

TR-H-063

発振ニューラルネットの遊走的な出力軌道の  
性質変化に関する実験的解析

林 幸雄

大森 洋一\*

(\* 奈良先端科学技術大学院大学)

1994. 3. 14

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

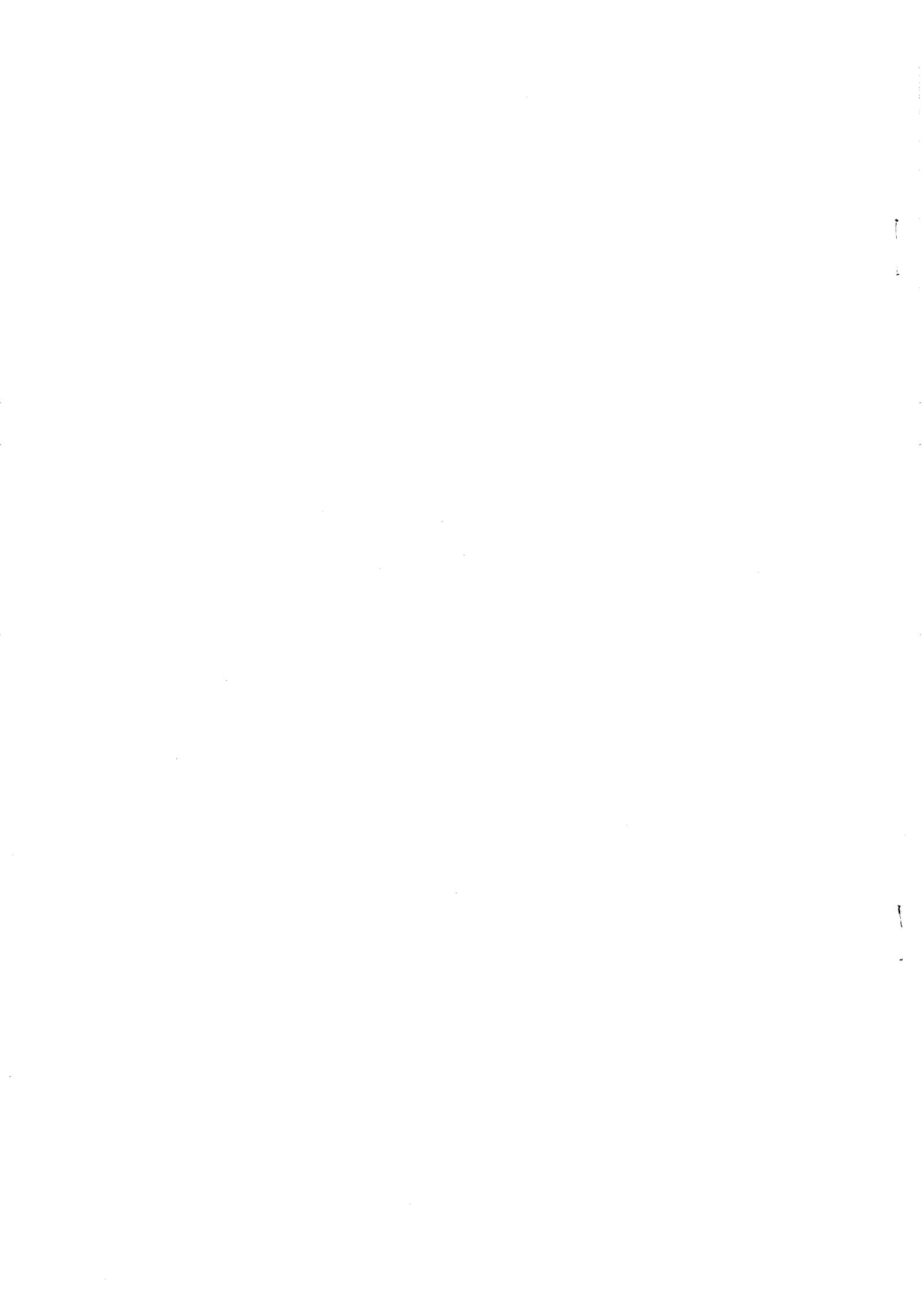
# 発振ニューラルネットの遊走的な出力軌道の 性質変化に関する実験的解析

林 幸雄  
大森 洋一\*

\* 奈良先端科学技術大学院大学 実習生

## 概要

(株) ATR 人間情報通信研究所では、生物の脳内の優れた情報処理能力の工学的実現を目指したニューラルネットモデルの研究を行っている。特に、脳内における時空間パターン処理に着目した、時間変化する出力を自律生成できるモデルを考えている。その一つとして、興奮性素子と抑制性素子のペアで構成される発振ニューラルネットが既に提案され、素子変数に関する  $N$  次元 ( $N \geq 3$ ) の状態空間内で、記憶パターン間を遊走する複雑な軌道を持つことが報告されている。一方、平衡状態に収束する従来の対称相互結合型の連想記憶モデルでは、記憶パターン数の増加に伴う記憶パターン想起の破綻現象が知られているが、同様のことが上記のモデルでも起こると予想される。すなわち、記憶パターン数の増加に伴った記憶パターンへの引力の弱体化によって、軌道の遊走領域が記憶パターン間から  $N$  次元超立方体の他の頂点間にまで広がると考えられる。本論文は、このような遊走領域の性質変化を  $N = 30$  と  $50$  の場合で実験的に示した、学外実習の結果をまとめたものである。



# 1 背景

人間の脳内では多数の神経細胞が結合した回路網によって、情報処理が行われている。通常のノイマン型コンピュータと比較すると、脳は曖昧なパターン情報の認識や、適応的な運動制御に関する優れた能力を持っている。こうした脳の情報処理能力の工学的実現を目指して、これまで数多くの人工的なニューラルネットのモデルが提案されている [5] [6]。しかし、従来のモデルでは、最終的な静止状態から入力パターンを分類するものが多く、この場合、出力の時間変化は情報処理に直接的には利用されていない。

これに対して、時空間パターン処理に適した、軌道の多様性と可制御性の特徴を有するモデルとして、互いに反対符号の重み係数で結合された興奮 - 抑制性ペア素子で構成される発振ニューラルネットがすでに提案されている [3]。ここで多様性とは、状態空間に漸近安定平衡点やリミテッドサイクル、カオスなどの多様なアトラクタが存在することを指す。また、可制御性とは上記のさまざまなアトラクタ上の軌道の遊走領域を外部入力などのパラメータによっておおまかに制御できることを指す。このような軌道の多様性と可制御性の特徴は、時空間パターン処理に利用できる。例えば、記憶パターン近傍のリミテッドサイクルを記憶パターンの想起機能に、記憶パターン間のカオスの軌道を記憶パターンの探索機能に対応づけることができる。

上記のモデルにおいて、パターンの想起機能や探索機能を効率的に実行するにあたり、モデルの動的な軌道が状態空間内でどのように記憶パターンに引き込まれたり、離れたりするかが重要となるが、これについてはまだ十分解明できていない。記憶パターンへの引き込みに関する性質の 1 つとして、初期状態にしたがった静止状態に収束する対称相互結合型の連想記憶モデルでは、記憶パターン数の増加に伴って、突然に記憶パターンへの引力の弱体化が起こって、他の点にトラップされる現象がある [1] [2]。上記の発振ニューラルネットにおいても同様の現象が起こることが予想される。本研究では、このような記憶パターン数の増加に伴う遊走軌道の性質変化として、軌道の遊走領域が記憶パターンの頂点間からそれら以外の頂点間まで広がることを、実験的な解析によって示した。

# 2 発振ニューラルネット

著者らによって、それぞれ  $N$  個ずつの興奮性素子と抑制性素子からなる発振ニューラルネットワークが提案され、その動的な挙動の性質が部分的に解明されている [3] [4]。このネットワークの構造的特徴は、各興奮性素子が他の興奮性素子と相互結合する一方、抑制性素子是对應する興奮性素子とのみ結合する点である。ネットワークの動作方程式は以下の式 (1) ~ (4) で規定される。

$$\frac{dx_i}{dt} = -x_i + G(\sum_{j=1}^N W_{ij}x_j - K_{EI}y_i + I_i) \quad (1)$$

$$\frac{dy_i}{dt} = -y_i + G(K_{IE}x_i) \quad (2)$$

$$G(z) = \frac{2}{\pi} \arctan\left(\frac{z}{a}\right) \quad (3)$$

$$W_{ij} = \frac{1}{N} \sum_{\mu=1}^M \xi_i^{\mu} \xi_j^{\mu} + \delta_{ij} \quad (4)$$

ここで、 $x_i$ ,  $y_i$  はそれぞれ興奮性素子と抑制性素子に対する細胞集団の平均パルス密度、 $g(z)$  は平均膜電位から平均パルス密度へのシグモイド関数で、 $a$  はその関数のパラメータである。また、 $W_{ij}$  は興奮性素子間の結合重みで、 $M$  個の記憶パターン  $\xi^{\mu}$  の相関行列で設定され、 $\delta_{ij}$  はクロネッカーのデルタである。 $-K_{EI}$  および  $K_{IE}$  はそれぞれ、興奮 - 抑制性ペア素子の抑制結合と興奮結合の重み係数で、 $I_i$  は興奮性素子の入力バイアスを表す。出力パターンは各興奮性素子  $x_i(t)$  の値とした。図 1 に本モデルのネットワーク構造を示す。

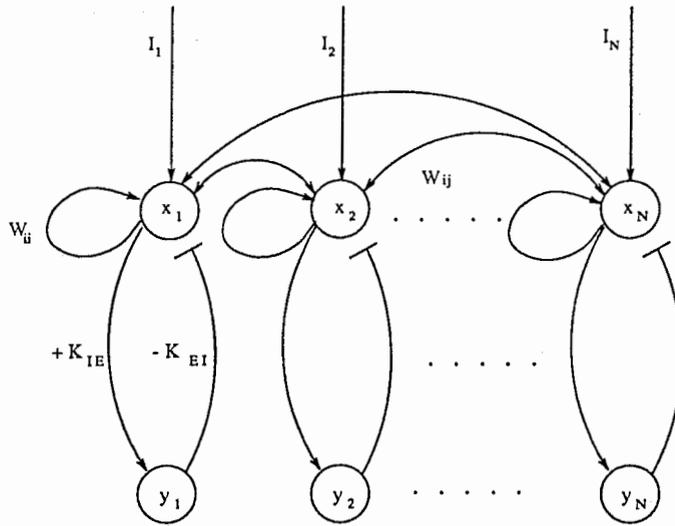


図 1: 発振ニューラルネットワーク (矢印  $\dashv$  と  $\rightarrow$  は興奮性結合と抑制性結合を表す)

この発振ニューラルネットワークは、初期状態  $x_i(0)$ 、 $y_i(0)$  や入力バイアス  $I_i$ 、結合重み係数等のパラメータの値に従って、漸近安定平衡点からリミットサイクル、カオスまでの多様な出力軌道を自律生成する。特に、入力バイアスをおある一つの記憶パターンの充分近くに設定した時、その記憶パターン付近にリミットサイクルまたは漸近安定平衡点が生成されることや、入力バイアスを記憶パターンから離れたパターンに設定した時には、記憶パターン間を遊走するカオス軌道が生成されることなどが知られている [3]。

### 3 実験内容

対称相互結合型の連想記憶モデルにおける記憶パターン数の臨界値  $M \approx 0.14N$  前後における突然の記憶パターン想起の破綻と同様の現象 [1] [2] が、発振ニューラルネットにも起こり得ると考えられる。すなわち、記憶パターン数の増加に伴って、記憶パターンへの引力の弱体化が起こり、出力軌道の遊走領域が記憶パターン間から  $N$  次元超立方体 ( $-1 \leq x_i \leq 1$ ) 上の記憶パターン以外の他の頂点にまで、広がるようになることが予想される。実際、次元数  $N = 22$  の場合で、このことが実験的に示されている [4]。

そこで、より大きな次元数の場合で上記と同様な、出力軌道に対する  $N$  次元超立方体の最近接頂点の時間的な遷移と各頂点への滞在時間を調べてみた。具体的には  $N = 30$  のときは  $M = 4 < 0.14N$  と  $M = 5 > 0.14N$  に対して、 $N = 50$  のときは  $M = 6 < 0.14N$  と  $M = 8 > 0.14N$  に対して以下の 2 つの項目についての解析を行った。

- 実験 1. 分布グラフの比較

分布グラフとして、軌道の最近接頂点への滞在時間を、基準とした頂点  $\xi^1$  から異なるビット数ごとの頂点に対して、集計したヒストグラムを作った。

この分布グラフにより、記憶パターンの頂点とそれら以外の頂点に対する軌道の滞在率の比較を行った。

この際、定常分布に落ち着くまでの反復回数と、分布グラフの形の初期状態  $\{x_i(0)\}$  に対する依存性についても検討した。

## ● 実験 2. 最近接頂点の時間的遷移

記憶パターン  $\xi^\mu$  とその反転パターン  $-\xi^\mu$  の頂点、さらにそれら以外の頂点に対する軌道の最近接頂点の時間的遷移から、軌道の周期性や、記憶パターンの頂点と他の頂点との滞在頻度の比較について検討を行った。

また、同じネットワークサイズ  $N$  と記憶パターン数  $M$  の場合で、異なる初期状態  $\{x_i(0)\}$  からの遷移の違いを比較した。

ここで、式 (1) ~ (4) の動作方程式の数値計算には、4 次のルンゲ・クッタ法を用いた (刻み幅  $\Delta t = 0.05$ )。ネットワークのパラメータは、興奮 - 抑制性ペア素子の発振条件を満足するように [3]、 $a = 0.1$ 、 $K_{EI} = K_{IE} = 2.0$  とし、入力バイアスは軌道の遊走領域を偏らせないように  $I_i = 0$  とした。各記憶パターン  $\{\xi_i^\mu\}$  は  $N$  個の要素中、同数の 1 と  $-1$  をもつベクトルで、任意の記憶パターン間の相関  $|\sum_{i=1}^N \xi_i^\alpha \xi_i^\beta|$  ( $\alpha \neq \beta$ ) が互いに  $\sqrt{N}$  以下になるような、ほとんど直交なものをランダムに選んだ。また、興奮性素子の初期状態は、原点近くの  $|x_i(0)| < 0.2$  のものをランダムに選び、抑制性素子の初期状態は、入力バイアスと等価的な効果を持つこと [4] から、 $y_i(0) \equiv 0$  とした。

## 4 実験結果

### 4.1 実験 1. の結果

まず、定常分布に落ち着くまでの反復回数に関する結果を述べる。 $N$  と  $M$  の各組合せに対して、反復回数を 10 万回から 50 万回まで 10 万回刻みで増加した場合の分布の形の変化を図 2 から 図 5 に示す。図中のグラフの黒棒は記憶パターンおよびその反転パターンに関する滞在率を、白棒はそれら以外の頂点に関する滞在率を表す。これらの結果から、反復回数が 40 万回以内で定常分布に落ち着いたと考えられる。

そこで、以下の検討では反復回数が 40 万回の時で、記憶パターン数の臨界値  $M \approx 0.14N$  前後における、軌道の最近接頂点に対する滞在率の分布や時間的遷移を比較した。

図 6 と図 7 を比較すると、分布の形が初期状態に依存して異なっているものの、記憶パターン数の臨界値  $M \approx 0.14N$  前後における分布の変化、すなわち記憶パターンへの引力の弱体化が読み取れる。但し、 $N = 30$  の場合 (図 6 と図 7) に比べて、 $N = 50$  の場合 (図 8 と図 9) の方が、臨界値前後に対する変化の傾向が顕著でなく、初期状態に対する分布の形の違いも大きくなっている。

### 4.2 実験 2. の結果

最近接頂点の時間的遷移を図 10 から 図 13 に示す。図中のグラフの縦軸は記憶パターンの番号であり、 $1 \leq \mu \leq M$  が記憶パターン  $\xi^\mu$  を、 $M + 1 \leq \mu \leq 2M$  が反転パターン  $-\xi^{\mu-M}$  に相当する。また、番号 0 は、記憶パターン以外の頂点を表す。これらの図から、軌道が周期性を持たないカオス的で複雑な遷移をしたことが伺える。

また、図 10 と 図 12 (図 12 と 図 13) の対応するグラフを比較すると、臨界値前の方がその後より記憶パターンの頂点への滞在頻度が少ない (縞模様が密でなく、記憶パターン以外の頂点に滞在する時間が多い) ことがわかる。これらの図から、 $M < 0.14N$  から  $M \geq 0.14N$  の記憶パターン数の増加に伴って、記憶パターンへの引力の弱体化による出力軌道の劇的な変化が起こったことが伺える。他の初期状態の場合も上記と同様な結果が得られたが、その頂点間の時間的遷移 (滞在時間や順序) は初期状態に大きく依存した異なるものであった。

## 5 結論

以上の結果を総合すると、文献 [4] の  $N = 22$  の場合と同様な記憶パターン数の増加に伴う軌道の性質変化が、 $N = 30$  と  $50$  の場合でも得られたものの、次元数が大きくなる程その傾向が顕著でなくなることや、軌道の最近接頂点の分布の形が初期状態に依存するという、前者とは異なる結果も得られた。これらの相違点に関しては、さらに詳細な検討を行う必要がある。

## 参考文献

- [1] Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1987). Statistical Mechanics of Neural Networks near Saturation. *Annals of Physics*, **173**, 30-67.
- [2] Geszti, T. (1990). *Physical Models of Neural Networks*. World Scientific.
- [3] Hayashi, Y. (1994a). Oscillatory Neural Network and Learning of Continuously Transformed Patterns. *Neural Networks* (in press).
- [4] Hayashi, Y. (1994b). Dynamical Search Domain Control According to the Input Bias of an Oscillatory Neural Network. *Submitted to Neural Networks*.
- [5] Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proceeding of the National Academy of Science, USA*, **81**, 3088-3092.
- [6] Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (1986). *Parallel Distributed Processing*. Volume 1 & 2, The MIT Press.

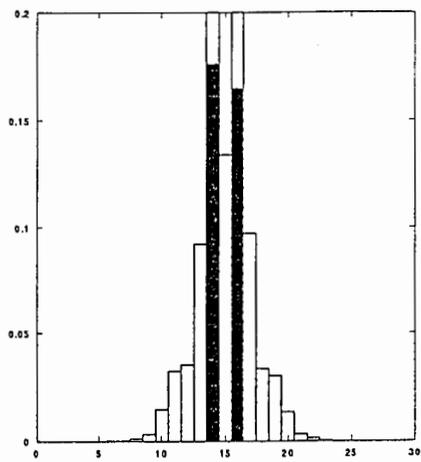


図 2-1: 100,000 回

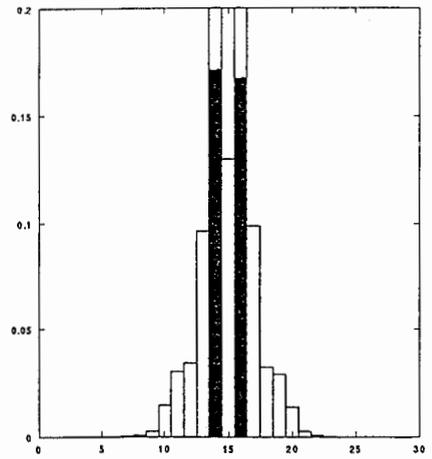


図 2-2: 200,000 回

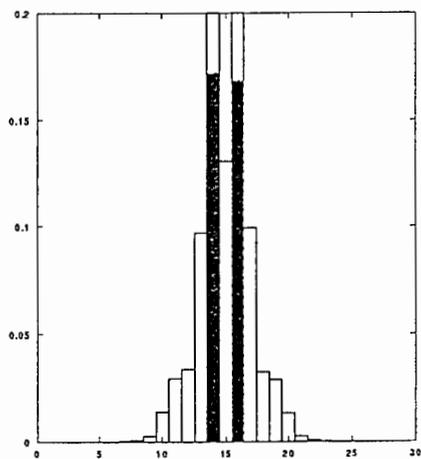


図 2-3: 300,000 回

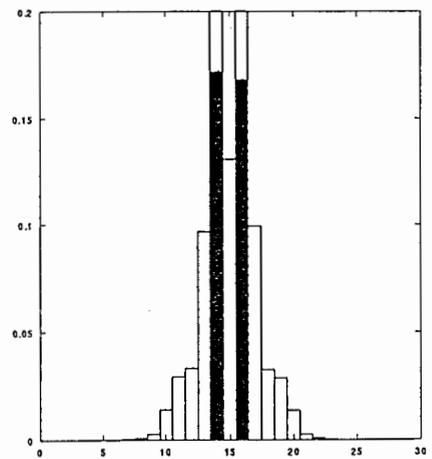


図 2-4: 400,000 回

図 2: 実行時間による分布の変化 ( $N = 30$   $M = 4$ )

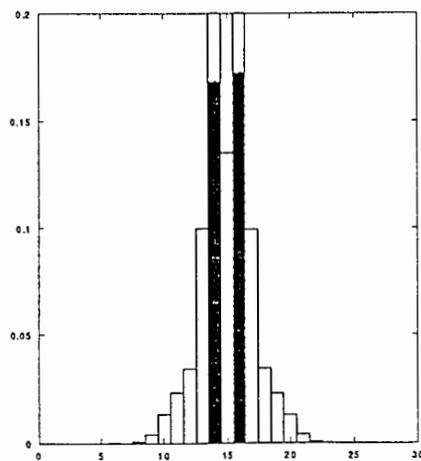


図 2-5: 500,000 回

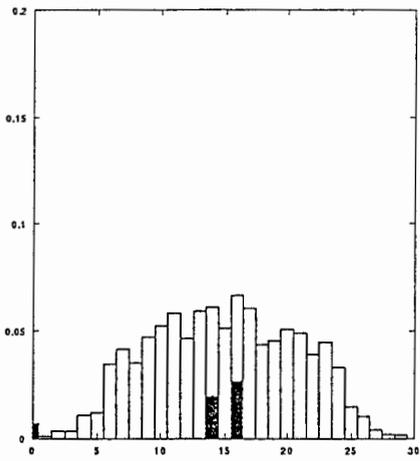


図 3-1: 100,000 回

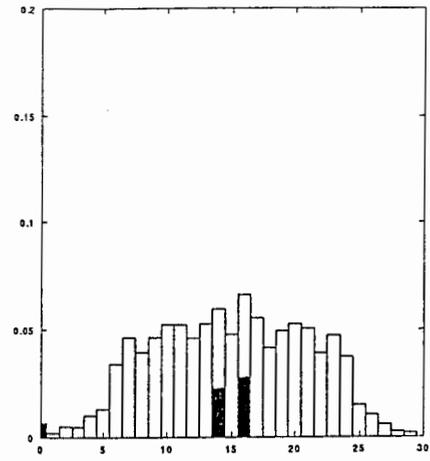


図 3-2: 200,000 回

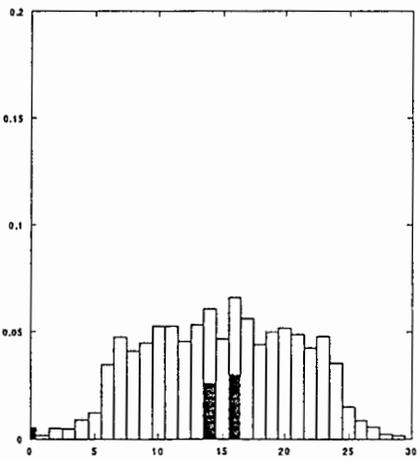


図 3-3: 300,000 回

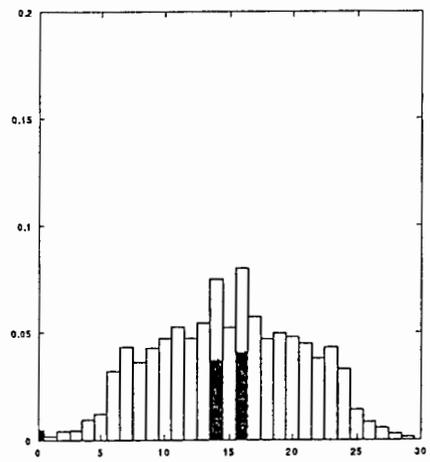


図 3-4: 400,000 回

図 3: 実行時間による分布の変化 ( $N = 30$   $M = 5$ )

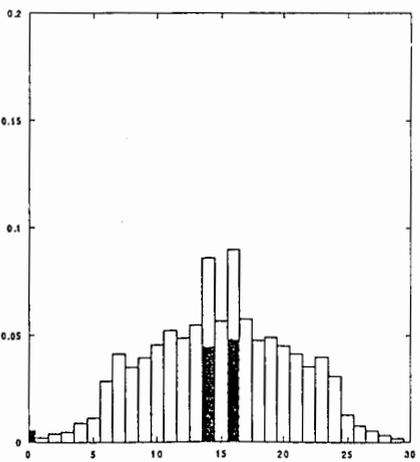


図 3-5: 500,000 回

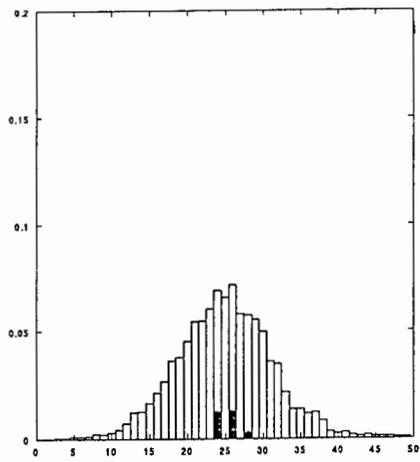


図 4-1: 100,000 回

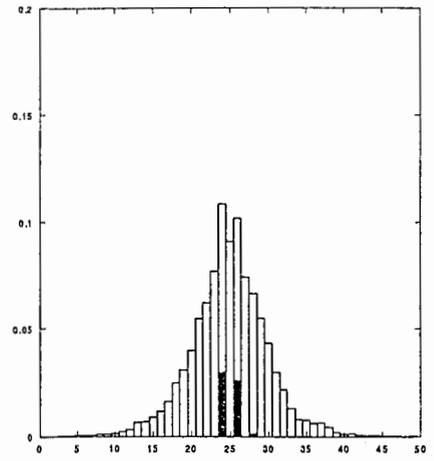


図 4-2: 200,000 回

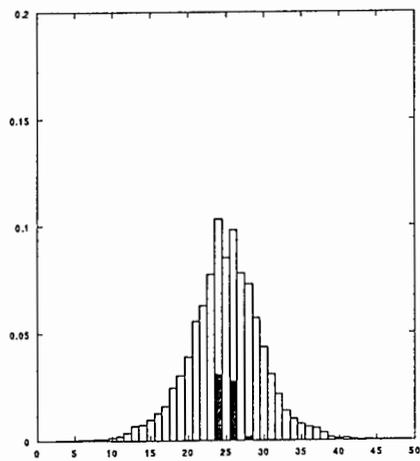


図 4-3: 300,000 回

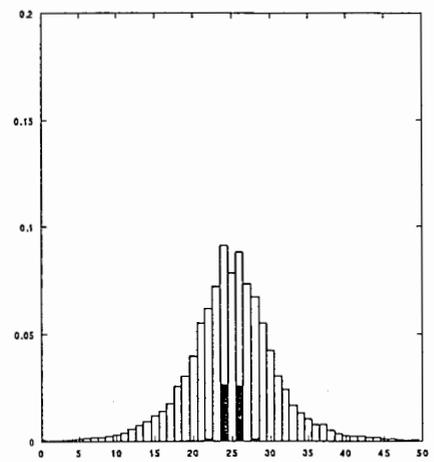


図 4-4: 400,000 回

図 4: 実行時間による分布の変化 ( $N = 50$   $M = 6$ )

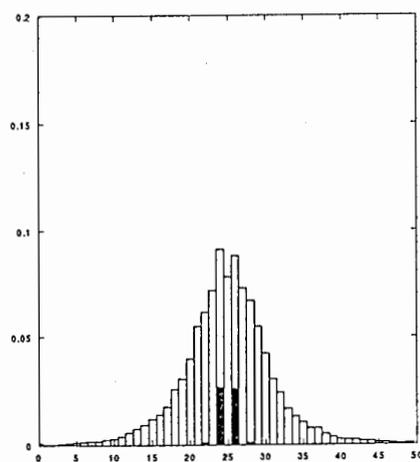


図 4-5: 500,000 回

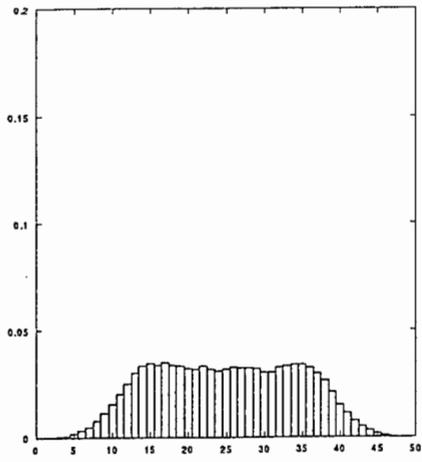


図 5-1: 100,000 回

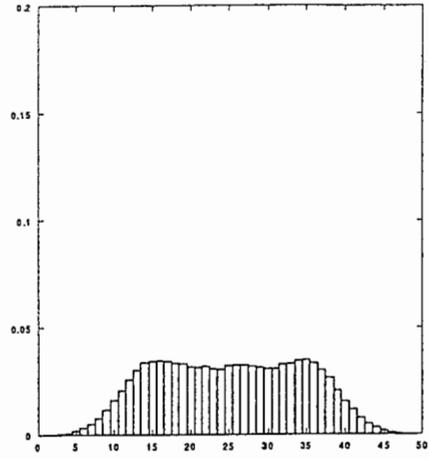


図 5-2: 200,000 回

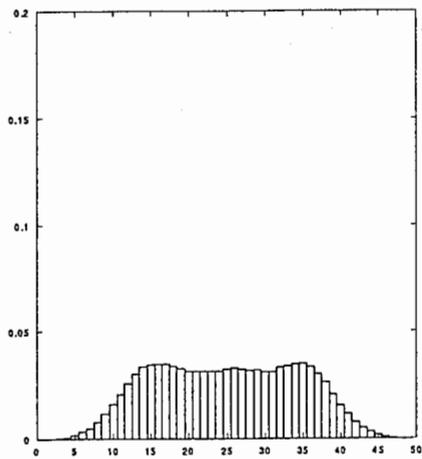


図 5-3: 300,000 回

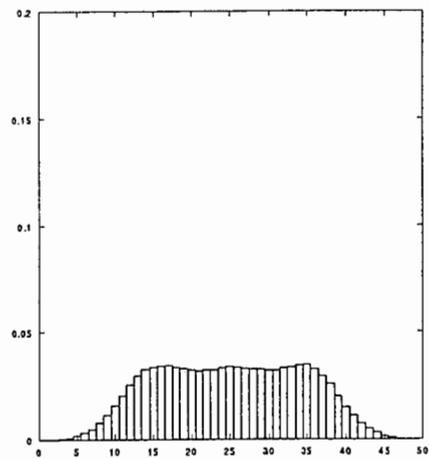


図 5-4: 400,000 回

図 5: 実行時間による分布の変化 ( $N = 50$   $M = 8$ )

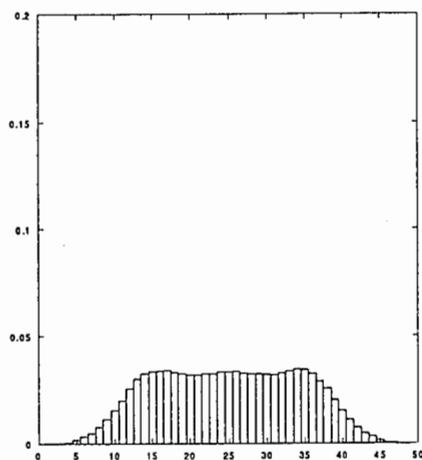


図 5-5: 500,000 回

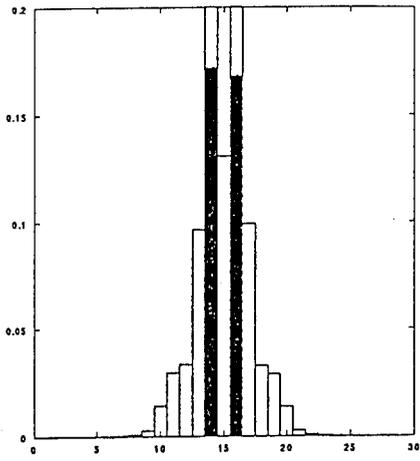


図 6-1:  $N = 30, M = 4$

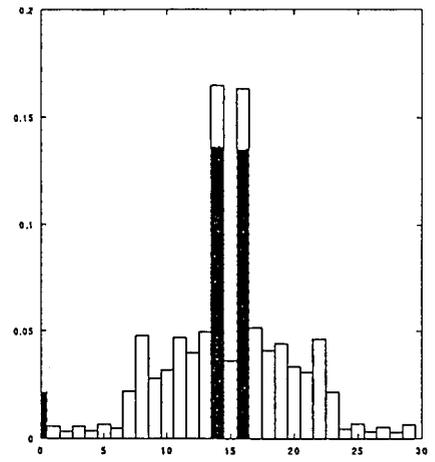


図 6-2:  $N = 30, M = 4$

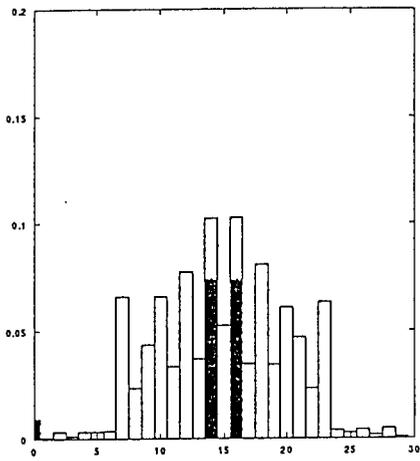


図 6-3:  $N = 30, M = 4$

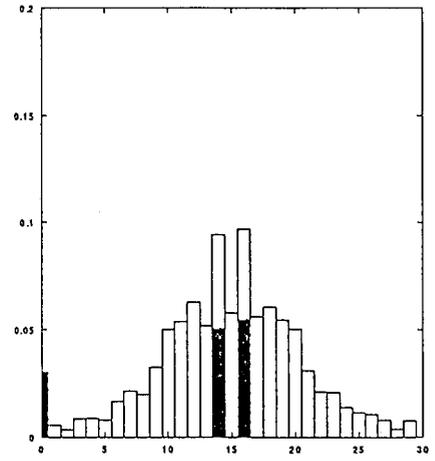


図 6-4:  $N = 30, M = 4$

図 6: 初期状態による分布の変化 (40 万回反復)

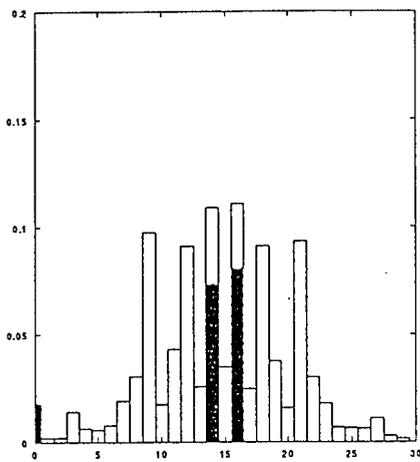


図 6-5:  $N = 30, M = 4$

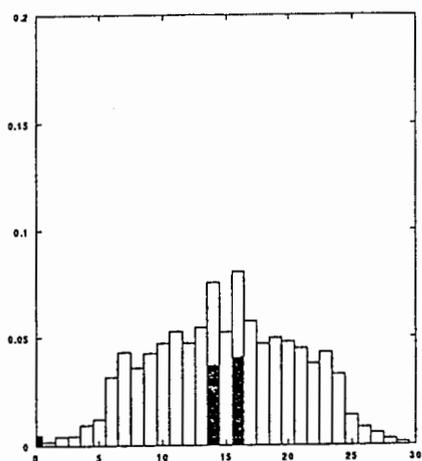


図 7-1:  $N = 30, M = 5$

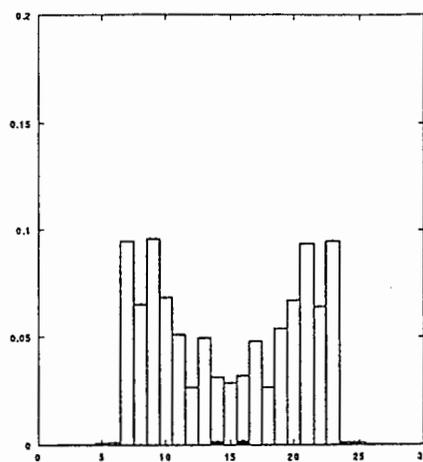


図 7-2:  $N = 30, M = 5$

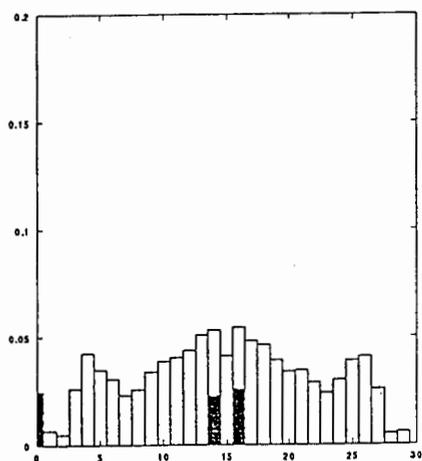


図 7-3:  $N = 30, M = 5$

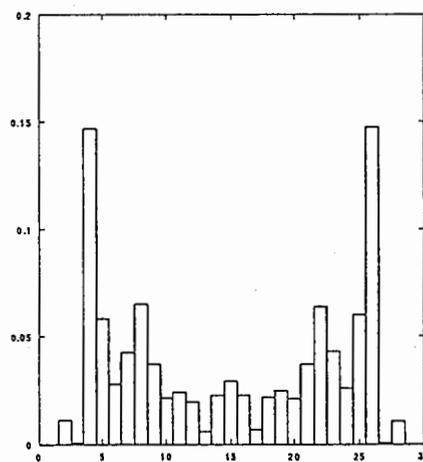


図 7-4:  $N = 30, M = 5$

図 7: 初期状態による分布の変化 ( $N = 30, M = 5$ )

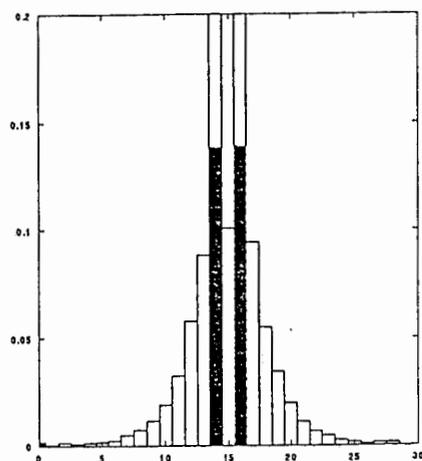


図 7-5:  $N = 30, M = 5$

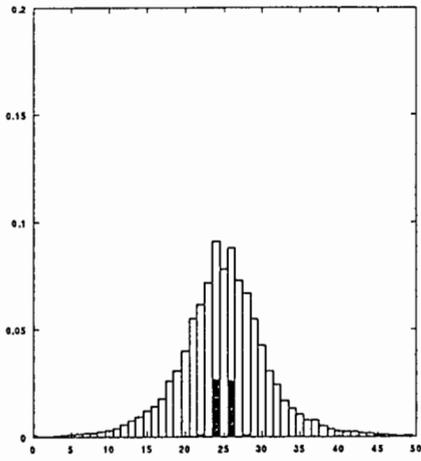


図 8-1:  $N = 50, M = 6$

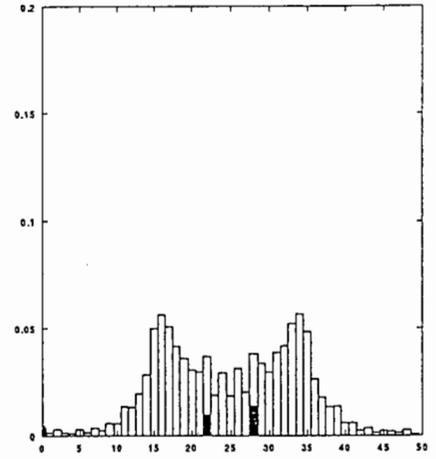


図 8-2:  $N = 50, M = 6$

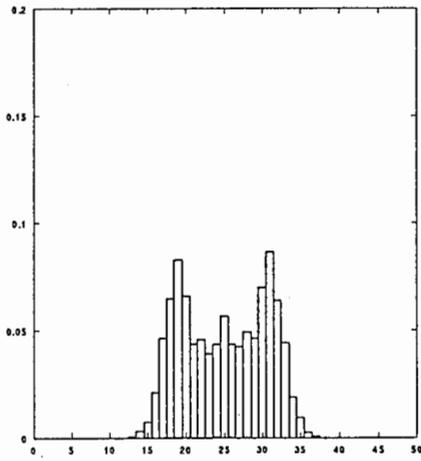


図 8-3:  $N = 50, M = 6$

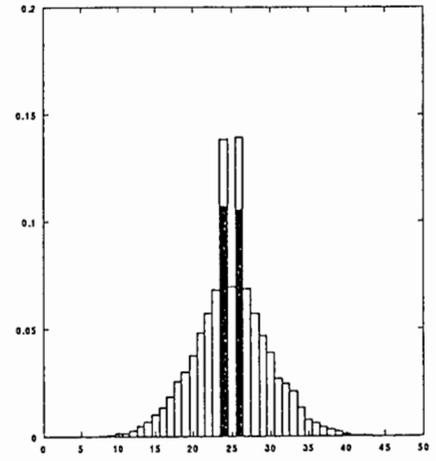


図 8-4:  $N = 50, M = 6$

図 8: 初期状態による分布の変化 ( $N = 50, M = 6$ )

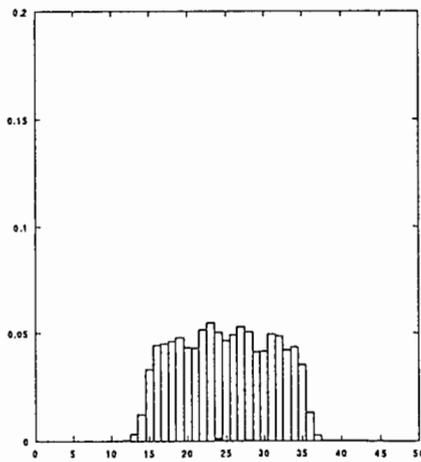


図 8-5:  $N = 50, M = 6$

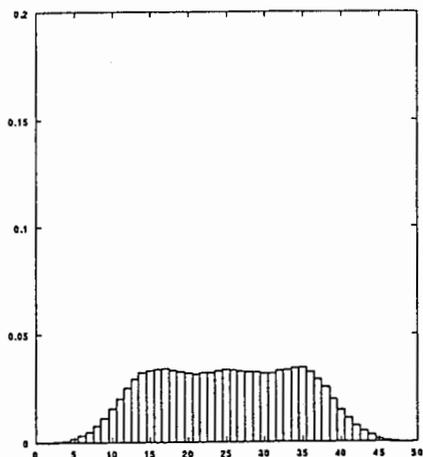


図 9-1:  $N = 50, M = 8$

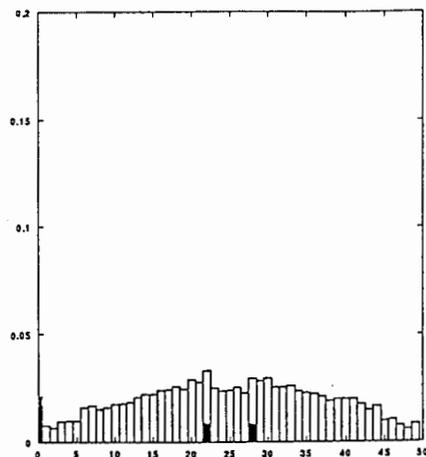


図 9-2:  $N = 50, M = 8$

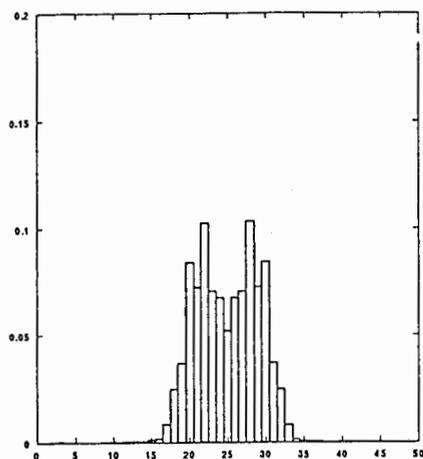


図 9-3:  $N = 50, M = 8$

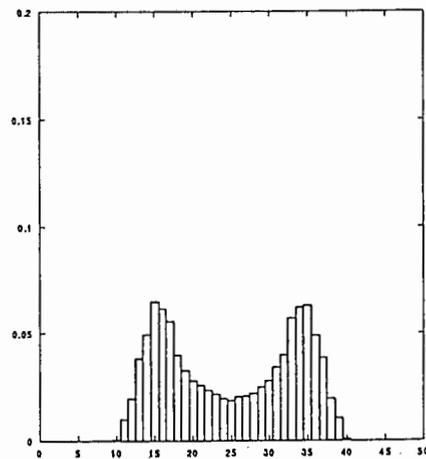


図 9-4:  $N = 50, M = 8$

図 9: 初期状態による分布の変化 ( $N = 50, M = 8$ )

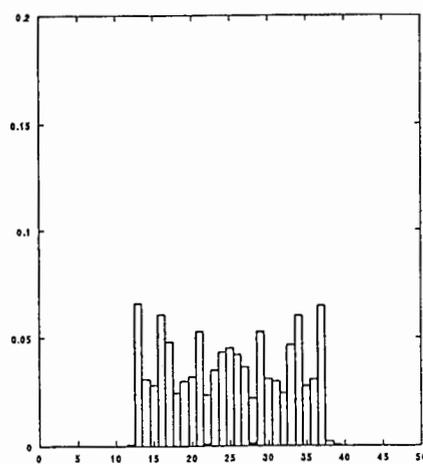


図 9-5:  $N = 50, M = 8$

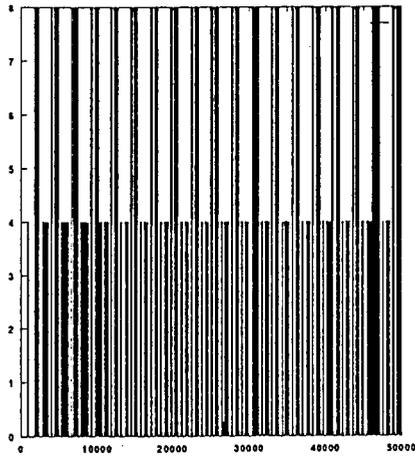


図 10-1: 50,000 回

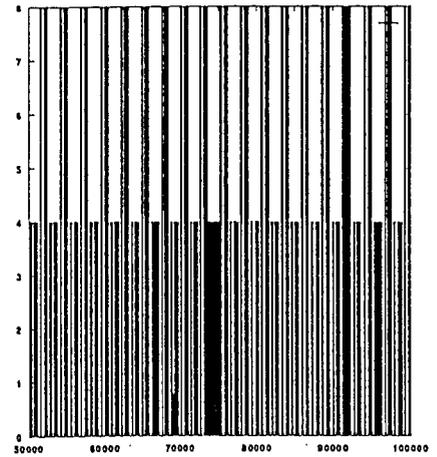


図 10-2: 100,000 回

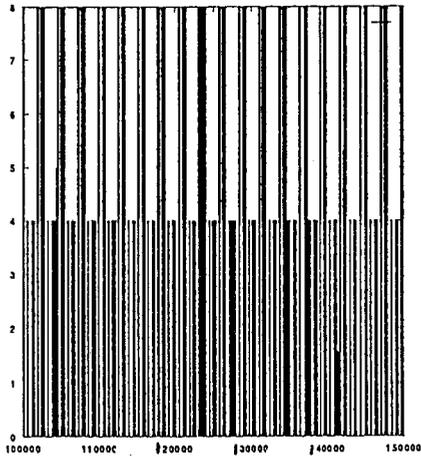


図 10-3: 150,000 回

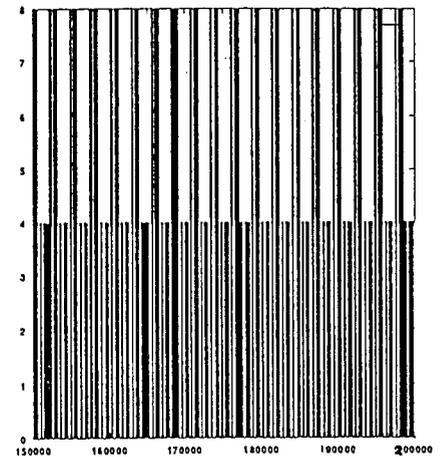


図 10-4: 200,000 回

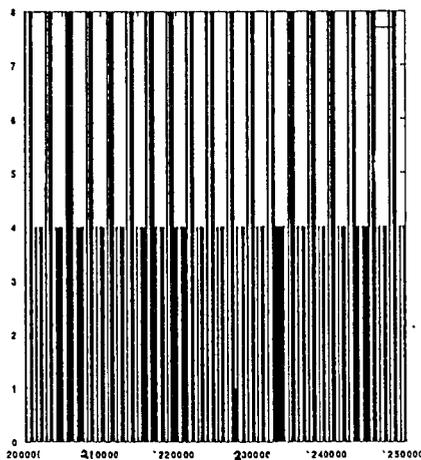


図 10-5: 250,000 回

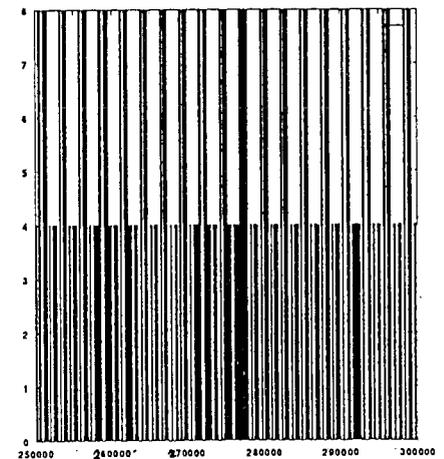


図 10-6: 300,000 回

図 10: 実行時間による頻度の変化 ( $N = 30$   $M = 4$ )

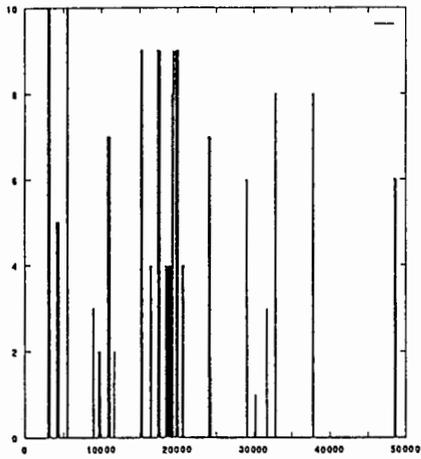


図 11-1: 50,000 回

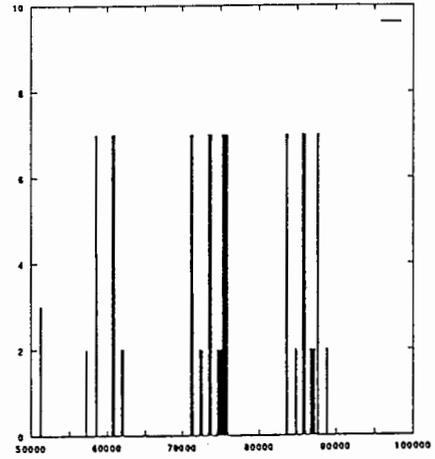


図 11-2: 100,000 回

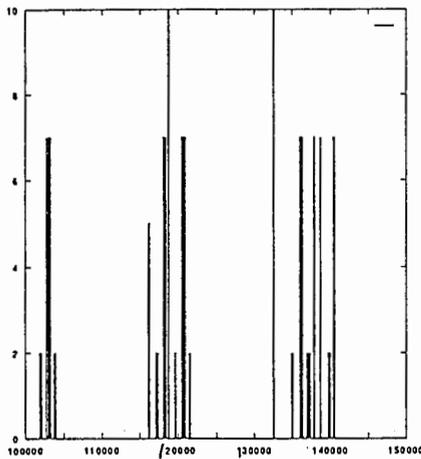


図 11-3: 150,000 回

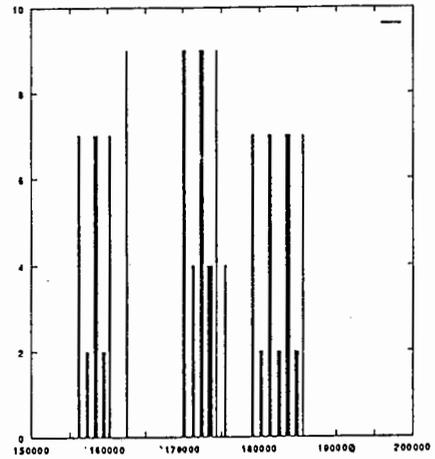


図 11-4: 200,000 回

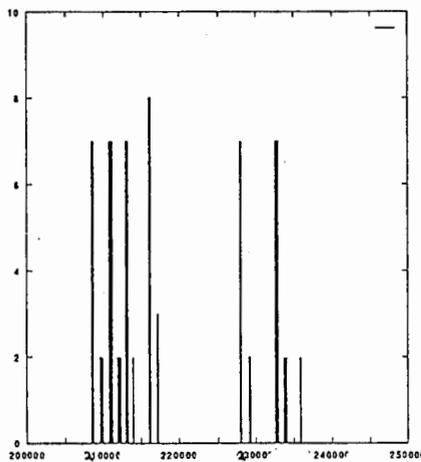


図 11-5: 250,000 回

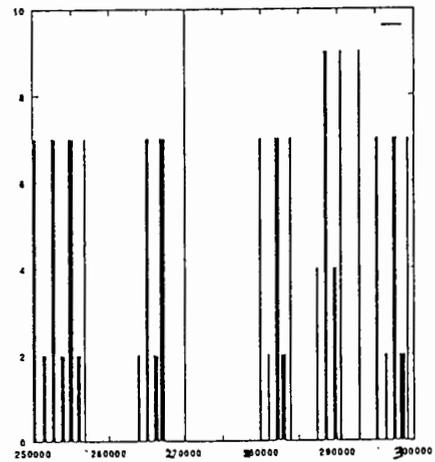


図 11-6: 300,000 回

図 11: 実行時間による頻度の変化 ( $N = 30$   $M = 5$ )

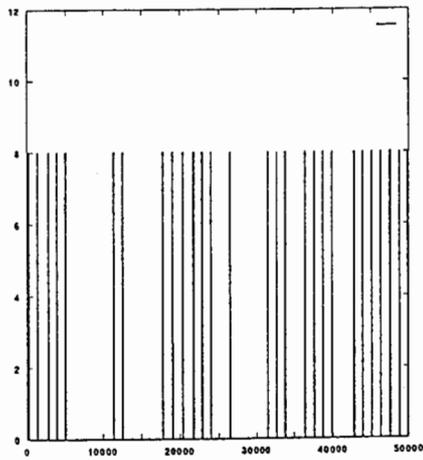


図 12-1: 50,000 回

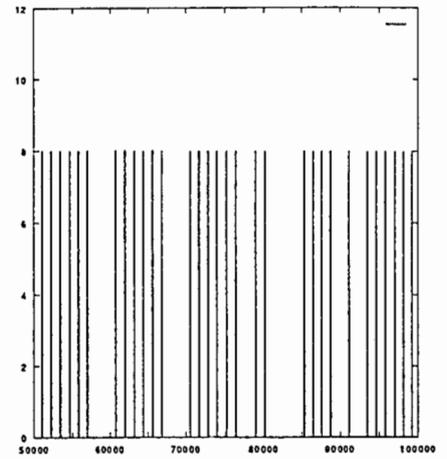


図 12-2: 100,000 回

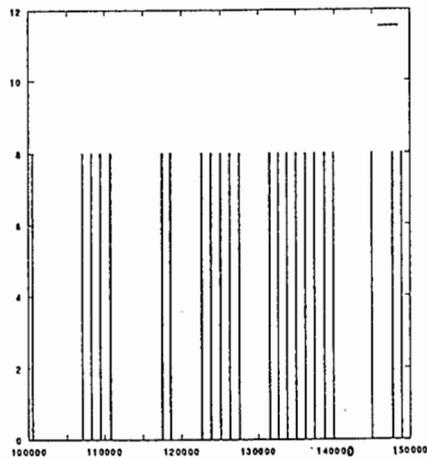


図 12-3: 150,000 回

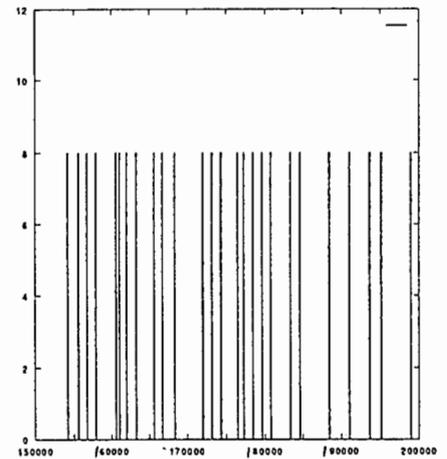


図 12-4: 200,000 回

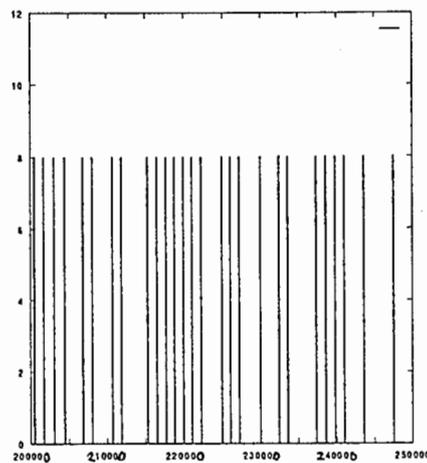


図 12-5: 250,000 回

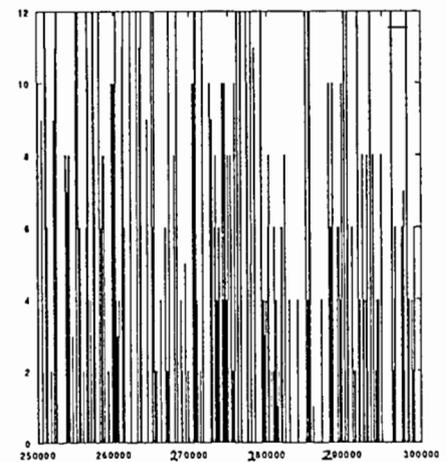


図 12-6: 300,000 回

図 12: 実行時間による頻度の変化 ( $N = 50$   $M = 6$ )

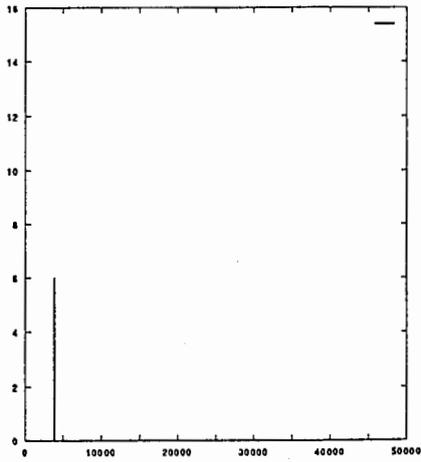


図 13-1: 50,000 回

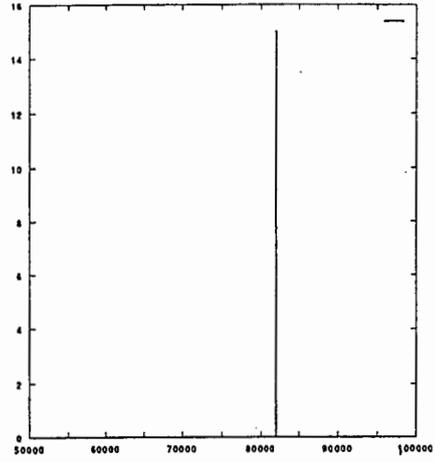


図 13-2: 100,000 回

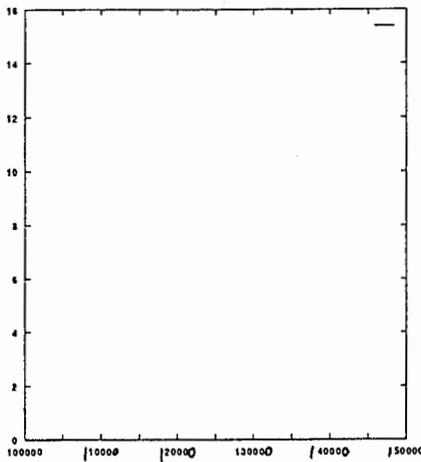


図 13-3: 150,000 回

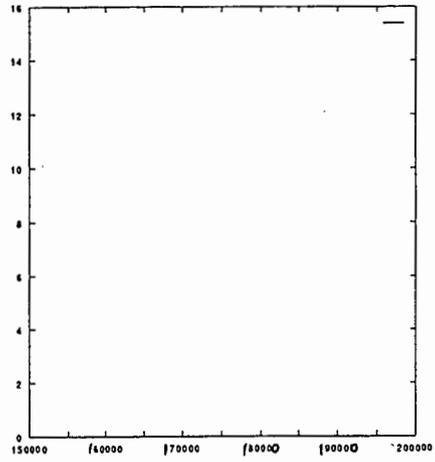


図 13-4: 200,000 回

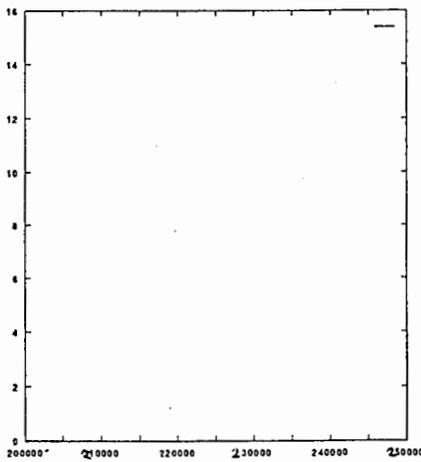


図 13-5: 250,000 回

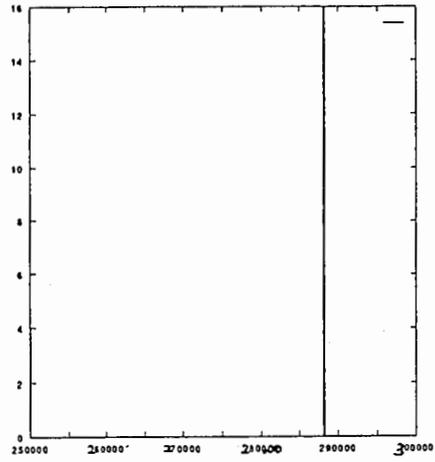


図 13-6: 300,000 回

図 13: 実行時間による頻度の変化 ( $N = 50$   $M = 8$ )