B = A_{150}$ となる．したがって，ルール 150 の軌道の状態を各々 A_{90} で遷移させた状態は再びルール 150 の軌道を作る．■

次に， A_{90} で遷移させる前後でルール 150 の軌道を比較した場合，それらにはどのような関係があるのかを考える．

その為に、ルール 90,150 の軌道の関係を図示した図 14 を使う．これを三角図と呼ぶことにする．

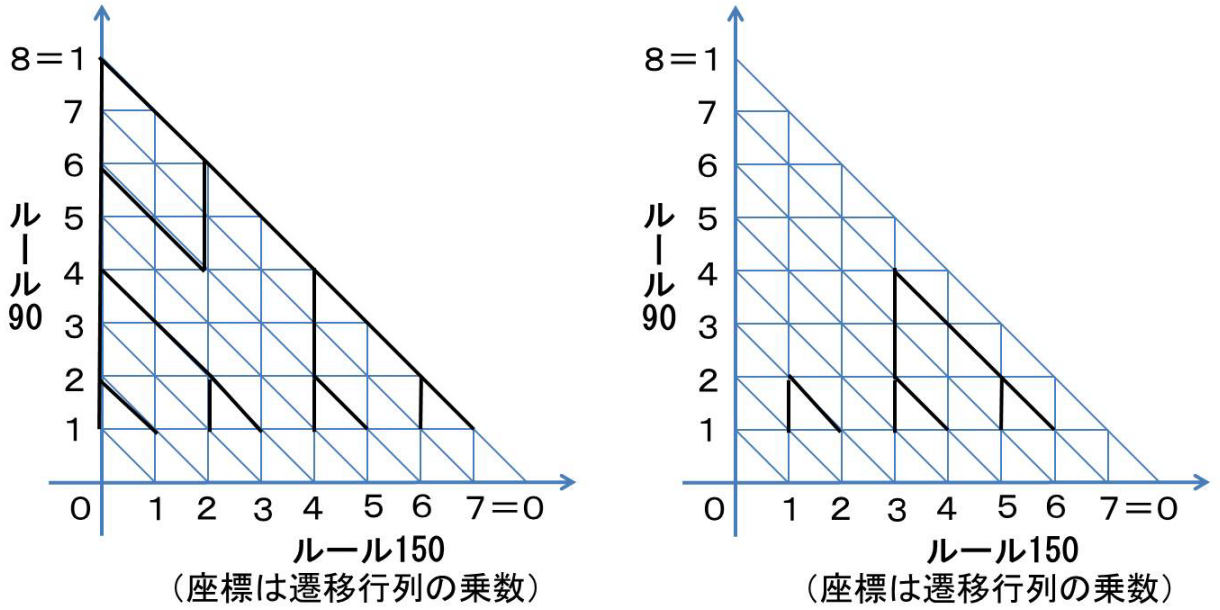


図 14 三角図

定義 5.5 三角図

横軸がルール 150 の軌道を表し，座標は遷移行列の乗数である ($7=0$ は $A_{150}^7 = A_{150}^0 = E$)．初期状態 \mathbf{x}^0 が座標原点で，縦軸はルール 90 の軌道を表す．縦軸のルール 90 の状態をルール 150 の軌道の状態の 1 次結合で表すことができる．

与えられたルール 150 の軌道の初めの 3 つの状態を $\mathbf{u}^0, \mathbf{u}^1, \mathbf{u}^2$ とすると， $A_{90}\mathbf{u}^0 = \mathbf{u}^0 + \mathbf{u}^1$ で， $A_{90}\mathbf{u}^1 = \mathbf{u}^1 + \mathbf{u}^2$ より，

$$A_{90}^2\mathbf{u}^0 = \mathbf{u}^0 + \mathbf{u}^1 + \mathbf{u}^1 + \mathbf{u}^2 = \mathbf{u}^0 + \mathbf{u}^2 \quad (46)$$

三角図では線形結合に寄与する部分を太線で表しており，太線がフラクタルパターンになるのは上式に基づいて計算を行なったからである．

補足として，本フラクタルパターンは原点のみを 1 として残りは 0 としたルール 60 の時間発展パターンに一致する．ルール 60,90,150 は全て線形ルールである．

次に， $A_{90}^8\mathbf{x}^0$ を縦軸の座標 = 1 のルール 150 の軌道の状態の線形結合で表すと，

$$A_{90}^8\mathbf{u}^0 = A_{90}(\mathbf{u}^0 + \mathbf{u}^1 + \mathbf{u}^2 + \mathbf{u}^3 + \mathbf{u}^4 + \mathbf{u}^5 + \mathbf{u}^6 + \mathbf{u}^7) \quad (47)$$

$\mathbf{u}^7 = \mathbf{u}^0$ 及び $A_{90}^8 = A_{90}$ より上式は，

$$A_{90}\mathbf{u}^0 = A_{90}(\mathbf{u}^1 + \mathbf{u}^2 + \mathbf{u}^3 + \mathbf{u}^4 + \mathbf{u}^5 + \mathbf{u}^6) \quad (48)$$

今度は逆に右辺の和を求める．右図のように上方向に計算すれば良いので，

$$A_{90}\mathbf{u}^0 = A_{90}^2A_{150}\mathbf{u}^0 + A_{90}^4A_{150}^3\mathbf{u}^0 \quad (49)$$

(49) 式右辺 $A_{90}^2 A_{150}^4 (A_{90}^2 A_{150}^2 + E)$ において,

$$A_{90}^2 A_{150}^2 + E = (\Lambda^2 + \Lambda^{-2})(1 + \Lambda^2 + \Lambda^{-2}) + E = \Lambda^3 + \Lambda^{-3} + \Lambda^2 + \Lambda^{-2} + E = A_{150}^3 \quad (50)$$

よって, (49) 式は

$$A_{90} \mathbf{u}^0 = A_{90} A_{150}^4 A_{90} \mathbf{u}^0 \quad (51)$$

ルール 90 の周期軌道の任意の状態を初期状態に取れるので $\mathbf{w}^0 = A_{90} \mathbf{u}^0$ とすれば

$$\mathbf{w}^0 = A_{90} A_{150}^4 \mathbf{w}^0 \quad (52)$$

(52) 式は, $N = 7$ でルール 90 の周期軌道に属す状態を 1 つ選んで, その状態を初期状態とするルール 150 の軌道を作り, 次に軌道の各状態をルール 90 でそれぞれ遷移させて新しいルール 150 の軌道を作ると, 新しい軌道はもとの軌道を時間方向に 4 つ分シフトしたのになっている事を意味する (時間方向のシフトであって, 空間方向のシフトではない点に注意する).

また, 時間方向のシフト数 4 と基本周期 7 は互いに素なので, ルール 90 の周期軌道には初めに作成したルール 150 の軌道を構成する状態が 1 回ずつ登場する.

$N = 7$ の場合, ルール 90 の周期状態に属す初期状態から時間発展したルール 90 とルール 150 の軌道は状態の並び方が異なるだけである.

5.4.2 $N = 14$ の場合

次に $N = 14$ のルール 90 とルール 150 の軌道の対応関係を考える. 基本的な方針は同じだが, システムサイズが偶数なので, 3.6.2 節にルール 90 の性質として記した, $T' = T/2$ という現象が起こるかを確認する.

【ルール 90】

$$\begin{aligned} A_{90} &= \Lambda + \Lambda^{-1} \\ A_{90}^2 &= \Lambda^2 + \Lambda^{-2} \\ A_{90}^3 &= \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} \\ A_{90}^4 &= \Lambda^4 + \Lambda^{-4} \\ A_{90}^5 &= \Lambda^3 + \Lambda^{-3} + \Lambda^5 + \Lambda^{-5} \\ A_{90}^6 &= \Lambda^2 + \Lambda^{-2} + \Lambda^6 + \Lambda^{-6} \\ A_{90}^7 &= \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} + \Lambda^5 + \Lambda^{-5} \\ A_{90}^8 &= \Lambda^6 + \Lambda^{-6} & \star \\ A_{90}^9 &= \Lambda^5 + \Lambda^{-5} \\ A_{90}^{10} &= \Lambda^4 + \Lambda^{-4} + \Lambda^6 + \Lambda^{-6} \\ A_{90}^{11} &= \Lambda^3 + \Lambda^{-3} \\ A_{90}^{12} &= \Lambda^2 + \Lambda^{-2} + \Lambda^4 + \Lambda^{-4} \\ A_{90}^{13} &= \Lambda + \Lambda^{-1} + \Lambda^5 + \Lambda^{-5} \\ A_{90}^{14} &= \Lambda^2 + \Lambda^{-2} + \Lambda^4 + \Lambda^{-4} + \Lambda^6 + \Lambda^{-6} \\ A_{90}^{15} &= A_{90} \end{aligned}$$

【ルール 150】

$$\begin{aligned}
A_{150} &= E + \Lambda + \Lambda^{-1} \\
A_{150}^2 &= E + \Lambda^2 + \Lambda^{-2} \\
A_{150}^3 &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^3 + \Lambda^{-3} \\
A_{150}^4 &= E + \Lambda^4 + \Lambda^{-4} \\
A_{150}^5 &= E + \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} + \Lambda^4 + \Lambda^{-4} \\
A_{150}^6 &= E + \Lambda^4 + \Lambda^{-4} + \Lambda^6 + \Lambda^{-6} \\
A_{150}^7 &= E + \Lambda + \Lambda^{-1} + \Lambda^3 + \Lambda^{-3} + \Lambda^4 + \Lambda^{-4} + \Lambda^6 + \Lambda^{-6} \\
A_{150}^8 &= E + \Lambda^6 + \Lambda^{-6} \\
A_{150}^9 &= E + \Lambda + \Lambda^{-1} + \Lambda^5 + \Lambda^{-5} + \Lambda^6 + \Lambda^{-6} \\
A_{150}^{10} &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^4 + \Lambda^{-4} \\
A_{150}^{11} &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^4 + \Lambda^{-4} + \Lambda^5 + \Lambda^{-5} \\
A_{150}^{12} &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^6 + \Lambda^{-6} \\
A_{150}^{13} &= E + \Lambda^2 + \Lambda^{-2} + \Lambda^3 + \Lambda^{-3} + \Lambda^5 + \Lambda^{-5} + \Lambda^6 + \Lambda^{-6} \\
A_{150}^{14} &= E
\end{aligned}$$

三角図を描いて計算すると、今回は下式が得られる。

$$A_{90}\mathbf{u}^0 = A_{90}^2 A_{150}^8 A_{90}\mathbf{u}^0 \quad (53)$$

ルール 90 の周期軌道の任意の状態を初期状態に取れるので $\mathbf{w}^0 = A_{90}\mathbf{u}^0$ とすれば

$$\mathbf{w}^0 = A_{90}^2 A_{150}^8 \mathbf{w}^0 \quad (54)$$

よって、ルール 90 で初期状態を 2 回遷移させた状態が属すルール 150 の軌道は、初期状態の属していた軌道を時間 8 つ分シフトしたものである事が分かる。 $14 = 7 \times 2$ と $8 = 2 \times 2 \times 2$ より、7 回のシフトで初期状態に回帰する。

問題になるのは、 A_{90} 方向 2 回の遷移はもとの軌道のシフトであるが、1 回遷移した軌道はどのような状態から構成されているかである。ここで、ルール 90 の☆印に着目する。 $\Lambda^8 = \Lambda^{-6}$ より、

$$A_{90}^8 = \Lambda^6 + \Lambda^{-6} = \Lambda^8 + \Lambda^6 = \Lambda^7(\Lambda + \Lambda^{-1}) = \Lambda^7 A_{90} \quad (55)$$

$A_{90}^7 \mathbf{w}^0 = \Lambda^7 \mathbf{w}^0$ であり、 A_{90} で初期状態を 7 回遷移した状態は初期状態を空間方向に半分シフトした状態になっている。7 回遷移は 2 回遷移の間にあるので、問題にした A_{90} で 1 回遷移した軌道は、元の軌道の状態を半分シフトしたもので構成されていることが分かる。1 度シフトパターンが出現すればそれ以降はシフトパターンなので $T' = T/2$ である。

ルール 90 の周期状態に属す状態から作ったルール 150 の軌道の $t = 7$ の状態が初期状態の $S = 7$ のシフトパターンならば、 $A_{90}^8 \mathbf{w}^0 = \mathbf{w}^{10} = \Lambda^7 \mathbf{w}^3$ と $A_{90}^7 \mathbf{w}^0 = \Lambda^7 \mathbf{w}^0$ から $\Lambda^7 \mathbf{w}^3 = A_{90} \Lambda^7 \mathbf{w}^0$ となる。よって、 $\mathbf{w}^3 = A_{90} \mathbf{w}^0$ 、これは A_{90} で 1 回遷移させた軌道は元の軌道を時間 3 つ分シフトしたものであることを意味しており、2 回遷移で時間 8 つ分シフトするという性質と合わせると、 $T = 2$ である。これは、 $A_{150}^2 \mathbf{w}^0 = \mathbf{w}^0$ より、任意の n に対して $w_n^t = w_{n+2}^t$ を満たさねばならず、そのようなものでかつルール 90 の周期状態に属す元は $\mathbf{0}$ のみである。

5.4.3 三角図から得られる結果の一般化

5.4 節の締め括りとして、 $N = 7, 14$ でルール 90 と 150 を結ぶ役割を果たした三角図が $N = 7 \cdot 2^m$ の場合にどのようなものかを議論する。

三角図に於いて、 $A_{90}^{7 \cdot 2^m}$ の値をルール 150 の軌道の状態の和で表すとき、和に現れる状態の間隔は 2^m になっている ($m = 0$ も許す)。これは三角図がフラクタルパターンとなることから明らかであるが、以下に文字を使った証明を記しておく。

定理 5.5 $A_{90}^{7 \cdot 2^m} = \Lambda^{2^m} + \Lambda^{2 \cdot 2^m} + \Lambda^{3 \cdot 2^m} + \Lambda^{4 \cdot 2^m} + \Lambda^{5 \cdot 2^m} + \Lambda^{6 \cdot 2^m}$ が成り立つ。

証明 5.5 該当する箇所の A_{150} の冪乗を計算すると、

$$\begin{aligned} A_{150}^{2^m} &= E + \Lambda^{2^m} + \Lambda^{-2^m} & A_{150}^{2 \cdot 2^m} &= E + \Lambda^{2 \cdot 2^m} + \Lambda^{-2 \cdot 2^m} \\ A_{150}^{3 \cdot 2^m} &= E + \Lambda^{2^m} + \Lambda^{-2^m} + \Lambda^{3 \cdot 2^m} + \Lambda^{-3 \cdot 2^m} & A_{150}^{4 \cdot 2^m} &= E + \Lambda^{4 \cdot 2^m} + \Lambda^{-4 \cdot 2^m} \\ A_{150}^{5 \cdot 2^m} &= E + \Lambda^{2^m} + \Lambda^{-2^m} + \Lambda^{2 \cdot 2^m} + \Lambda^{-2 \cdot 2^m} & A_{150}^{6 \cdot 2^m} &= E + \Lambda^{2^m} + \Lambda^{-2^m} + \Lambda^{3 \cdot 2^m} + \Lambda^{-3 \cdot 2^m} \end{aligned}$$

これらを加えると定理 5.5 の右辺に一致する。■

ルール 90 には緩和状態が存在する (可逆ではない)。緩和時間を上から押さえる定理を示す。

定理 5.6 $N = 7 \cdot 2^m$ のとき、ルール 90 の緩和時間は 2^{m-1} 以下である。

証明 5.6 $A_{90}^{2^{m-1}} A_{90}^{7 \cdot 2^m} = (\Lambda + \Lambda^{-1})^{2^{m-1}} (\Lambda + \Lambda^{-1})^{7 \cdot 2^m}$ ここで、定理 5.5 を用いて展開し、整理すると、 $A_{90}^{2^{m-1}} A_{90}^{7 \cdot 2^m} = \Lambda^{2^{m-1}} + \Lambda^{-2^{m-1}} = A_{90}^{2^{m-1}}$ を得る。■

ルール 90 の最大基本相対周期の上限を与える定理を示す。

定理 5.7 $N = 7 \cdot 2^m$ のとき、ルール 90 の最大基本相対周期は $7 \cdot 2^{m-1}$ 以下である。

証明 5.7 $A_{90}^{2^{m-1}} A_{90}^{7 \cdot 2^m} = \Lambda^{4 \cdot 2^m} + \Lambda^{3 \cdot 2^m} = \Lambda^{8 \cdot 2^{m-1}} + \Lambda^{6 \cdot 2^{m-1}} = \Lambda^{7 \cdot 2^{m-1}} (\Lambda^{2^{m-1}} + \Lambda^{-2^{m-1}}) = \Lambda^{7 \cdot 2^{m-1}} A_{90}^{2^{m-1}}$ を得る。■

最後に、(52),(54) 式の一般化に当たるものを導出する。

定理 5.8 \mathbf{w}^0 がルール 90 の Ω_c の元であるとする。このとき、 $\mathbf{w}^0 = A_{90}^{2^m} A_{150}^{4 \cdot 2^m} \mathbf{w}^0$ が成立する。

証明 5.8 証明 5.5 に記した A_{150} の冪乗の展開を用いると、

$$A_{150}^{2^m} + A_{150}^{2 \cdot 2^m} = A_{150}^{2^m} (E + \Lambda^{2^m} + \Lambda^{-2^m} + E) = A_{150}^{2^m} A_{90}^{2^m} \quad (56)$$

$$A_{150}^{3 \cdot 2^m} + A_{150}^{4 \cdot 2^m} + A_{150}^{5 \cdot 2^m} + A_{150}^{6 \cdot 2^m} = (A_{150}^{3 \cdot 2^m} + A_{150}^{5 \cdot 2^m}) A_{90}^{2^m} \quad (57)$$

(56),(57) 式の和を計算することで下式を得る。

$$A_{90}^{2^m} (A_{150}^{2^m} + A_{150}^{3 \cdot 2^m} + A_{150}^{5 \cdot 2^m}) = A_{90}^{2^m} A_{150}^{2^m} (E + \Lambda^{2^m} + \Lambda^{-2^m} + E) = A_{90}^{2^m} A_{150}^{2^m} A_{150}^{3 \cdot 2^m} = A_{90}^{2^m} A_{150}^{4 \cdot 2^m} \quad (58)$$

(58) 式右辺の値は $A_{90}^{7 \cdot 2^m}$ に等しい。ここで、定理 5.6 より、 $A_{90}^{7 \cdot 2^m} \mathbf{w}^0 = \mathbf{w}^0$ となり題意成立。■

5.3 節でルール 150 の特別なシステムサイズを議論したが [38]、このシステムサイズはルール 90 の場合にも特殊な性質を持つことが分かった。

5.5 興味深い ECA4 : ルール 170 の固有値分布

ルール 150 と同様に, ECA の分類を一旦忘れ, U_c の特異値間隔分布に着目したとき, その分布の形がポアソン分布によく似ているという結果は興味深い. 現時点では表層的な類似に過ぎないが, 量子系のエネルギー固有値の間隔分布は, 対応する古典系が可積分の場合にポアソン分布に, カオスの場合にウィグナー分布になることが知られている (Appendix-C 参照).

本節ではルール 170 の固有値分布を解析的な導出を試みる. 間隔分布まで解析的に求めることはできなかったが, アンフォールディングを解析的行うためには, 固有値分布が必要となる.

尚, 以下で扱う巡回行列は正確にはルール 240 の出力パターンであるが左右反転で 2 つのルールは同一視されるため, この分野の慣習に従い, 若い方のルール番号 170 を用いる.

5.5.1 巡回行列

シフトパターンはそれを行列とみなせば, 巡回行列である. その性質を復習する. 巡回行列 U_c の形は以下の通りで, 1 行目が右方向に循環的に 1 つずつシフトする. 空間変数を表す添え字は $1 \leq n \leq N$ としていたが, 本節では便宜上 $0 \leq n \leq N-1$ とする. また, 時間変数は省略する.

$$U_c = \begin{pmatrix} u_0 & u_1 & u_2 & \cdots & u_{N-2} & u_{N-1} \\ u_{N-1} & u_0 & u_1 & \cdots & u_{N-3} & u_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ u_2 & u_3 & u_4 & \cdots & u_0 & u_1 \\ u_1 & u_2 & u_3 & \cdots & u_{N-1} & u_0 \end{pmatrix} \quad (59)$$

これはユニタリ行列 U によって対角化され, 固有値 $\lambda_j (0 \leq j \leq N-1)$ は,

$$f(x) = \sum_{k=0}^{n-1} u_k x^k, \quad \omega = \exp \frac{2\pi i}{N} \quad (60)$$

とすると, 固有値は以下のように表される.

$$\lambda_j = f(\omega^j) \quad (61)$$

5.5.2 ${}^t U_c U_c$ の固有値

次に, ${}^t U_c U_c$ の固有値を求める. ${}^t U_c U_c$ は実対称行列であるから, その固有値は実数となる.

U_c を対角化して得られる対角行列を D とすると, \dagger を複素転置共役記号として,

$$U^\dagger U_c U = D \quad (62)$$

より,

$$(U^\dagger U_c U)^\dagger U^\dagger U_c U = D^\dagger D \quad (63)$$

ここで, U_c の要素は実数なので, $U_c^\dagger = {}^t U_c$ に注意すると, $U U^\dagger = E$ を用いて,

$$U^\dagger {}^t U_c U_c U = D^\dagger D \quad (64)$$

$(D^\dagger D)^\dagger = D^\dagger D$ より, 上式右辺は実対角行列. よって, ${}^t U_c U_c$ の固有値を Λ_j とすると, $\Lambda_j = \lambda_j^* \lambda_j$ となる. ここで, $*$ は複素共役記号. また, Λ は前節でシフト演算子を表す記号として用いたが, 本節にはシフト演算子は登場しないので混乱は生じないだろう.

Λ_j を具体的に書くと次のようになる.

$$\begin{aligned}
\Lambda_j &= \lambda_j^* \lambda_j = \left(\sum_{k=0}^{N-1} u_k \omega^{kj} \right)^* \left(\sum_{l=0}^{N-1} u_l \omega^{lj} \right) = \left(\sum_{k=0}^{N-1} u_k \exp -\frac{2\pi i}{N} kj \right) \left(\sum_{l=0}^{N-1} u_l \exp \frac{2\pi i}{N} lj \right) \\
&= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} u_k u_l \exp \frac{2\pi i}{N} j(-k+l) = \frac{1}{2} \left(\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} u_k u_l \{ \exp \frac{2\pi i}{N} j(k-l) + \exp \frac{2\pi i}{N} j(l-k) \} \right) \\
&= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} u_k u_l \cos \frac{2\pi}{N} j(k-l)
\end{aligned} \tag{65}$$

5.5.3 ${}^t U_c U_c$ の固有値分布

定理 5.9 ルール 170 の固有値分布 $\rho(\Lambda)$ は

$$\rho(\Lambda) = 4e^{-4\Lambda}$$

証明 5.9 $\delta(x)$ をディラックのデルタ関数とする. U_c を 1 つ固定した (= U_c の初期値 $\{u_0, \dots, u_{N-1}\}$ を固定した) 元で, その条件付き固有値分布 $\rho(\Lambda|u_0, \dots, u_{N-1})$ は,

$$\rho(\Lambda|u_0, \dots, u_{N-1}) = \frac{1}{N} \sum_{j=0}^{N-1} \delta(\Lambda - \Lambda_j) \tag{66}$$

(66) 式のデルタ関数を積分表示すると,

$$\delta(\Lambda - \Lambda_j) = \frac{1}{2\pi} \int e^{i(\Lambda - \Lambda_j)y} dy \tag{67}$$

よって, (66) 式は,

$$\rho(\Lambda|u_0, \dots, u_{N-1}) = \frac{1}{2\pi N} \int e^{i\Lambda y} \sum_{j=0}^{N-1} e^{-i\Lambda_j y} dy \tag{68}$$

したがって, 次の値を評価しなければならない.

$$\sum_{j=0}^{N-1} e^{-i\Lambda_j y} \tag{69}$$

ここで, これまでの議論は初期値 $\{u_0, \dots, u_{N-1}\}$ を固定した元での条件付き分布に関するものだったことを思い出す. 実際は $\{u_0, \dots, u_{N-1}\}$ の各値 u_n は確率分布 $p(u_n)$ からそれぞれ独立に出力されたものであるから, 次の積分を実行し, $\rho(\Lambda)$ を求める.

$$\rho(\Lambda) = \int \dots \int p(u_0) \dots p(u_{N-1}) du_0 \dots du_{N-1} \rho(\Lambda|u_0, \dots, u_{N-1}) \tag{70}$$

条件付き固有値分布の u_n に依存する部分は (69) 式の部分なので, まずは次の値,

$$\int \dots \int p(u_0) \dots p(u_{N-1}) du_0 \dots du_{N-1} \sum_{j=0}^{N-1} e^{-iy\Lambda_j} \tag{71}$$

を計算する. 指数の肩は $\mathbf{u} = (u_0, \dots, u_{N-1})$ とすると, 行列 $C(j)$ を用いて以下のように書ける.

$$-i {}^t \mathbf{u} C(j) \mathbf{u} \tag{72}$$

行列 $C(j)$ の k, l 成分を $c_{kl}(j)$ とすると,

$$c_{kl}(j) = \cos \frac{2\pi}{N} j(k-l) \quad (73)$$

$C_{kl}(j)$ は次のような実対称テプリッツ行列である.

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_{N-2} & c_{N-1} \\ c_1 & c_0 & \cdots & c_{N-3} & c_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{N-2} & c_{N-3} & \cdots & c_0 & c_1 \\ c_{N-1} & c_{N-2} & \cdots & c_1 & c_0 \end{pmatrix} \quad (74)$$

巡回行列はテプリッツ行列の一種であるが, (実対称) テプリッツ行列だからと言って巡回行列とは限らない点に注意する. (74) が巡回行列なら $c_1 = c_{N-1}$ である. これを確認する.

$$c_1(j) = \cos \frac{2\pi}{N} j = \cos \frac{2\pi}{N} j(N-1) = c_{N-1}(j) \quad (75)$$

同様に行列サイズが奇数 $N = 2m + 1 (m \geq 1)$ の場合, $c_{1+k} = c_{N-k} (1 \leq k \leq m)$ が成り立つ. 又, $N = 2m$ の場合は, c_m はペアを持たないが, これは今は本質的ではなく, 結局 $C(j)$ は巡回行列であることが分かる.

次に, $C(j)$ の固有値の性質を調べよう. $\alpha = 2\pi/N$ として, その上 2 行を具体的に書き出すと,

$$\begin{pmatrix} 1 & \cos j\alpha & \cdots & \cos(N-2)j\alpha & \cos(N-1)j\alpha \\ \cos j\alpha & 1 & \cdots & \cos(N-3)j\alpha & \cos(N-2)j\alpha \end{pmatrix} \quad (76)$$

$N-1$ 列に注目してみる.

$$\cos(N-3)j\alpha = \cos(N-2)j\alpha \cos j\alpha + \sin(N-2)j\alpha \sin j\alpha$$

右辺第 1 項は 1 行の $\cos j\alpha$ 倍を引くことで消去できる. また, $\sin j\alpha \neq 0$ すなわち, $N = 2m + 1$ のときは $j = 0$, $N = 2m$ のときは $j = 0, m$ を除いて, 行基本変形によって,

$$\begin{pmatrix} 1 & \cos j\alpha & \cdots & \cos(N-2)j\alpha & \cos(N-1)j\alpha \\ 0 & \sin j\alpha & \cdots & \sin(N-2)j\alpha & \sin(N-1)j\alpha \end{pmatrix} \quad (77)$$

となる. 3 行目以降はこの 2 行の線形結合で書けるので階数は 2. 従って, 非零の固有値の個数は 2 個. $C(j)$ の固有値を $\lambda'_g (0 \leq g \leq N-1)$ とする.

【 $N = 2m + 1 (j \neq 0)$ 】 のとき

$C(j)$ が巡回行列であることを用いて

$$\lambda'_g = 1 + 2 \sum_{k=1}^m \cos kj\alpha \cos kg\alpha \quad (78)$$

ここで, $j = g$ のとき,

$$\lambda'_j = 1 + 2 \sum_{k=1}^m \cos^2 kj\alpha > 0 \quad (79)$$

次の公式を用いると,

$$\sum_{r=1}^n \cos^2 rx = \frac{n}{2} + \frac{\cos(n+1)x \sin nx}{2 \sin x} \quad (80)$$

$$\sum_{k=1}^m \cos^2 kj\alpha = \frac{2m-1}{4} = \frac{N-2}{4} \quad (81)$$

よって,

$$\lambda'_j = \frac{N}{2} \quad (82)$$

また, $\cos k(N-j)\alpha = \cos kj\alpha$ より,

$$\lambda'_{N-j} = \frac{N}{2} \quad (83)$$

【 $N = 2m + 1(j = 0)$ 】 のとき

全ての要素は 1 になるので, 階数は 1 で非零の固有値は 1 個, その固有値の値は N .

【 $N = 2m(j \neq 0, m)$ 】 のとき

$$\lambda'_g = 1 + 2 \sum_{k=1}^{m-1} \cos kj\alpha \cos kg\alpha + \cos mj\alpha \cos mg\alpha \quad (84)$$

再び $j = g$ を考えると, 上式右辺の第 3 項は $\cos^2 mj\alpha = 1$ より, 結局,

$$\lambda'_j = \frac{N}{2} \quad , \quad \lambda'_{N-j} = \frac{N}{2} \quad (85)$$

【 $N = 2m(j = 0)$ 】 のとき

全ての要素は 1 になるので, 階数は 1 で非零の固有値は 1 個, その固有値の値は N .

【 $N = 2m(j = m)$ 】 のとき

1, -1 が市松模様に配置された行列になり, 階数は 1 で非零の固有値は 1 個, 固有値の値は N .

これらの固有値に対応する固有ベクトルは巡回行列を対角化するユニタリ行列 U を見ればよいので, U を具体的に書いておく.

$$U = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-2} & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(N-2)} & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \omega^{N-2} & \omega^{2(N-2)} & \cdots & & \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & & \omega^{(N-1)^2} \end{pmatrix} \quad (86)$$

$C(j)$ を対角化して得られた行列を $D(j)$ とすれば, $U^\dagger C(j)U = D(j)$ であり, (72) 式は,

$$-i^t \mathbf{u} C(j) \mathbf{u} = -i^t \mathbf{u} U D(j) U^\dagger \mathbf{u} \quad (87)$$

バイナリベクトル \mathbf{u} から, 新しいベクトル \mathbf{b} を次のようにして作る.

$$\mathbf{b} = U^\dagger \mathbf{u} \quad (88)$$

(87) 式右辺は, \mathbf{b} を用いて,

$$-i \mathbf{b}^\dagger D(j) \mathbf{b} \quad (89)$$

$D(j)$ の固有値は先ほどまでの考察により, 殆どが零であることに注意する.

(71) 式の被積分関数は,

$$\sum_{j=0}^{N-1} e^{-iy\Lambda_j} = \exp(-iy\mathbf{b}^\dagger D(0)\mathbf{b}) + \exp(-iy\mathbf{b}^\dagger D(1)\mathbf{b}) + \cdots + \exp(-iy\mathbf{b}^\dagger D(N-1)\mathbf{b}) \quad (90)$$

右辺初項から考える．まず $D(0)$ は第 0 固有値が N でそれ以外は零固有値だから,

$$\exp(-iy\mathbf{b}^\dagger D(0)\mathbf{b}) = \exp(-iyN\mathbf{b}^\dagger\mathbf{b}) = \exp(-iyNb_0^*b_0) \quad (91)$$

ただし, b_0 は \mathbf{b} の第 0 成分である． $b_0 = \sum_{k=0}^{N-1} u_k/\sqrt{N}$ より, $b_0^* = b_0$

次に, $D(1)$ は第 1, $N-1$ 固有値が $N/2$ でそれ以外は零固有値だから, 同様にして,

$$\exp(-iy\mathbf{b}^\dagger D(1)\mathbf{b}) = \exp(-iy\frac{N}{2}(b_1^*b_1 + b_{N-1}^*b_{N-1})) \quad (92)$$

ここで, $b_1 = \sum_{k=0}^{N-1} u_k\omega^{-k}/\sqrt{N} = b_{N-1}^*$, $b_1^* = \sum_{k=0}^{N-1} u_k\omega^k/\sqrt{N} = b_{N-1}$ より, (92) 式は,

$$\exp(-iy\mathbf{b}^\dagger D(1)\mathbf{b}) = \exp(-iyNb_1^*b_1) \quad (93)$$

他の成分も同様に $b_k^*b_k = b_{N-k}^*b_{N-k}$ となって, 結局 (90) 式は,

$$\sum_{j=0}^{N-1} e^{-iy\Lambda_j} = \exp(-iyNb_0^2) + \exp(-iyNb_1^*b_1) + \cdots + \exp(-iyNb_{N-1}^*b_{N-1}) \quad (94)$$

CA の解析では周期パターン U_c をその周期の平方根で割る．今の場合は $T = N$ なので, $u_k \rightarrow u_k/\sqrt{N}$ とすると, (94) 式は,

$$\sum_{j=0}^{N-1} e^{-iy\Lambda_j} = \exp(-iyb_0^2) + \exp(-iyb_1^*b_1) + \cdots + \exp(-iyb_{N-1}^*b_{N-1}) \quad (95)$$

次に, バイナリ値の確率変数 $u_k \in \{0, 1\}$ から作った b_k の性質を調べる．

$$b_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} u_r\omega^{-kr} = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} u_r \cos \frac{2\pi rk}{N} - \frac{i}{\sqrt{N}} \sum_{r=0}^{N-1} u_r \sin \frac{2\pi rk}{N} \quad (96)$$

(96) 式右辺の実部を s_k , 虚部を t_k として,

$$b_k = s_k + it_k \quad (0 \leq k \leq N-1), t_0 = 0 \quad (97)$$

ここで, s_k, t_k の従う分布を求める． N が非常に大きいとき, 中心極限定理より, s_k, t_k がガウス分布に従うことは明らかである．よって, その平均と分散が求まれば良い．初めに $k \neq 0$ のときの s_k を考察の対象とする．ただし, 以降の $\langle \cdots \rangle$ は \cdots の平均を表すとする．まず, $\langle s_k \rangle$ を求める．

$$\langle s_k \rangle = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} \cos \frac{2\pi rk}{N} \langle u_r \rangle = \frac{1}{\sqrt{2N}} \sum_{r=0}^{N-1} \cos \frac{2\pi rk}{N} = 0 \quad (98)$$

次に, 分散 σ^2 を求める． $\sigma^2 = \langle s_k^2 \rangle - \langle s_k \rangle^2$ ．ここで, (98) 式の結果より, $\sigma^2 = \langle s_k^2 \rangle$ を求める．

$$\langle s_k^2 \rangle = \frac{1}{N} \sum_{q=0}^{N-1} \sum_{r=0}^{N-1} \cos \frac{2\pi qk}{N} \cos \frac{2\pi rk}{N} \langle u_q u_r \rangle \quad (99)$$

$q \neq r$ のとき $\langle u_q u_r \rangle = \langle u_q \rangle \langle u_r \rangle = 1/4$, $q = r$ のとき $\langle u_q u_r \rangle = \langle u_r^2 \rangle = \langle u_r \rangle = 1/2$ に注意すると (99) 式は,

$$\langle s_k^2 \rangle = \frac{1}{2N} \sum_{r=0}^{N-1} \frac{1 + \cos \frac{4\pi r}{N}}{2} + \frac{1}{4N} \sum_{q=0}^{N-1} \sum_{r=0(r \neq q)}^{N-1} \cos \frac{2\pi q}{N} \cos \frac{2\pi r}{N} \quad (100)$$

(100) 式の第 1 項の値は $1/4$ である. 第 2 項は,

$$\frac{1}{4N} \sum_{q=0}^{N-1} \sum_{r=0}^{N-1} \cos \frac{2\pi q}{N} \cos \frac{2\pi r}{N} - \frac{1}{4N} \sum_{r=0}^{N-1} \frac{1 + \cos 2\pi r}{2} = -\frac{1}{8} \quad (101)$$

従って, $s_k (k \neq 0)$ の分散は,

$$\sigma^2 = \frac{1}{4} - \frac{1}{8} = \frac{1}{8} \quad (102)$$

となる. 虚部 t_k に関しても同様の計算から, 平均 0 分散 $1/8$ のガウス分布となることが分かる. また, $q \neq r$ のとき, $\langle s_q s_r \rangle = \langle s_q \rangle \langle s_r \rangle$ より, s_q, s_r は独立. 同様に s_q, t_r, t_q, t_r も独立.

以上から $k \neq 0$ で実部 s_k と虚部 t_k はそれぞれ独立なガウス分布 ($\sigma^2 = 1/8$ より $a = 4$)

$$P(s_k) = \sqrt{\frac{a}{\pi}} \exp(-as_k^2), \quad P(t_k) = \sqrt{\frac{a}{\pi}} \exp(-at_k^2) \quad (103)$$

に従う. $b_0 = s_0$ については, これまで特に触れていないが, $k \neq 0$ のときと同様にの計算により $P(s_0)$ が平均 $\sqrt{N}/2$ 分散 $1/4$ のガウス分布に従っていることが分かる.

以上を踏まえ, (71) 式のバイナリ変数 u_k による積分を b_k によるガウス積分で置き換える. (95) 式を参照し, まず, 次の値を求めてみる. ただし, $b_1 = s_1 + it_1$ である.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-iyb_1^* b_1} P(s_1) P(t_1) ds_1 dt_1 = \frac{a}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-iy(s_1^2 + t_1^2)} e^{-a(s_1^2 + t_1^2)} ds_1 dt_1 \quad (104)$$

上式右辺を極座標表示して, r で積分することで, 下式右辺を得る.

$$\frac{a}{\pi} \int_0^{\infty} \int_0^{2\pi} e^{-iyr^2} e^{-ar^2} r dr d\theta = a \left(\frac{a}{a^2 + y^2} - i \frac{y}{a^2 + y^2} \right) \quad (105)$$

ここで, s_0 に関する次の積分を Ψ とおく.

$$\Psi = \int_{-\infty}^{\infty} e^{-iyNs_0^2} P(s_0) ds_0 \quad (106)$$

s_0 を除く $N - 1$ 個の変数の積分を (105) 式と同様に実行すると, 結局, 固有値分布 $\rho(\Lambda)$ は,

$$\rho(\Lambda) = \frac{1}{2\pi N} \int_{-\infty}^{\infty} \Psi dy + \frac{N-1}{2\pi N} \int_{-\infty}^{\infty} e^{iy\Lambda} a \left(\frac{a}{a^2 + y^2} - i \frac{y}{a^2 + y^2} \right) dy \quad (107)$$

(107) 式の第 1 項の積分の実部は $\mathcal{O}(1)$ なので $N \rightarrow \infty$ で第 1 項は無視できる.

次に (107) 式の第 2 項を計算する. $\rho(\Lambda)$ は実数なので, 虚部を無視すると,

$$\rho(\Lambda) = \frac{a(N-1)}{2\pi N} \int_{-\infty}^{\infty} \left\{ \frac{a \cos y\Lambda}{a^2 + y^2} + \frac{y \sin y\Lambda}{a^2 + y^2} \right\} dy$$

$N \rightarrow \infty$ のとき上式は,

$$\rho(\Lambda) = \frac{a}{2\pi} \int_{-\infty}^{\infty} \left\{ \frac{a \cos y\Lambda}{a^2 + y^2} + \frac{y \sin y\Lambda}{a^2 + y^2} \right\} dy \quad (108)$$

ここで, 次の公式を用いる.

$$\int_0^{\infty} \frac{\cos \theta x}{\beta^2 + x^2} dx = \frac{\pi}{2\beta} e^{-|\theta|\beta} \quad (\beta > 0) \quad (109)$$

(109) 式の指数の肩において, θ に対応する (108) 式の文字は Λ である. ここで, 今回扱う ${}^tU_c U_c$ の固有値は非負であるから, (109) 式の絶対値が外れて, (108) 式右辺第 1 項は,

$$\int_{-\infty}^{\infty} \frac{a \cos y\Lambda}{a^2 + y^2} dy = 2a \int_0^{\infty} \frac{\cos y\Lambda}{a^2 + y^2} dy = 2a \times \frac{\pi}{2a} e^{-a\Lambda} = \pi e^{-a\Lambda} \quad (110)$$

第 2 項に対しては次の公式を用いる.

$$\int_0^{\infty} \frac{x \sin \theta x}{\beta^2 + x^2} dx = \frac{\pi}{2} e^{-\theta\beta} \quad (\theta, \beta > 0) \quad (111)$$

これによって, (108) 式右辺第 2 項は,

$$\int_{-\infty}^{\infty} \frac{y \sin y\Lambda}{a^2 + y^2} dy = 2 \int_0^{\infty} \frac{y \sin y\Lambda}{a^2 + y^2} dy = 2 \times \frac{\pi}{2} e^{-a\Lambda} = \pi e^{-a\Lambda} \quad (112)$$

(110) 式, (112) 式を (108) 式に代入すると,

$$\rho(\Lambda) = \frac{a}{2\pi} \times 2\pi e^{-a\Lambda} = a e^{-a\Lambda} \quad (113)$$

(102) から求めた $a = 4$ を上式に代入して, $\rho(\Lambda) = 4e^{-4\Lambda}$ ルール 170 の周期パターンの固有値分布が求まった. ■

図 15 の実線定理の固有値分布である. 点は, ルール 170 で $N = 1001, M = 500$ の結果である. 右図は左図を対数プロットしたものである.

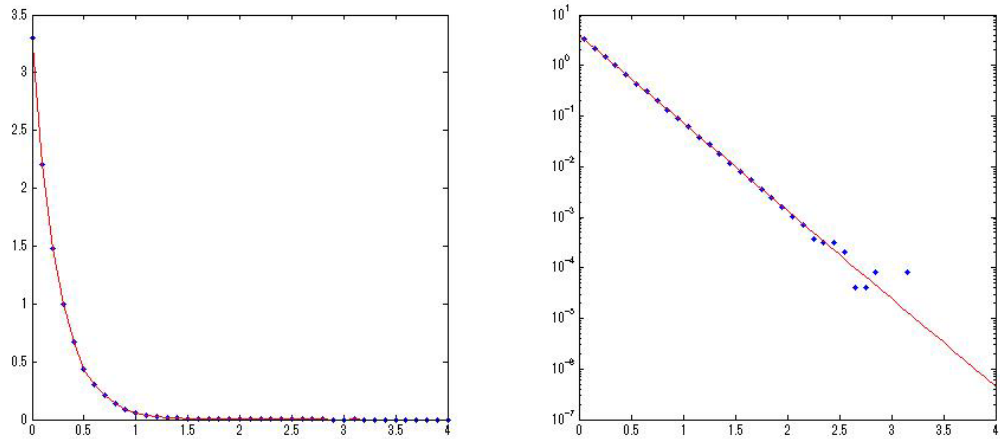


図 15 ルール 170 の周期パターンの固有値分布

5.6 興味深い ECA5 : ルール 73 の特異値間隔分布

ルール 73 では domain wall に挟まれた領域に於いて比較的複雑な時間発展が生じる．この性質より U_r はサンプリングできたが，Appendix-A の図から明らかな通り，ルール 73 の domain wall はセルの値が変化しない直線的なものである．したがって，セルの値が 0 の domain wall の空間変数を $n = n'$ とすると $\forall j \ ({}^tU_r U_r)_{jn'} = 0$ よりランク落ちが生じ，特異値の個数はかなり少ない．実際に $N = 200, M = 500$ として行った計算では，階数の平均値が 46.5，分散 207，最大値 94，最小値 16 であった．ここから更に半分の特異値を捨てるので，実質的に間隔分布の計算に用いることができる特異値の個数は 1 サンプル当たり 20 個程度となる．特に，特異値の個数が少ないと，アンフォールディングが難しい．このような理由から，ルール 73 に対しては信頼できる数値計算が現時点ではできていない．尚，ルール 73 の ${}^tU_r U_r$ は Appendix-A4 の左下の図を参照されたい．

5.7 ECA の特異値間隔分布のまとめ

相対周期パターンの特異値間隔分布を用いて，ウルフラムによる ECA の分類の定量化を試み，数値計算を行った範囲で以下の結果を得た．

クラス 2 相対周期がシステムサイズ以上の U_r がサンプリングできない． U_r に代えて U_c を用いると，極端に周期が短いもの等を除けばポアソン分布が観察された．ただし，ルール 73 では domain wall と呼ばれる時間変化しないセルに挟まれた領域に於いて比較的複雑な時間発展が生じるため， U_r はサンプリングできたが他のクラス 2 と同様に特異値の個数は非常に少ない．

クラス 3 特異値間隔分布は擬ウィグナー分布になる．しかし，ルール 18,146 ではシステムサイズが偶数のとき擬ウィグナー分布とは異なる分布が観察された．この分布の特徴は，擬ウィグナー分布が $P_N(0) = 0$ であるのに対して，特異値の縮退により $P(0) \simeq 0.5$ となっている点である．また，ルール 54 に関しては擬ウィグナー分布が観察された．

クラス 4 擬ウィグナー分布と異なり，かつ，ルール 18,146 でシステムサイズが偶数のときに見られた分布とも異なる間隔分布が観察された．原点付近の値は $P(0) \simeq 0$ である．

以上より，相対周期パターンの特異値間隔分布は分類の指標として有効と思われる．重要とされるクラス 4 の探索に於いては，まず，原点付近で $P(0) \simeq 0$ であることを確認し，次に統計的検定で擬ウィグナー分布ではないものを求めれば良い．尚，本研究によりウルフラムがクラス 4 の特徴として挙げた，effectively have very long transients の情報を除いても，クラス 3 とクラス 4 の違いを定量化できる事が分かった．

経験則の範囲で $P(s)$ の原点付近の挙動は ${}^tU U$ の持つ規則性に強く依存することが分かった．

ルール 170 (シフト写像) の周期パターン U_c (左右反転で巡回行列) の ${}^tU_c U_c$ はテーパーリッツ行列となり，その場合はポアソン分布が現れる． $P(0) = 1$ である．

ルール 18 の相対周期パターン U_r はそれ自体に一見規則性は無い様に見えるが ${}^tU_r U_r$ の要素には 0 がチェッカーボード状に並ぶ． $P(0) \simeq 0.5$ である．

ルール 90 の ${}^tU_c U_c$ は $P(0) \simeq 0.5$ である．ただし，この規則性は上の 2 つと異なり，直接 ${}^tU_c U_c$ を見て確認できる程ははっきりしたものではない．

6 ECA 以外の CA に於けるクラス 4 の探索

本章では 5 章で得た特異値間隔分布を用いた ECA の分類基準を ECA 以外の CA に対して適用する．具体的には，ECA では非常に稀なクラス 4 のルールの探索を行う．

6.1 クラス 4 と 5 近傍 CA

1 章に記した通り，ウルフラムの分類は単にパターン（ルール）の分類だけを目的としたものではなく，興味深い CA を発見する為の前段階という側面を持っている．その際に重要視されるのがクラス 4 に属すルールであり，ウルフラムは，その特徴とされる localized structure が見出されるルールとして ECA のルール 110 に着目した．

ECA 以外の CA に対して，クラス 4 に該当するルールを探索する研究も行なわれている． u_n^t の 5 近傍のセルの総和 $u_{n-2}^t + u_{n-1}^t + u_n^t + u_{n+1}^t + u_{n+2}^t$ の取り得る値 5 ～ 0 に対して，各々 0, 1 を対応着け次の時刻の値とする 5 近傍総和型 CA を対象にした数値的な研究では，コード 20 と 52 がクラス 4 と推測されている [4]（5 近傍総和型 CA ではルールではなくコードと呼ぶ慣習がある）．コード 20 での対応着けは $(5, 4, 3, 2, 1, 0) \rightarrow (0, 1, 0, 1, 0, 0)$ である．コード番号の 20 は 010100 を 2 進数とみなし，それを 10 進数表示したものである．尚，[4] で研究対象となったのは，全 64 種類のコードのうち，総和 0 に対して 0 を対応させる 32 種類のルールである．

ただし，この推測は ECA の場合と同様に恣意的であり，ルール 110 のように特別な性質が証明されている訳でもない．コード 20, 52 に対して本稿の分類手法を適用しようとしたところ，相対周期が非常に短く U_r がサンプリングできなかった．この原因の 1 つは，コード 20 や 52 で生じる localized structure 同士が衝突すると，消滅したり時間変化しない構造に変化したりして，最終的に得られる U_r が単純なものになることにある．

補足として，ルール 110 にも衝突によって他の localized structure を消滅させるものは存在し，rejector と呼ばれる [3]．しかし，Appendix-A3 右下図に見られるようなフェイズシフトは起きても消滅は起こらない衝突も存在する．

6.2 クラス 4 と箱玉系

ウィグナー分布でもなく分布 A（原点付近の値が 0 ではない）でもない，クラス 4 の特異値間隔分布を観察できる CA を探す試みは，5 近傍総和型 CA の場合にはうまくいかなかった．ここで，研究者によってはウルフラムの言う localized structure を（擬）ソリトン [21] と呼ぶことがあることに注意する．ソリトンとは元々は古典可積分系に見られる粒子性を持つ孤立波のことである．

ソリトンの挙動を示す CA は，ソリトンセルオートマトンと呼ばれ，[24] を初めとして様々なものが考案されている．但し，それらの多くは特定の初期値でソリトンのパターンが見られるに過ぎない．これに対して箱玉系（Box-Ball-System）と呼ばれる CA では任意の初期値でソリトンの時間発展パターンが観測でき [25]，かつ，衝突によるソリトンの消滅は起こらない．さらに，時間発展は可逆である為，数値計算に適している．以上が，箱玉系に着目して，クラス 4 の特徴が見られるかを調べる主な動機である．境界条件はこれまで同様，周期境界条件を用いる．周期境界条件を課した箱玉系を周期箱玉系（pBBS）という．時間発展則は次章に記す．

6.3 pBBS に対する数値計算方法

定義 6.1 pBBS の時間発展則は次の通りである [26].

- (i) 時刻 t に於ける状態に存在する 10 対に注目し, その 1 と 0 を線で結ぶ.
- (ii) 結ばれた 10 対を無視して現れる 10 対を線で結ぶ.
- (iii) 以上の操作を全ての 1 が 0 と結ばれるまで続ける.
- (iv) 全ての結ばれた 1 と 0 を入れ換えてできる状態を時刻 $t+1$ の状態とする.

この時間発展則は, ECA (3 近傍) や 5 近傍 CA とは異なり, 次の時刻のセルの値が近傍のセルだけでは決まらない. したがって, 1.1.2 節に記した微小摂動に対する安定性を pattern of difference から評価する方法は意味のある結果をもたらさないことが予想される. 実際にこの方法を pBBS に用いた結果の例が Appendix-A4 の右下の図である. 次に, pBBS の保存量について述べる.

定義 6.2 pBBS の保存量

時間発展則の説明で 10 対を導入した. これに対し, 初めにできた 10 対の個数を p_1 とする. 次に, それらは無視して作られた 10 対を p_2 , この操作を続けて, 最後の 10 対を p_f とする. $p_1 \sim p_f$ をまとめて (p_1, p_2, \dots, p_f) と書く. (p_1, p_2, \dots, p_f) は時間発展の保存量となる. 保存量は左から順に $p_i (1 \leq i \leq f)$ 個の箱を縦に並べたヤング図によって表されることがある.

箱玉系は可逆であり, 任意の初期状態は時間発展するといずれ初期状態に回帰する. 即ち, 2^N 個の状態からなる全状態空間は周期軌道 U_c により分解される. $\sum_i^f p_i$ は状態に含まれる 1 の総数で, これは周期軌道を通して保存されるので 1 の個数が異なる周期軌道を選ぶと, それらの保存量は異なる. 一般に, 同じ保存量を持つ周期軌道であっても, その周期が等しいとは限らないが, $p_f = 1$ の場合は同じ保存量を持つ軌道の (相対) 周期の長さは 1 種類となる. 本性質は数値計算を効率良く行うのに非常に都合が良い為, 本稿では $p_f = 1$ の保存量のみを用いる.

尚, 性質の違う軌道の特異値間隔を混ぜ合わせて間隔分布を描けば, それが無意味なものになる危険性は十分予測できる. これに対して, 与えられた保存量を持つ状態を作る 10-insertion という方法 [27] を用いて保存量を固定し $N = 140$ で玉の個数を 64 個に統一してサンプリングを行った.

保存量は以下の A, B, C の 3 種類とし, Y_A, Y_B, Y_C は [26] の公式を用いて, 与えられた保存量を持つ状態の総数を求め, 数値計算から得た周期でその総数を割ることで算出した A, B, C に属す軌道の個数の常用対数を取ったものである.

A. (40,11,3,3,3,1,1,1,1) $T' = 190, Y_A = 33.5160$

B. (35,13,3,2,2,2,2,1,1,1,1) $T' = 14630, Y_B = 32.0411$

C. (34,11,6,5,4,2,1,1) $T' = 159600, Y_C = 33.2694$

A, C ではこれまで同様 $M = 500$ として特異値列の端 $1/4$ を捨てている. B に関しては A, C と同様の数値計算を行なったところ次頁の中央のグラフに似た中間的な分布が観測された. しかし, 異なる統計的挙動をする準位間隔たちについて平均を行った事により, この様な分布が観測された可能性があるため, $M = 4000$ で固有値密度が最も高い領域 $52 \leq \mu \leq 57$ に入った特異値の 77405 個の規格化された間隔を使用して分布を描いた. その場合も次頁に示す中間的な分布となった.

また, 特異値間隔分布のシステムサイズ依存性を確認するために, $N = 53, 61, 71, 83$ の 4 つの場合に対して, 各々の玉の密度を固定し相対周期を変化させて特異値間隔分布を調べた.

6.4 実験結果

6.4.1 $N = 140$ の場合

A = 上段, B = 中段, C = 下段. 右図は左図の縦軸を対数プロットしたものである.

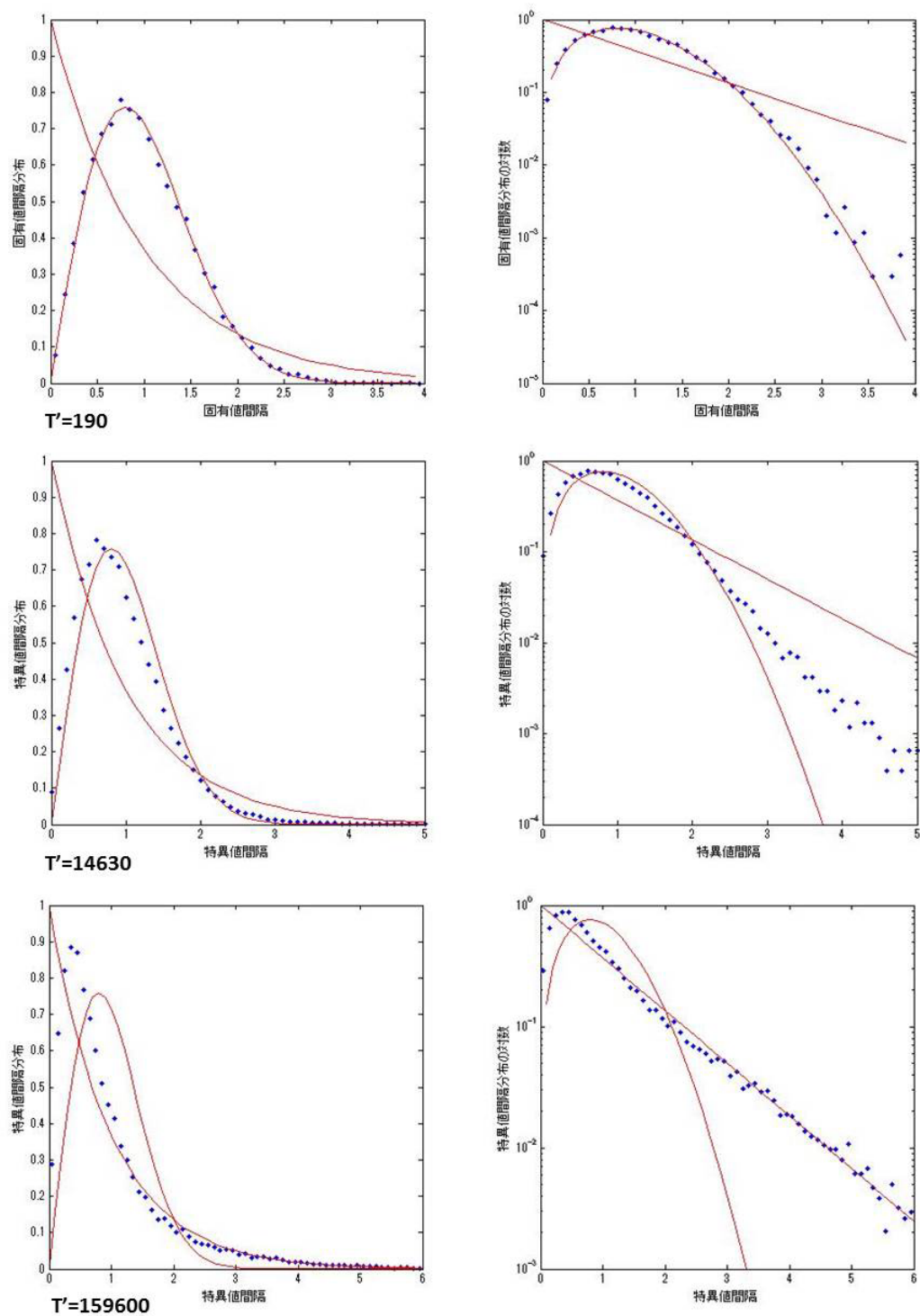


図 16 相対周期(保存量)を変化させたときの pBBS に於ける U_r の特異値間隔分布. $N = 140$, 玉の個数 64 に固定

6.4.2 $N = 53, 61, 71, 83$ の場合

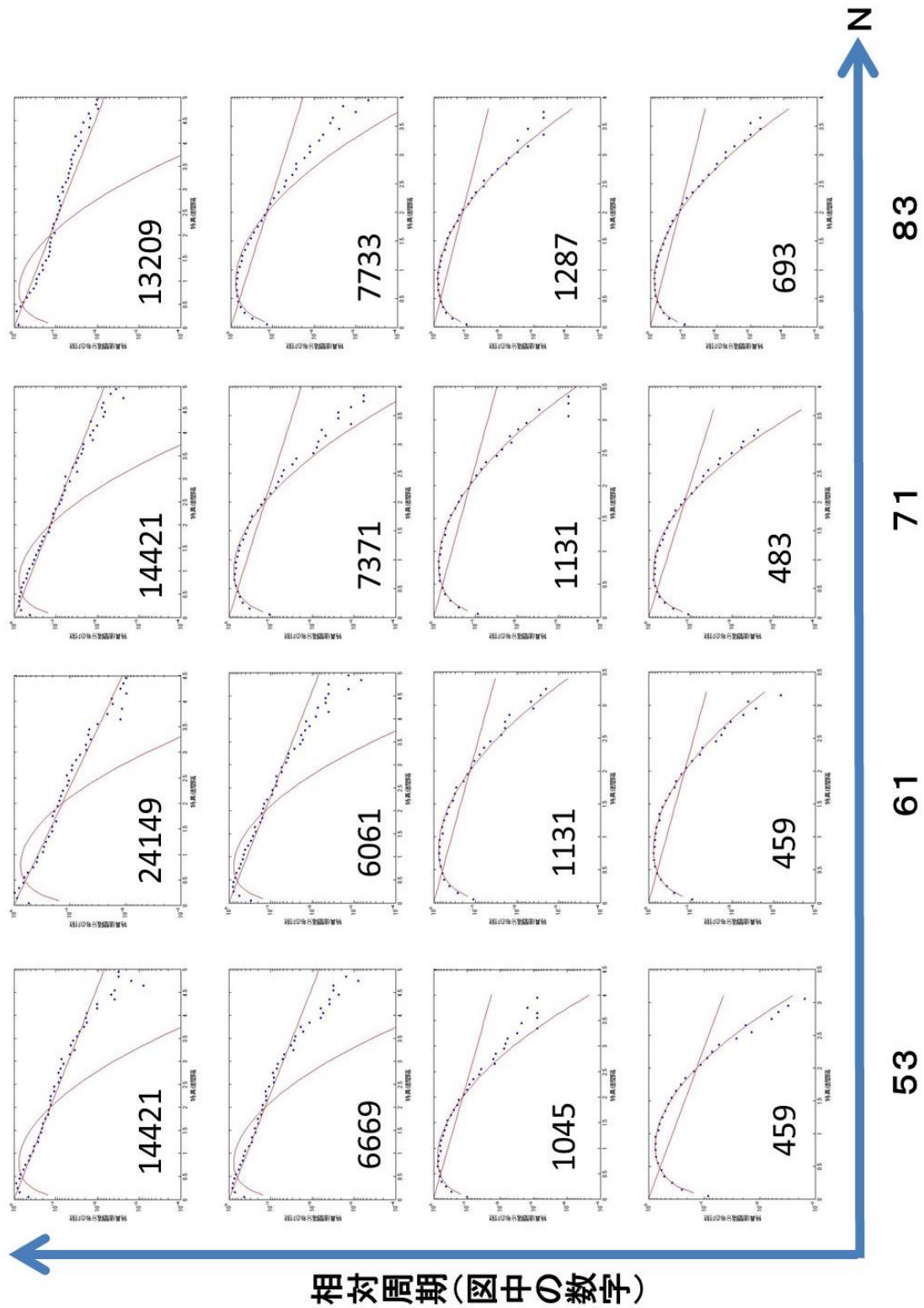


図 17 N 及び T' (保存量) を変化させた場合、各システムサイズで玉の個数を固定

6.5 実験結果のまとめ

数値計算した範囲で相対周期によっては、原点付近で $P(0) \simeq 0$ で、かつ、サイズ 140 の擬ウィグナー分布ではない、クラス 4 の特異値間隔分布が観察された。

また、当初は想定していなかったこととして、相対周期が短い場合にはウィグナー分布に、相対周期が長い場合はポアソン分布によく似た分布が観察された。これに関しては、2 標本 KS 検定を行い、 T' が短いときの分布はサイズ 140 の擬ウィグナー分布である可能性が否定できなかった。

一方、 T' が長いときの分布は、ポアソン分布であるという仮説が棄却された。pBBS から得られたポアソン分布に似た分布では間隔 0 の頻度が急激に 0 に近づいている。一方、 $P_{\text{poisson}}(0) = 1$ であるから、棄却はこの原点付近の挙動の違いによるものと思われる。この違いを考慮し、pBBS から得られたポアソン分布に似た分布を擬ポアソン分布と名付けることにする。擬ポアソン分布の特徴は、裾がポアソン分布と同様に指数関数の 1 次でディケイする事である（ウィグナーは 2 次）。

尚、今回の実験では、手持ちの計算機の性能的な限界の為、相対周期が 20 万以下となる保存量を使用している。これは相対周期としては比較的短い。よって、保存量をランダムに選ぶと殆どの場合、その特異値間隔分布は擬ポアソン分布になると予想している。

加えて、相対周期の変化に伴う特異値間隔分布の変化の様子に着目することで、pBBS は相対周期（或いは相対周期を制御する $p_f = 1$ の保存量）という内部パラメタに関して、擬ウィグナー分布から擬ポアソン分布への遷移を起こす事が分かった。この遷移の過程でクラス 4 の特徴を示す特異値間隔分布が現れる。実は、6.6.2 節で紹介する通りランダム行列理論に於いても同じような遷移が知られている。

6.6 結果に対する考察

6.6.1 クラス 4 の別の側面

7.2 節で pBBS に着目する動機の一つとして、localized structure が衝突で壊れない点を挙げたが、実は別の理由も存在する。ウルフラムが分類を行った後に発表された [3] のインパクトは非常に大きく、CA の分類は、CA の計算能力という側面に傾斜して議論されていた感がある。実際、[4] のあとで出版されたウルフラムの大著 [9] は、この点にかなりのページ数を割いている。

しかし、1 章に記したクラス 4 の定義は CA の計算能力とは独立で、計算能力に偏った議論は、クラス 4 の持つ別の側面を見逃す危険性をはらんでいる。そこで、 U_r がクラス 4 の特徴を持ち、かつ、計算能力以外の点で数学的な価値が明らかな CA として、pBBS に着目したのである。以下に、(p)BBS の価値を補足しておく。

連続系に見られるソリトンの衝突が BBS により再現されるのは偶然ではなく、BBS は連続なソリトン系と超離散化という変換によって結びついていることが知られている。さらに、pBBS はリーマン予想との関連が指摘される等、(物理的な背景を考慮せずとも) それ自体が数学的に興味深い CA である [28]。

これまで、研究者ごとに異なる基準のもとで、クラス 4 に属すと主張されたルールは複数あるが、ルール 110 の計算能力のように具体的にその数学的な特性が示されたものは希少である。この点に関して、pBBS の特異値間隔分布が（保存量によるが）クラス 4 の特徴を示すという事実は、クラス 4 が確かに数学的に重要な CA を含むことの証左と言えよう。

6.6.2 ランダム行列と周期箱玉系に於ける Poisson-Wigner 遷移

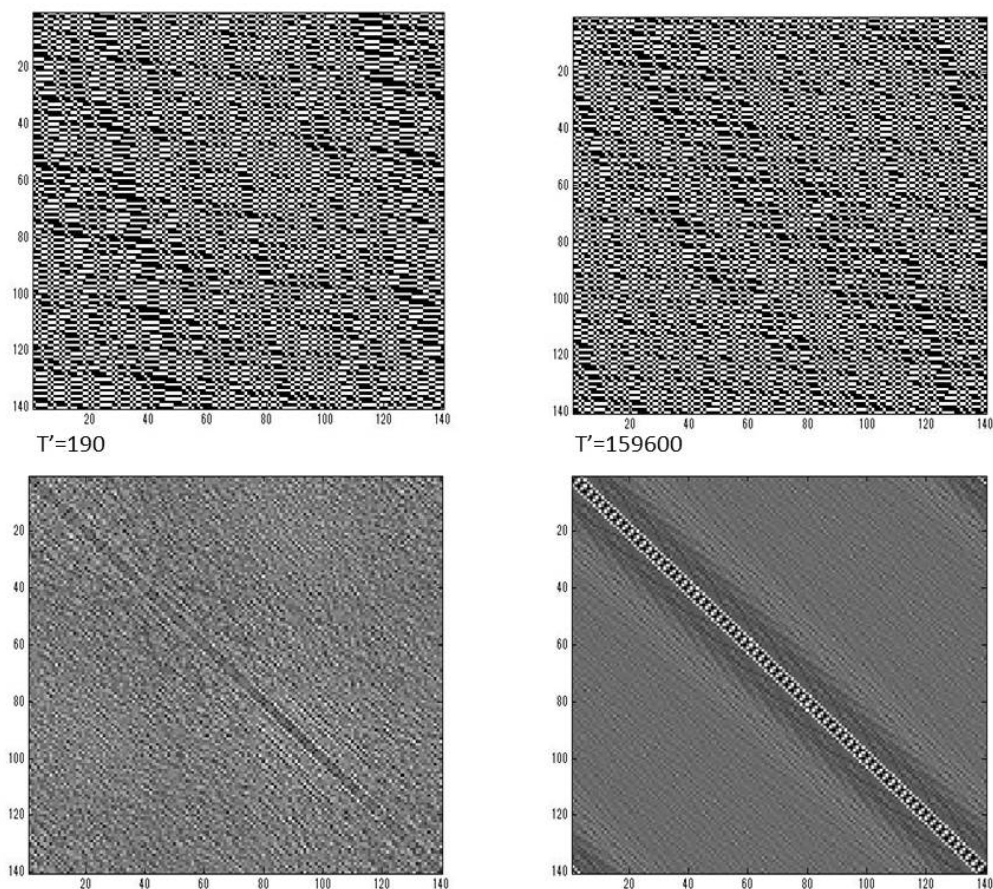


図 18 pBBS の U_r サンプルの一部（上段）と，その ${}^tU_r U_r$ (下段)

pBBS は 5.6 節に述べたクラス 4 の特徴を示す分布だけでなく，擬ウィグナー分布や擬ポアソン分布が観察されることもある．5.7 節の経験則を念頭に置き，まず， ${}^tU_r U_r$ の様子を具体的に観察してみる．

図 18 の上段は pBBS のサンプルパターンの一部で左側が擬ウィグナー分布になるもの ($T' = 190$)，右側が擬ポアソン分布になるもの ($T' = 159600$) である．下段は，それぞれに対応する ${}^tU_r U_r$ をグレースケール表示したものである．ここで，単純にグレースケール表示をすると対角要素の値（自己相関）が他の成分より大きいので，それ以外の要素の差異が見えづらくなってしまう．図 18 の下段ではこの点を考慮し，対角要素の値を行列要素の平均値で置き換えたものをグレースケールで表示している．

複数のソリトンが，周期境界条件のもとで追い越しとフェイズシフトを繰り返すので，サンプルパターン自体は共に複雑な様相を呈している．一方， ${}^tU_r U_r$ は相対周期が長い場合，テーブルリッツ行列に見られるような，斜め方向に走る規則性が見られる．

但し，テーブルリッツ行列では斜め方向の要素が全て同じ値になるのに対して，先に述べた規則性では要素の値が一致するには至っていない．ポアソン分布と擬ポアソン分布の違いである特異値の縮退の有無はこの差が原因なのではないだろうか．

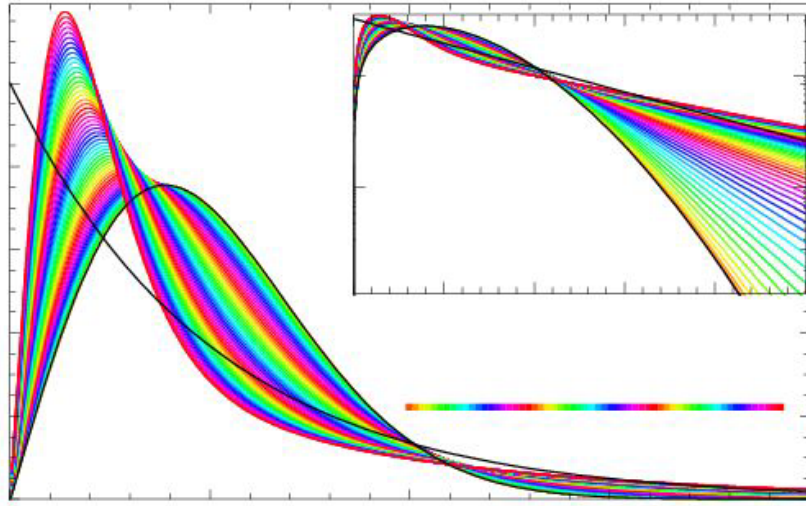


図 19 ランダム行列に於ける Poisson-Wigner 遷移 西垣真祐氏のご厚意による

ここで、擬ウィグナー分布から擬ポアソン分布への遷移について関連すると思われる事項をまとめておく．固有値間隔分布の普遍性は非常に強力であるが，では逆に，ランダム行列ならば，固有値間隔分布は全てウィグナー分布かというそうではない．各行列要素の従う確率分布をガウス分布から 1 パラメタ拡張することで，それを遷移パラメタとして，ウィグナー分布からポアソン分布への遷移が観察できる（図 19 参照；[29] より引用）．ここで，上図の原点付近では pBBS と同様に $P(0) = 0$ を保って分布形が変化している．この遷移を **Poisson-Wigner 遷移** という [30]．

これに対して pBBS の遷移を擬 Poisson-Wigner 遷移と呼ぶことにする．次の 2 点は興味深い．

(I) ランダム行列の遷移と比較すると，pBBS では外部からパラメタを導入する事なく，pBBS に内在した性質から擬 Poisson-Wigner 遷移が生じる．

尚，擬ポアソン分布が観察された pBBS の U_r と同じサイズのバイナリランダム長方形列を使用して数値計算を行ったところ，その特異値間隔分布は擬ウィグナー分布になっていた．但し，間隔分布以外の統計量の類似性については現時点でほぼ不明である． $T' = 190, M = 500$ で作成した pBBS の ${}^tU_r U_r$ の固有値分布を (9) 式で $Q = 190/140$ とした $\rho(\lambda)$ と比較してみたところ，一致は見られなかった．

(II) pBBS の時間発展は決定論的であり，初期状態を除くと確率的な要素の入る余地が無い．

与えられた保存量 α に対し pBBS の取り得る状態の集合を Ω_α とすると， Ω_α は周期軌道により一意に分解される．今回の数値実験では 1 つの保存量に属す軌道の周期は 1 種類なので，初期状態に対する等しい重みのサンプリングは周期軌道を等しい重みでサンプリングする事と等価である．同じ保存量を持つ 10^{30} 個を超える全軌道をサンプリングする事は不可能なため，ランダムサンプリングを行っているが，数値計算の制約を忘れると，原理的にはこのランダムネスを排除できる．その場合，決定論的な系の統計性を研究の対象としていることになる．

その様な研究で，ランダム行列とかかわりが深く，かつ，Poisson-Wigner 遷移が登場する分野に量子カオス論がある [17]．そこでは確率的要素を含まないハミルトニアンエネルギー固有値の間隔分布としてウィグナー分布や Poisson-Wigner 遷移が観察されることが知られている [31]．

尚，本稿の研究対象はあくまでパターンの特異値であり，今のところ ${}^tU_r U_r$ にハミルトニアンといった物理的な意味は見出されていない．

6.6.3 （補足） pBBS と遷移ランダム行列の比較

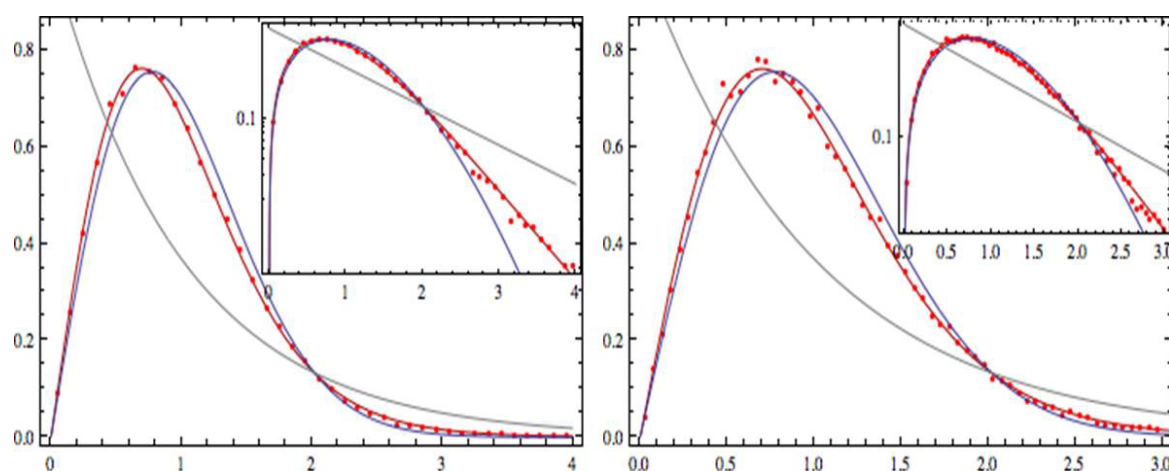


図 20 pBBS の中間的な特異値間隔分布と遷移ランダム行列の比較 西垣真祐氏のご協力による

6.6.2 節では 1 パラメタ拡張されたランダム行列（遷移ランダム行列と呼ぶ）と pBBS に於いてともにポアソン分布とウィグナー分布を結ぶような遷移が観察されることについて考察した。

しかし、遷移ランダム行列と pBBS の直接的な比較を行った訳ではない。

最後に、図 19 の遷移ランダム行列と pBBS から求めた特異値間隔分布を比較したグラフを示す。図 20 の左図と右図はそれぞれ、範囲 $[0, 4]$ ビン 0.1 及び範囲 $[0, 3]$ ビン 0.05 で遷移ランダム行列の解析結果と最小 2 乗フィットした結果である。図の右上には縦軸を対数プロットしたグラフを添えている。

これらのグラフを描くために図 6.4.1 の中間的な分布を描いた特異値間隔に対して、使用する特異値の領域（ウィンドウ）をさらに狭めて $53 \leq \mu \leq 56$ とし、45361 個の間隔を用いている。

いずれもよくフィットできており対応する遷移パラメータ a ([29] を参照されたい) はそれぞれ $a = 1.20(13), 1.14(19)$ である。

この結果は pBBS と遷移ランダム行列の関係を示唆するものである（ポアソン分布とウィグナー分布の間の遷移というだけならば、さまざまな遷移の可能性がある点に注意する）。今後、さらにシステムサイズを大きくした系での数値実験が望まれる。

7 アンフォールドされた特異値間隔の分布の利用

前章までは特異値間隔分布を用いた CA の分類について議論してきた。本研究で用いた方法は、画像等からその特徴を抽出する特異値分解に基礎を置いている。しかし、特徴としては捨てられるべき背景パターンの方を対象としており、特に、特異値や特異値の間隔そのものではなくアンフォールディングされた特異値の間隔という量が、本来の、特異値分解を特徴量抽出に使うという文脈でどのような意味を持つのかは判然としない。

最終章の本章では、この点に関して特異値のカットオフの問題に絡めて意味づけを行う。

加えて、アンフォールドされた特異値の間隔を用いた新しいカットオフ基準を提案する。

特異値分解と累積寄与率基準については 3.1 節を参照されたい。

7.1 累積寄与率基準の欠点

累積寄与率基準の欠点は以下の通りである。特異値を大きいものから順に並べ i 番目の特異値で初めて累積寄与率が 80 % を超えたとすると、 i 番目の特異値まで用いることになるが、 i 番目と $i+1$ 番目の特異値の大きさにはほとんど違いがない（＝特異値の減少が穏やかな領域に i がある）場合、重みがほぼ変わらないところでカットすることに積極的な理由を見出すことができない。

少々極端な例を挙げると、 X がシフトパターンの場合、その特異値には縮退が起こる。 $\mu_i = \mu_{i+1}$ であるのに、 i 番目と $i+1$ 番目の特異値の扱いが異なる、といった事態が生じ得る。

7.2 スクリーンプロット基準

累積寄与率基準の欠点は、重みが緩やかに減少する領域でのカットオフを正当化しにくい事であった。ここから、重みの変化に着目した別の基準を導入することは極めて自然である。

そこでスクリーンプロットを用いた基準を考える。スクリーンプロットは主成分分析でカットオフの位置を決めるために用いられる。ここで、主成分分析とは p 次元での分布を最もよく保存した q 個の合成変数の組を求める方法である。具体的には、 p 次元のデータの分散共分散を対角化し、その固有値（合成変数の分散に対応）の大きい方から q 個の固有値に対応する固有ベクトルを、求める合成変数の組として採用する。 i 番目の固有値 d_i に対応する固有ベクトルを第 i 主成分ベクトルと呼び、 $d_i / \sum_{k=1}^p d_k$ を第 i 主成分の寄与率と呼ぶ。ここで、 q の値をどの様を選ぶかが問題となる。その 1 つの基準として特異値分解の場合と同様に、累積寄与率 S''_q がある。

$$S''_q = \left(\sum_{k=1}^q d_k / \sum_{k=1}^p d_k \right) \times 100 (\%) \quad (114)$$

問題点も同様であり、これに対して別の基準としてスクリーンプロットを用いた基準がある。

定義 7.1 スクリーンプロット

スクリーンプロットとは、単に $d_i \geq d_{i+1}$ とソートした固有値を添え字の順番に並べたものに過ぎない。この減り方の変化が特異値の何らかの質的な変化を表しているとみなし、減り方が変化する点を基準にカットオフをかける方法をスクリーテストと呼ぶ。

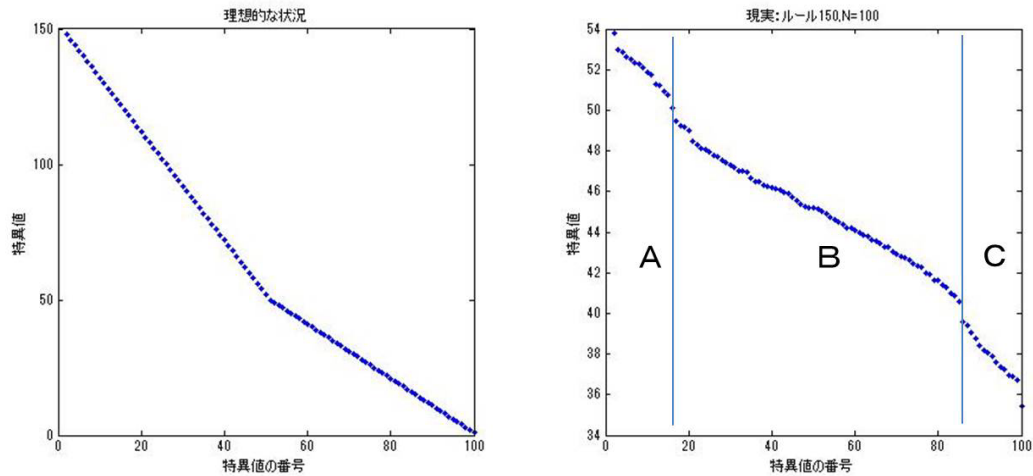


図 21 スクリープロット：理想的な状況（左図）vs 現実的な状況（右図）

このスクリーテストを特異値のカットオフ基準に採用してみる。

具体例を見てみよう。左図はスクリーテストが想定する理想的な状況である。この場合、傾きに変化する $i = 50$ でカットオフをかける。

しかし、右図のルール 150 の出力パターンのスクリープロットが示す通り、現実の問題では傾きに変化する点は複数生じ得る。ここで、特異値が A, B, C の 3 種類のクラスターに分離されていると考え、 A のみを用いる、或いは C を捨てる、という議論になる。しかし、これこそがスクリーテストの欠点であり、本基準は使用する特異値を過少（前者）または過大（後者）に推定する傾向にある。

この傾向は、特異値分布がウィグナーの半円則のようなものの場合、分布端の密度が低いことが特異値の平均間隔が大きい（＝特異値の差が急激に変化する）ことを意味する事から理解される。

解決方法の 1 つは A の使用と C の不使用を確定し、 B に対して再びスクリープロットを試みることである。

クラスター分割の目安になるのが特異値の間隔であるが、先ほどの特異値分布の話から明らかにように、ナイーブには特異値の間隔はどの領域で間隔を計るかに依存する。

例えば、特異値が急激に変化する特異値分布の端を除去したとしても、やはり特異値が密な領域（半円則なら半円の頂点）に近付くほど、間隔は密になり、したがってカットオフしにくくなる。

結局、 B が穏やかに減少して見えるのは、急減少する領域と（大域的に）比較している為であるから、このような特異値の大域的な性質を消し去って、特異値の変化を全ての領域に於いて同等に比較するにはどうしたらよいかを考えたときに登場するのが、特異値間隔の『局所的な』規格化である。これは、アンフォールディングの操作に他ならない。

アンフォールディングには複数の方法があるが、本章では次の移動平均による方法を用いる [19]。

$$(\mu_1 - \mu_{2R}) / (2R - 1) \quad (115)$$

上は特異値 μ_1 と μ_{2R} の間（これを規格化領域と呼ぶ）にある $2R - 1$ 個の間隔の平均値を与えることに注意する。このとき、中央の間隔 $s_{R+1} = \mu_{R+1} - \mu_{R+2}$ を次のように規格化する。

$$(2R - 1)(\mu_{R+1} - \mu_{R+2}) / (\mu_1 - \mu_{2R}) \quad (116)$$

この規格化された特異値間隔を s'_{R+1} と書く。

次に、規格化領域を一つずらし、特異値 μ_2 から μ_{2R+1} に関しても同様に、

$$s'_{R+2} = (2R - 1)(\mu_{R+2} - \mu_{R+3})/(\mu_2 - \mu_{2R+1}) \quad (117)$$

以上を繰り返して、 N 個の特異値に対して $N - 2R - 1$ 個の規格化された間隔を得ることができる。

規格化領域内の特異値の配置が全て等間隔ならば、規格化された間隔は全て 1 になることに注意して、前ページにスクリーンプロットを示した 2 種類の特異値の規格化された間隔を見てみよう。

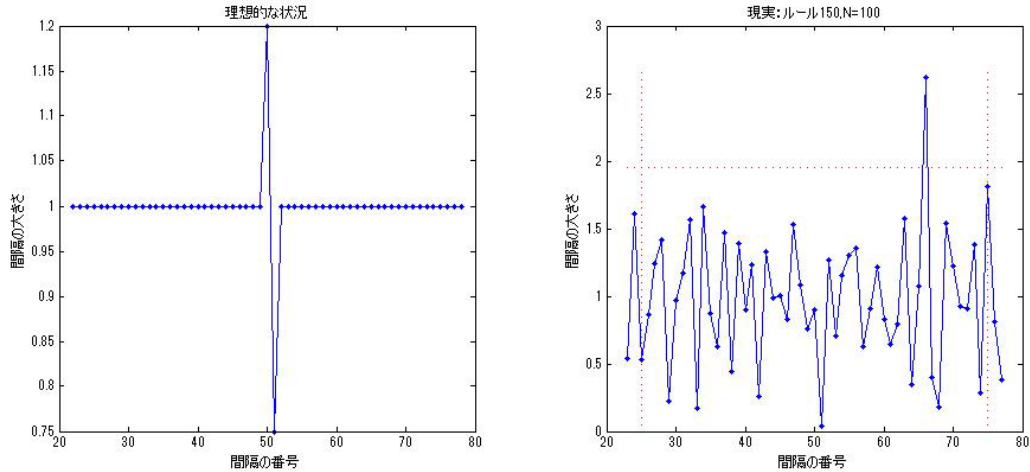


図 22 特異値の規格化された間隔：理想的な状況（左図）vs 現実的な状況（右図）

理想的な例では、確かに $s'_i (i = 50, 51)$ を除き、全ての規格化された間隔が 1 になっていることが分かる。また、飛びの位置は前節のスクリーテストによるカットオフの位置に一致している。

次に、ルール 150 の出力パターンに対して本解析を試みる。ただし、第一特異値は系統的に他の特異値に比べてオーダーが非常に大きいので除外する。 $\mu_2 \sim \mu_{100}$ を $R = 5$ で規格化した特異値間隔を求めると、 $i = 15, 66, 85$ で $s'_i > 2$ の飛びが観測された。 $i = 15, 85$ に関しては前節の右図に於いてカットオフの位置としたところなので、ここでもスクリーテストの基準が適切に検出されている。そして、残った $i = 66$ こそが、求めたかった前節右図の領域 B の局所的な特異値の飛びである。右図は、使用する特異値を $\mu_{21} \sim \mu_{80}$ に絞り（分布端付近の特異値間隔が疎な領域を捨てる）、 $R = 2$ で規格化した特異値間隔分布である。この場合も、 $i = 66$ で飛びが生じ、本節で導入した規格化手法は R の値に対して鋭敏な差異を示さないことが分かる。

局所的な特異値の変化によるスクリーテストをクラスタースクリーテスト (CST) と名付ける。

7.3 CST に於ける『飛び』の定量化

CST は通常のスクリーテストでは相対的に見えなくなってしまう特異値の変化を局所的規格化によって検出し、パターンから得られた特異値列を複数のクラスターに分割することにより、従来のスクリーテストに比べカットオフの柔軟性が増すという利点がある。

しかし、特異値間隔がどの程度であれば『飛び』とみなされるかについて曖昧さが残っている。本節では、これに対して以下に一応の目安を与える。

本手法の目的はパターンの特徴量の抽出であるから、抽出すべき情報を持たないようなパターンでは滅多に出現しない特異値間隔を『飛び』の基準に取ることは自然であろう。ランダムパターンは明らかにそのようなパターンの 1 つであると考える。

ランダムパターンを行列とみなせばランダム行列であり、長方ランダム行列の特異値間隔分布はウィグナー分布により極めてよく近似される。

本研究の対象となるのは、従来のスクリーテストでは検出されないような特異値が分布端に比べ穏やかに減少する領域での特異値の変化である。これに対応するのはランダム行列の特異値分布に於いて、分布端から十分離れた、特異値が密な領域である。

そのような領域に於けるアンフォールディングは次の形に書き表すことができる。

$$s = \mu'_i - \mu'_{i+1} = \rho(\mu_i)(\mu_i - \mu_{i+1}) \quad (118)$$

但し $\rho(\mu)$ は特異値密度関数である。上式では、対象とする領域の特異値密度がほぼ一定で $\rho(\mu_i) \simeq \rho(\mu_{i+1})$ となる事と、特異値密度関数が局所的な平均間隔（平均準位間隔という [32]）の逆数である事を用いている。具体的には、 μ_i, μ_{i+1} を含む微小領域 $\Delta\mu$ 内に含まれる特異値の個数は $\rho(\mu_i)\Delta\mu$ 個であるから、 $\Delta\mu$ での特異値の平均間隔は $\Delta\mu/\rho(\mu_i)\Delta\mu = 1/\rho(\mu_i)$ 。 (115) 式との対応関係は、

$$(\mu_1 - \mu_{2R})/(2R - 1) \longleftrightarrow 1/\rho(\mu_i) \quad (119)$$

である。次に、棄却域に当たるものを設定する。ウィグナー分布の累積分布関数 $F(x)$ は、

$$F(x) = \int_0^x P(s)ds = 1 - \exp(-\frac{\pi}{4}x^2) \quad (120)$$

有意水準を 5 % とすると、棄却域は $F(x) = 0.95$ を満たす x となり、その値は $x \simeq 1.953$ である。前節の右図の水平な点線は、規格化された特異値間隔 $= 1.953$ を示すものである。尚、ウィグナー分布に関する間隔の期待値は 1 である。

しかし、ランダム行列理論では通常、行列アンサンブルの性質を議論する。即ち、非常に沢山の行列の特異値の間隔に関する統計性を扱うのである。一方、この手法では 1 サンプルのパターンに対してランダム行列理論に準拠した基準を用いているため、一見問題があるように思えるが、実はランダム行列には 1 サンプルの統計性がアンサンブルの統計性に一致する自己平均性という性質 [32] が知られており、ウィグナー分布を 1 サンプルに関する分布だと見なせばよい。

本手法は特異値の大域的な性質によらない局所的な変化を検出し、スクリーテストを用いることに比べ柔軟なカットオフを実行できる可能性がある。

7.4 補足

本稿で用いた移動平均によるアンフォールディングは、これまでの計算で用いていた多項式フィッティングによるアンフォールディングとは一見異なる。ランダム行列理論から導かれる間隔分布の理論曲線を数値実験で確認したり、物理的な実験から得られた結果と比較したりするためには、実データをアンフォールドする必要がある。状況に応じて使いやすい方法を選択する。先述のもの以外にも、スプライン曲線を用いる手法などが提案されている [33]。尚、特異値が密な域から少々離れた領域でアンフォールドしたい場合、(118) 式を拡張した下式のアンフォールディングを行う。積分範囲は微小領域 $\Delta\mu$ である。

$$s = \mu'_i - \mu'_{i+1} = (\mu_i - \mu_{i+1}) \int^{\Delta\mu} \rho(\mu)d\mu \quad (121)$$

8 本研究のまとめ

ウルフラムの ECA の分類の問題点である目視の恣意性を除去した定量的な分類を試みた。提案手法では、CA の出力パターンを特異値分解し特異値の大きい部分を削除した残り（背景パターン）を研究対象とする事で目視の恣意性を取り除く。その結果、数値計算を行った範囲でウルフラムの分類に於いてクラス 3（カオス的なパターンを生成）のルールの背景パターンの特異値間隔分布は、ランダム行列の特異値間隔分布の近似式であるウィグナー分布に良い一致を見せる事が分かった。一方、クラス 3 との区別が難しいとされるクラス 4（localized structure を生じる）の特異値間隔分布はクラス 3 の特異値間隔分布とは異なる事が KS 検定により示された。特異値間隔分布は分類の指標として有効と思われる。

尚、CA の時間発展が決定論的であることを考慮すると、ランダム行列との一致は自明ではない。これに関して、クラス 3 のルール 150 は $N = 7 \cdot 2^m$ に於いてランダム選んだ初期値の殆ど全てが $T' = N$ なる最大基本相対周期軌道に属していることを証明した。将来、CA とランダム行列をつなぐ解析的な研究を行う上で役立つことを期待している。

次に、上の分類指標を用いて ECA 以外の CA についてクラス 4 に属すルールの探索を行った。同様の探索が多く研究者によって成されているが、クラス 4 と推定され、かつ、数学的な価値が明確なルールはほぼ皆無であった。これに対して、数学的な価値が明らかで、かつ、localized structure を生じる箱玉系と呼ばれる CA に対して、ECA のクラス 4 に見られた特異値間隔分布の特徴が生じるかを確認した。数値実験から、箱玉系の挙動はその保存量によって異なり、保存量を適当に選ぶとクラス 4 の特徴が観察されることが分かった。

最後に、特異値間隔分布の CA の分類以外への応用について考えた。固有値や特異値の間隔分布は特徴量抽出等のカットオフの基準として利用できるのではないかと考える。

9 今後の課題

クラス 3 のルール 22,30,45,106 に関しては相対周期が非常に長い等の理由で U_r がサンプリングできず解析を断念した。これらのルールに於いてクラス 3 の分布が見られるか否かは確認する必要がある。更に、本研究では主に $N = 200$ 付近のシステムサイズで数値実験を行っており、今後はより多くのシステムサイズで同様の実験を行う予定である。ただし、ウルフラムの分類は元来非常に大まかなもので、当初からすべてのシステムサイズに対してクラスごとに同じ挙動が見られるとの主張が成されている訳ではない。どのような例外が存在して、そのときの特異値間隔分布がどのような分布になるのかを詳細に調べることは、ウルフラムの分類自体の妥当性の検証に繋がるだろう。

また、今回は自明な特異値の減少を避けるという目的でシステムサイズ以上の相対周期パターンをサンプリングしているが、例えば、 $T' = N - 1$ のサンプルが生じた場合に、たった 1 個の自明な特異値の減少を理由にそのサンプルを捨てる事が適切と言えるのか疑問が生じる。これは、そのようなサンプルがどの程度の割合で生じるかを絡めて議論する必要があるだろう。

本研究では、ウルフラムが分類の目的とした CA の計算能力とクラス 4 の関係については議論の対象としなかったが、pBBS がクラス 4 の特徴を示す事から、その計算能力も今後の研究課題として興味深い。pBBS のアルゴリズムとしての側面を議論した先行研究として [34] がある。また、可逆 CA の計算能力に関しては [35] がある。

10 謝辞

時弘哲治先生，WILLOX Ralph 先生，間田潤先生，望月真祐先生に感謝致します。

時弘先生には，約 3 年前に研究室を志望したときから言葉に表せないほどお世話になりました。研究のご指導は勿論のこと，私の分かりにくい説明にも，黙って耳を傾けてくださいました。私が博士課程に進学後に教員免許の取得で大変なときにも，励ましとご配慮を賜りました。本当に言葉にできないことばかりでしたが，敢えて，誤解を恐れず表現するとすれば，時弘先生が示されたものは，真の豊かさであったと思います。3 年間，誠にありがとうございました。

WILLOX 先生には，いつも気さくに質問等にご対応頂きました。セミナーでつまらない結果を報告したときも，前向きなコメントをしてくださいました。

間田先生には，発表の準備や論文作成の際に，いつもの的確で具体的なアドバイスを頂きました。3 年間，御心配ばかりお掛けした気がします。

望月先生には，島根大学を卒業してからもランダム行列等について有益な情報を頂きました。

最後に，261 院生室の皆さんに感謝します。特に，時弘研究室の神吉さん，WILLOX 研究室の間瀬さん，坂井研究室の勝島さんと中村さん，吉田研究室の野村さんにお世話になりました。

付録 A サンプルパターン

A.1 クラス 3 (ルール 18,22,30,45,60,90)

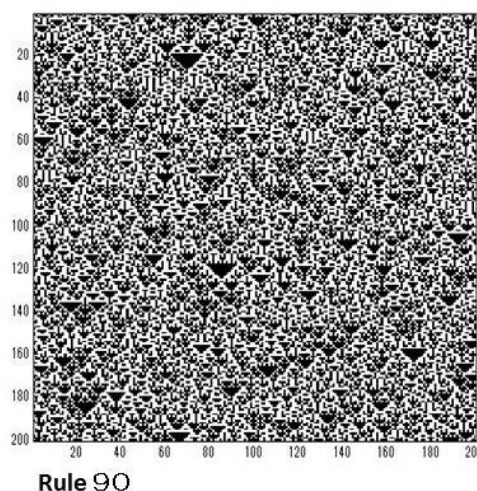
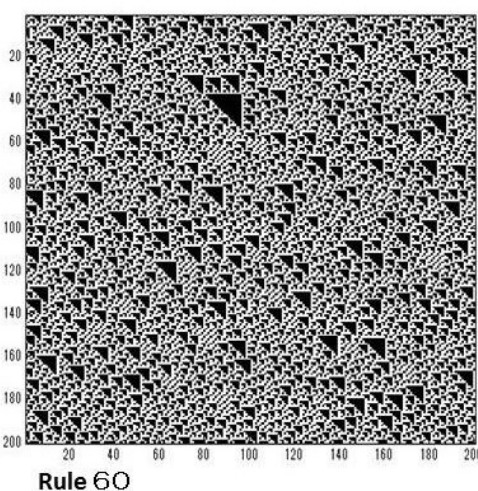
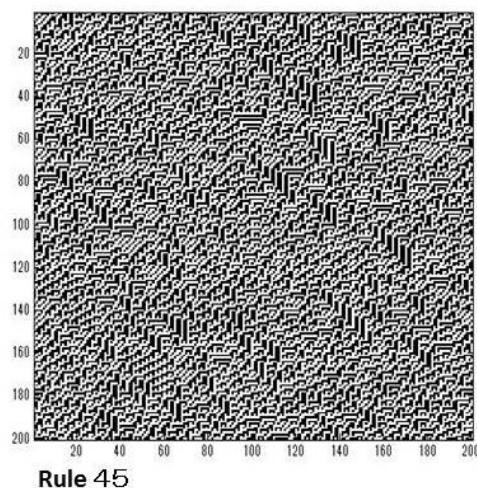
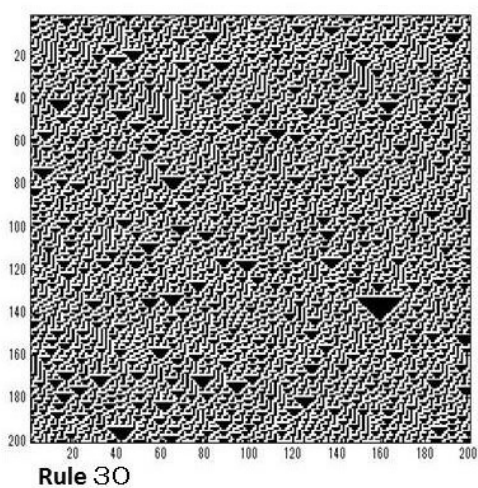
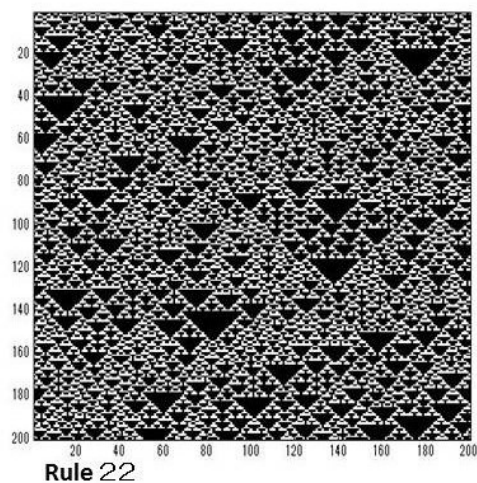
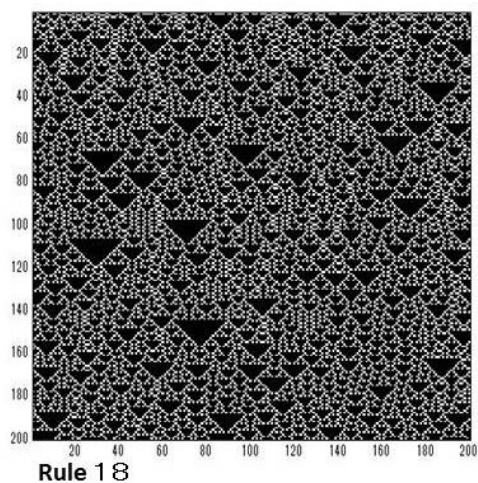


図 23 クラス 3 のサンプルパターン : ルール 18,22,30,45,60,90

A.2 クラス 3 (ルール 105,106,150,122,129,146)

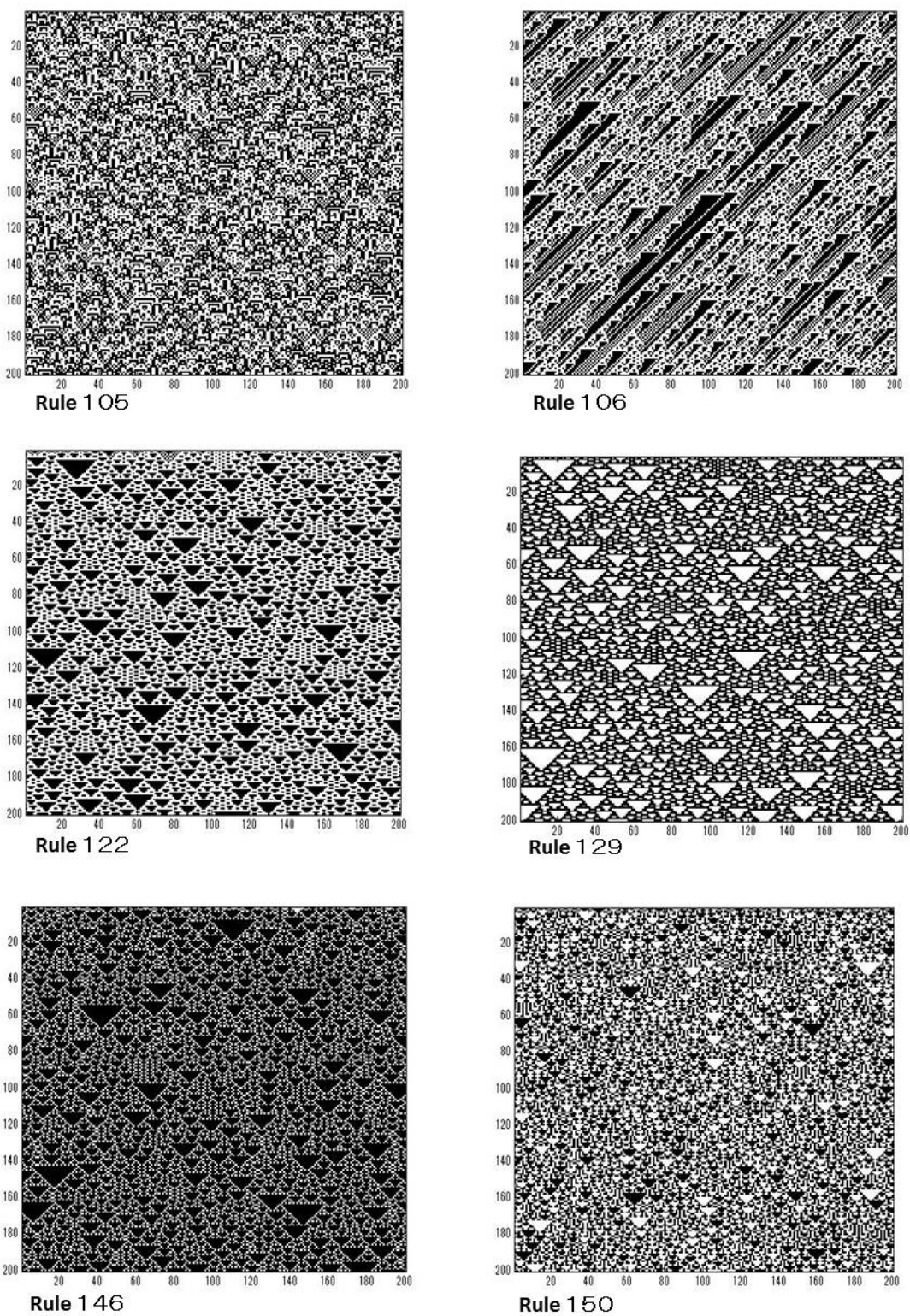


図 24 クラス 3 のサンプルパターン : ルール 105,106,122,129,146,150

A.3 クラス 2 (ルール 154,170,73), クラス 4 ([54],110)

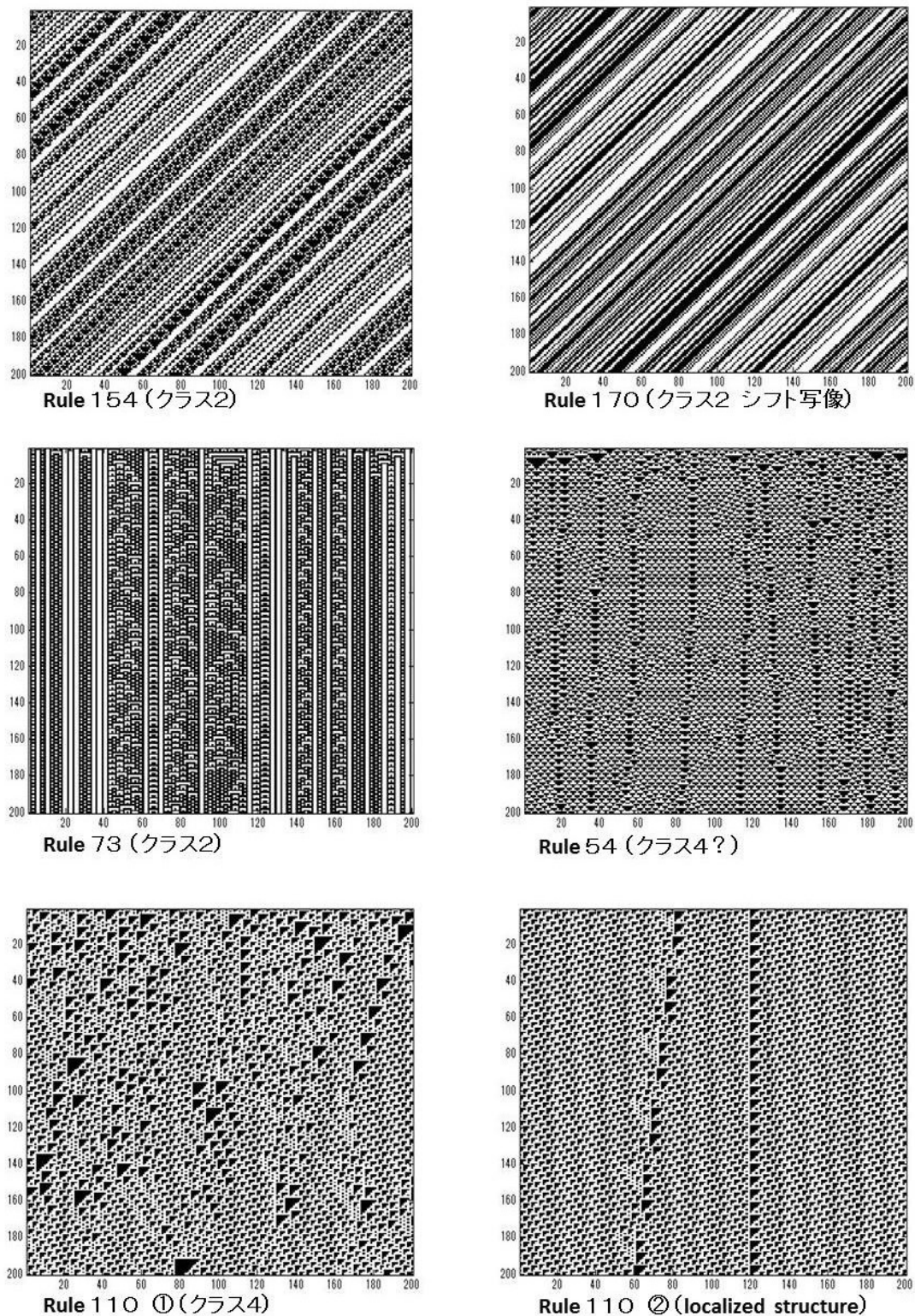


図 25 クラス 2 と 4 サンプルパターン (Rule73,154 は domain wall を生じる)

A.4 その他

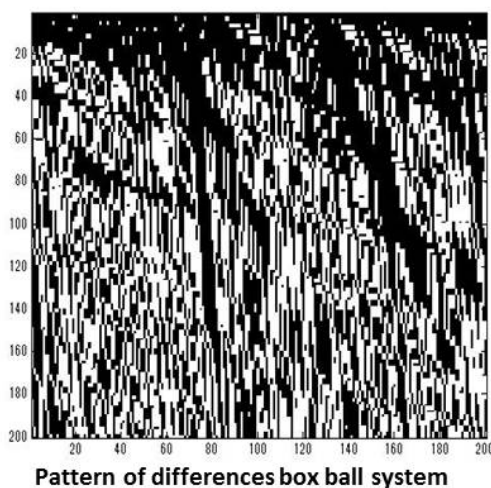
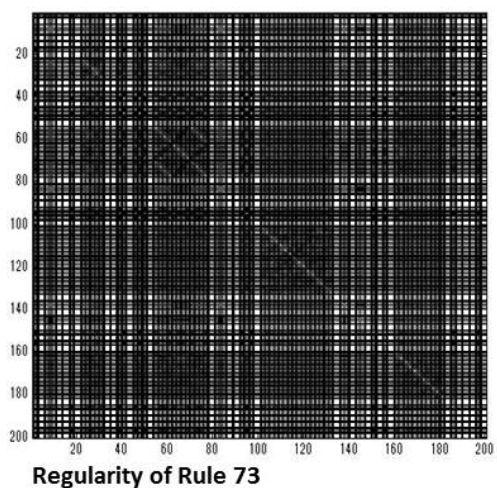
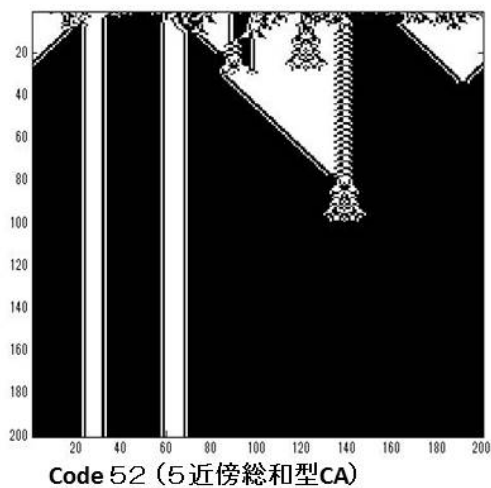
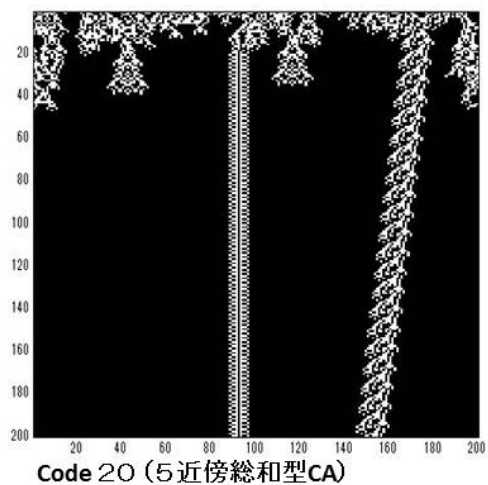
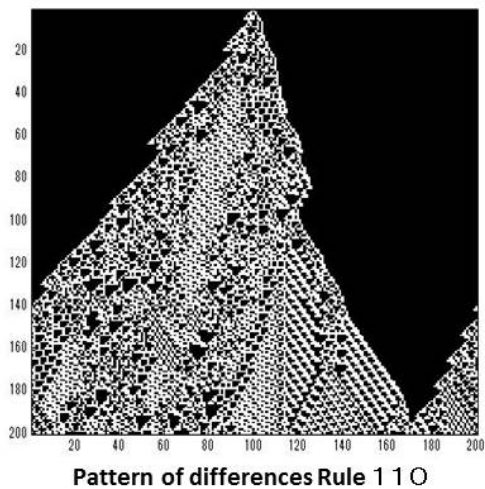
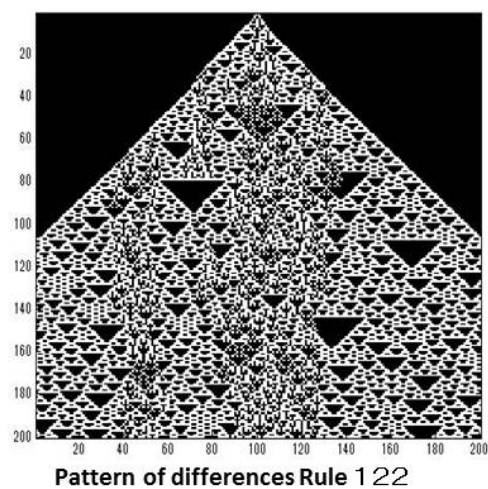


図 26 その他のパターン：Pattern of differences, 5 近傍 CA 等

付録 B ウィグナー分布の導出

ウィグナー分布は行列要素 ψ_{ij} が独立にガウス分布に従うサイズ 2 の対称行列 ψ の固有値間隔分布である。以下に導き方を記しておく [17].

対角成分 ψ_{11}, ψ_{22} の分散は 1, 非対角成分 ψ_{12} の分散が $1/2$, 平均はともに 0 であるとする。

ここで, $\text{tr}\psi = \psi_{11}^2 + 2\psi_{12}^2 + \psi_{22}^2$ より,

$$P(\psi_{11}, \psi_{12}, \psi_{22}) \propto \exp\left(-\frac{1}{2}\text{tr}\psi^2\right) d\psi_{11} d\psi_{12} d\psi_{22} \quad (122)$$

ψ は対称行列なので直交変換によって対角化され, その固有値を λ_1, λ_2 とすると, ψ^2 の固有値は λ_1^2, λ_2^2 なので (122) 式の指数関数の中身は $\text{tr}\psi^2 = \lambda_1^2 + \lambda_2^2$ となる。同様に測度も書き換えたものを $P(\lambda_1, \lambda_2, \theta)$ とする。

$$P(\lambda_1, \lambda_2, \theta) \propto \exp\left(-\frac{1}{2}(\lambda_1^2 + \lambda_2^2)\right) |J| d\lambda_1 d\lambda_2 d\theta \quad (123)$$

ヤコビアン J を求めると, $J = \lambda_2 - \lambda_1$ となるので, (123) 式は

$$P(\lambda_1, \lambda_2, \theta) \propto \exp\left(-\frac{1}{2}(\lambda_1^2 + \lambda_2^2)\right) |\lambda_2 - \lambda_1| d\lambda_1 d\lambda_2 d\theta \quad (124)$$

となる。次に変数変換 $s = \lambda_2 - \lambda_1, t = \lambda_2 + \lambda_1$ を行い,

$$P(s, t, \theta) \propto \exp\left(-\frac{1}{4}(s^2 + t^2)\right) s ds dt d\theta \quad (125)$$

(125) 式で s のスケールを任意に変更し s 以外の変数について積分すると, a, b を定数として,

$$P(s) = as \exp(-bs^2) \quad (126)$$

確率の和が 1 になる条件と, 間隔 s の期待値が 1 になる条件を課せば $a = \pi/2, b = \pi/4$ となり,

$$P(s) = \frac{\pi}{2} s \exp\left(-\frac{\pi}{4}s^2\right) \quad (127)$$

ウィグナー分布を得る。行列サイズが無限大の場合の固有値間隔分布は, ウィグナー分布によって非常によく近似されることが知られている。行列サイズが 2 ではなく一般的な N の場合は

$$P(\lambda_1, \lambda_2, \dots, \lambda_N) = C \left(\prod_{i>j} |\lambda_i - \lambda_j| \right) \exp\left(-\frac{1}{a^2}(\lambda_1^2 + \lambda_2^2 + \dots + \lambda_N^2)\right) \quad (128)$$

について以下の積分を実行し,

$$P(\lambda_1, \lambda_2) = \int \dots \int_{\mathbb{R} \setminus [\lambda_1, \lambda_2]} P(\lambda_1, \lambda_2, \dots, \lambda_N) d\lambda_3 \dots d\lambda_N \quad (129)$$

そこから間隔分布を得ることになる。

ウィグナー分布は 1956 年にテネシーで行われた中性子物理の国際会議においてウィグナーによって推測的に述べられた。その際に, 質問に答える形で (128), (129) 式について言及した。(128) 式は後にガウス型直交集団を特徴づける分布密度と呼ばれ, その導き方が詳しく論ぜられた。又, 困難な積分 (129) も特殊な技巧を用いることで実行され, そこから得られる正確な準位間隔分布とウィグナー分布との比較が定量的になされた [17].

付録 C ランダム行列理論と量子カオス論の歴史的な背景

ランダム行列理論と量子カオス論の歴史的な背景を踏まえながら、Poisson-Wigner 遷移の登場までをごく簡単に概観する。尚、ここに記した内容の殆どは [17] によっている。

C.1 原子核物理とランダム行列

50 年代の原子核物理において、重原子核のハミルトニアン of 不可知性は根本的な問題であった。

但し、量子状態の寿命に関する実験データから、ハミルトニアン of 行列要素の統計性は、ガウス分布でよく表されることは分かっていた。

そこで、各要素が独立にガウス分布に従う行列を作り、その統計性を議論する立場が現れた。尚、独立性の正当性については現在でも曖昧さが残っているといわれている [36]。

ランダムに生成したサイズ N のハミルトニアンを一つ固定したもとで、固有値 $\lambda_i^{(H)} (1 \leq i \leq N)$ の条件付き固有値分布 $\rho(\lambda|H)$ を求め、それを H の出現確率 $P(H)$ で平均することで固有値分布 $\rho(\lambda)$ を求める。 H がランダムエルミート行列のとき、 $N \rightarrow \infty$ の漸近分布は解析的に求めることができる。その分布は半円形になる為、これをウィグナーの半円則という。

ウィグナーの半円則は行列要素の従う確率分布の変化等に対して鋭敏な差異を示さない事が知られている。 $\rho(\lambda)$ より更に強い普遍性を持つものとして固有値間隔分布 $P(s)$ は非常に重要である。 $P(s)$ はウィグナー分布で近似される。

C.2 量子可積分系とポアソン分布

70 年代に入ると、原子核物理とは独立に、古典的に可積分な系を量子化すると、そのエネルギー固有値の間隔の分布はポアソン分布になることがベリー等によって示された [37]。これによって、古典-量子対応が意識されるようになった。即ち、上の対応関係に照らした場合、ウィグナー分布を導く要因は、古典系の非可積分性（カオス）にあるのではないかと推測が成された。古典カオスに対応する量子系のランダムネスを含まないハミルトニアンからウィグナー分布を観測することができるはずである。この主張の正しさは半古典的に証明された [45]。

C.3 量子カオスとビリヤード系

前節の推測を確かめるに当たって、古典系がカオスを示し、かつ、対応する量子系のエネルギー固有値を計算できる系を用意する必要性が生じ、スタジアムビリヤードが着目された。この名前は、長方形の両端に半円を貼りつけたビリヤードの形がスタジアムに似ていることに由来する。長方形部分の長さが 0（このときは可積分）でなければ、このビリヤード内の質点の運動はカオティックである。また、量子系のエネルギー固有値は数値的に求める事ができる。数値実験の結果は、推測を肯定するものであった。その後、ポアソン分布（可積分）とウィグナー分布（カオス）の間は、どの様になっているのかが問題にされるようになり、Poisson-Wigner 遷移の議論へと続いていく。

尚、Poisson-Wigner 遷移を観察できる数理モデルは 1 つではなく、ランダム行列を拡張する事でウィグナー分布からポアソン分布への遷移を観察する方法 [29] や、逆に、可積分な系に非可積分な摂動を加えることでポアソン分布からウィグナー分布への遷移を観察する方法 [31] 等がある。

参考文献

- [1] K.Nishinari, D.Takahashi "Analytical properties of ultradiscrete Burgers equation and rule 184 cellular automaton" J.Phys. **A 31**, 5439-5450 (1998)
- [2] E.R.Berlekamp, J.H.Conway, R.K.Guy "Winning Ways for your Mathematical Plays Vol II. Games in Particular" London: Academic Press (1982)
- [3] M.Cook "Universality in Elementary Cellular Automata." Complex Systems **15**, 1-40 (2004)
- [4] S.Wolfram "Cellular Automata and Complexity" Addison-Wesley, Reading (1994)
- [5] W.Li, N.H.Packard, C.Langton "Transition Phenomena in Cellular Automata Rule Space" Physica D **45**, 77-94 (1990)
- [6] M.Grant, J.D.Gunton "Cellular Automata, Langevin Equations, and Unstable States" Phys. Rev. Lett. vol **57**, 1970-1973 (1986)
- [7] 小山信也 "散乱行列式と数論的量子カオス" (別冊・数理科学 量子カオスの物理と数理) サイエンス社 (2000)
- [8] W.Li, M.G.Nordahl "Transient Behavior of Cellular Automaton Rule 110" SFI WORKING PAPER (1992)
- [9] S.Wolfram "A New Kind of Science" Wolfram Media Inc (2002)
- [10] 竹村彰通 "統計" (統計 共立講座 21 世紀の数学 14) 共立出版 (1997)
- [11] V.Plerou, P.Gopikrishnan, B.Rosenow, L.A.N.Amaral, T.Guhr and H.E.Stanley "Random Matrix Approach to Cross-Correlations in Financial Data" Phys.Rev.E **65**, 066126 (2002)
- [12] 金子勇治 時弘哲治 間田潤 "エレメンタリーセルオートマトンの相互相関行列の解析" 九州大学応用力学研究所 研究集会報告 **24O-S3**, 103-126 (2012)
- [13] 金子勇治 "ランダム行列と周期箱玉系に於ける Poisson-Wigner 遷移" RIMS 講究録別冊 (2013 投稿中)
- [14] S.Takesue "セルオートマトンの保存量" 数理解析研究所講究録 1020 巻 (1997)
- [15] 柳井晴夫 竹内啓 "射影行列 一般逆行列 特異値分解" (UP 応用数学選書 10) 東京大学出版会 (1983)
- [16] 野邊厚 由良文孝 "可逆エレメンタリーセルオートマトンの可積分性について" 数理解析研究所講究録 第 1541 巻, 178-191 (2007)
- [17] 長谷川洋 "量子系の準位統計－量子カオス序論" (大槻義彦編 物理学最前線 28) 共立出版 (1991)
- [18] M.Jimbo, T.Miwa, Y.Mori, M.Sato "Density matrix of impenetrable bose gas and the fifth Painlevé transcendent" Physica D**1**, 80-158 (1980)
- [19] L. Nemes "Statistical studies of level correlations and chaotic phenomena in spectroscopy" Acta Physica Hungarica **73**, 95-117 (1993)
- [20] C.Hammond, S.J.Miller "Distribution of eigenvalues for the ensemble of real symmetric toeplitz matrices" arXiv:math/0312215v1 (2003)

- [21] G.J.Martnez, A.Adamatzky, F.Chen, L.Chua "On Soliton Collision between Localization in Complex Elementary Cellular Automata: Rule 54 and 110 and Beyond" *Complex Systems* **21**, 117-142 (2012)
- [22] 伏見正則 "乱数" (UP 応用数学選書 12) 東京大学出版会 (1989)
- [23] M.L.Mehta "Random Matrices, Volume 142, Third Edition (Pure and Applied Mathematics)" Academic Press (2004)
- [24] K.Park, K.Steiglitz and P.Thurston "Soliton like behavior in automata" *Physica D* **19**, 423-432 (1986)
- [25] D.Takahashi, J.Satsuma "A soliton cellular automaton" *J.Phys.Soc.Jpn.* **59**, 3514 (1991)
- [26] 時弘哲治 "箱玉系の数理" 朝倉書店 (2010)
- [27] J.Mada, T.Tokihiro "Correlation function for a periodic box-ball system" arXiv:0911.3953v1 (2009)
- [28] T.Tokihiro, J.Mada "Asymptotic behavior of fundamental cycle of periodic box-ball systems : A number theoretical aspect" *Glasgow Math.J.* **47A**, 199-204 (2005)
- [29] S.M.Nishigaki, M.Giordano, T.G.Kovács, F.Pittler "Critical statistics at the mobility edge of QCD Dirac spectra" arXiv:1312.3286v1 (2013)
- [30] S.M.Nishigaki "Accuracy and range of validity of the Wigner surmise for mixed symmetry classes in random matrix theory" arXiv:1209.0696v2 (2012)
- [31] H.Ishio, K.Nakamura "Quantum irregular spectra: An alternative interpretation from the viewpoint of nonlinear dynamics" *Phys. Rev.* **A46**, R2193 (1992)
- [32] 永尾太郎 "ランダム行列の基礎" 東京大学出版会 (2005)
- [33] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes, the Art of Scientific Computing*. Cambridge University Press, Cambridge (1989)
- [34] 由良文孝 "ソリトンセルオートマトンと量子コンピューティング" 数理解析研究所講究録 1221 巻, 70-89 (2001)
- [35] 森田憲一 "可逆計算" (ナチュラルコンピューティング・シリーズ 第5巻) (2012)
- [36] 首藤啓 "擬可積分系とレベル統計 (別冊・数理科学 量子カオスの物理と数理)" サイエンス社 (2000)
- [37] M.V.Berry, M.Tabor "Level clustering in the regular spectrum" *Proc. Roy. Soc.*, A356, 375 (1977)
- [38] Y.Kaneko "A property of the rule150 elementary cellular automaton" arXiv:1401.4779v1 (2014)
- [39] N.Fatés "Experimental study of elementary cellular automata dynamics using the density parameter" *Discrete Mathematics and Theoretical Computer Science AB(DMCS)*, 155-166 (2003)
- [40] S.Ninagawa "Power spectral analysis of elementary cellular automata" *Complex Systems* **17**, 399-411 (2008)
- [41] M.Schule, R.Stoop "A full computation-relevant topological dynamics classification of elementary cellular automata" *Chaos* **22**, 043143 (2012)
- [42] 広田良吾 高橋大輔 "差分と超離散" 共立出版株式会社 (2003)

- [43] Shinsuke M. Nishigaki "Universality crossover between chiral random matrix ensembles and twisted $SU(2)$ lattice Dirac spectra" *Phys. Rev. D* **86**, 114505 (2012)
- [44] P.J. Forrester "Painlevé transcendent evaluation of the scaled distribution of the smallest eigenvalue in the Laguerre orthogonal and symplectic ensembles" *arXiv:nlin/0005064v1* (2000)
- [45] T.Nagao, S.Müller "The n-level spectral correlations for chaotic systems" *J. Phys. A: Math. Theor.* vol **42** 375102 (2009)