

TR - A - 0159

03

”リカレントネット、
カオス、
ウェーブレット”

佐藤 雅昭

1993. 1. 8

(1992.12.18 受付)

ATR 視聴覚機構研究所

〒619-02 京都府相楽郡精華町乾谷 ☎ 07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

Inuidani, Sanpeidani, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411

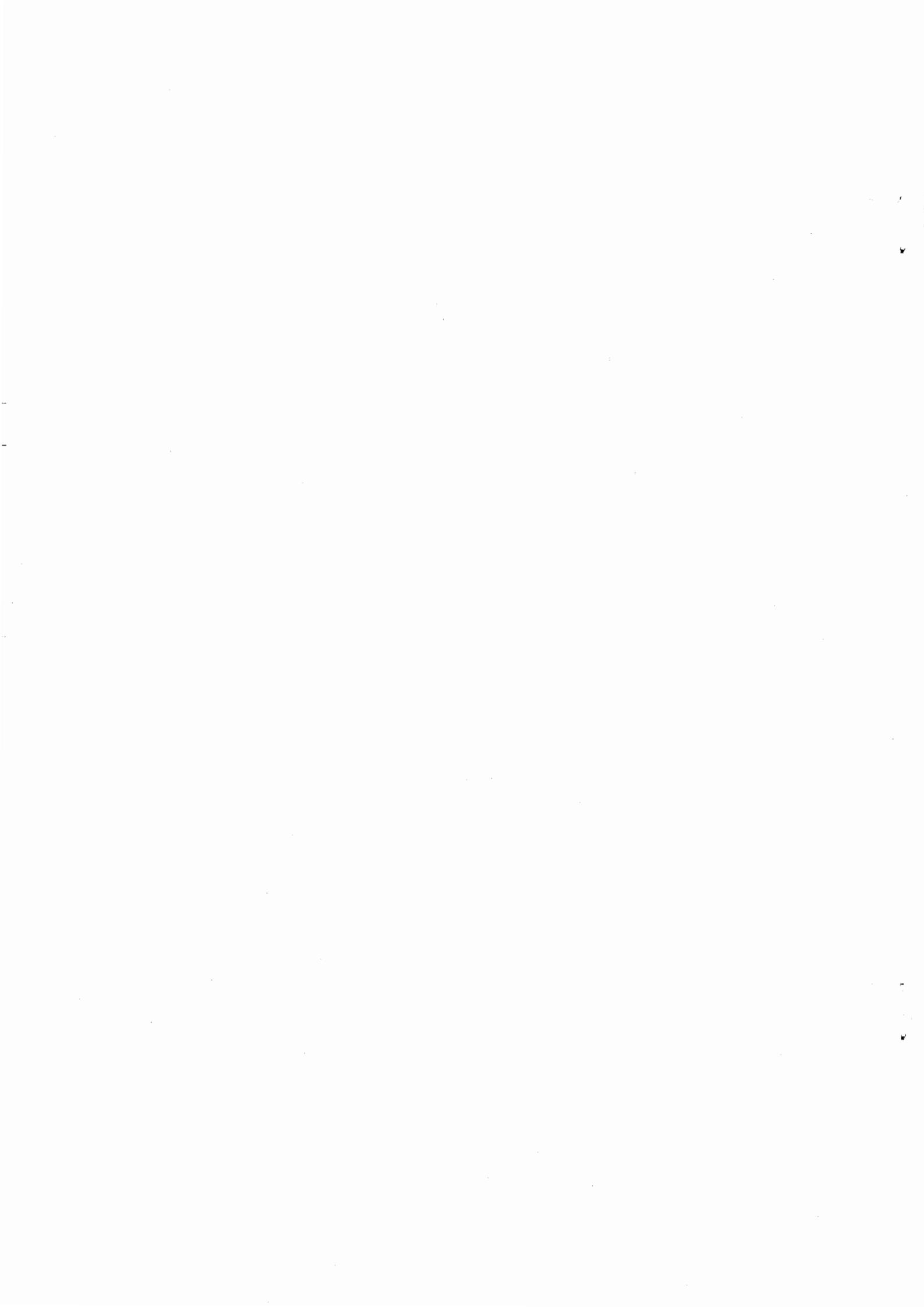
Faxsimile: +81-7749-5-1408

Telex: 5452-516 ATR J

“リカレントネット、
カオス、
ウェーブレット”

佐藤 雅昭

ATR 視聴覚機構研究所



目 次

1. はじめに
2. リカレントネットのダイナミクスと学習則
3. カオスダイナミクスの学習
4. 音声ゆらぎの学習
5. ウェーブレット



1. はじめに

本研究は、比較的単純なユニットが相互にフィードバック結合されたニューラルネットワーク（リカレントネット）の動的振る舞いを調べ、これを情報処理に応用することを目的としている。一方、生体システムの動的振る舞いの特徴としてカオスがある。カオスは、生体の神経細胞や神経回路網など生体のさまざまなレベルで観測されており、カオスが生体の情報処理にとって重要な役割を果たしていることが指摘されている。そこで、本報告ではカオスの情報処理への応用も試みている。

本編では、まず始めにリカレントネットのダイナミクスと学習則を説明し、リカレントネットを情報処理に応用するための枠組みを与える。次にその応用例としてカオスをリカレントネットに学習させるシミュレーションを解説する。また、人間の音声の自然さにとって重要な要素である音声波形のゆらぎをリカレントネットに学習させ、これをより自然な人工音声の合成に応用するための試みも解説する。最後に、カオスやゆらぎの自己相似的な性質を解析するための手段として、ウェーブレット理論を紹介する。

Neural Network

Type	Connection	Output	Information Processing
MLP	Feedforward	Nonlinear trans. of input patterns	Stable & Passive
Relaxation	Symmetric feedback	stable output (fixed point)	
Recurrent	Arbitrary feedback	Temporal output (limit cycle, chaos)	Dynamic & Active

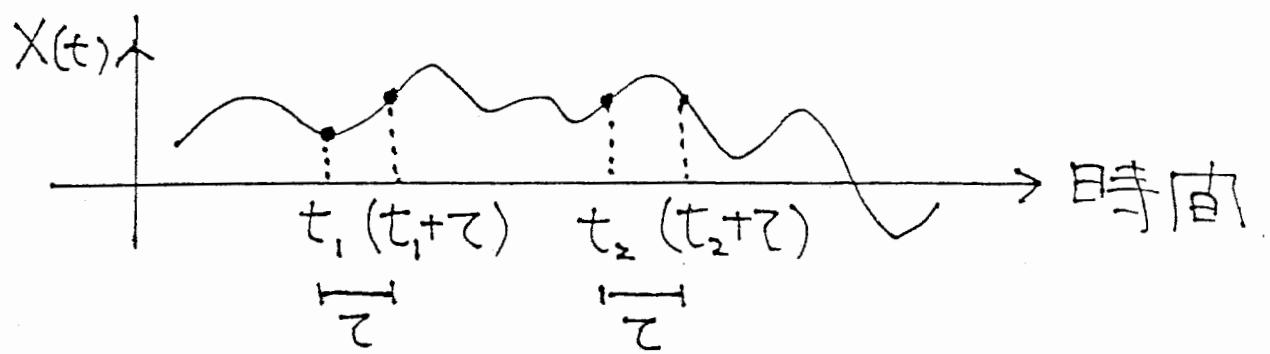
Nonlinear behavior of a recurrent network

→ Active and dynamic information processing

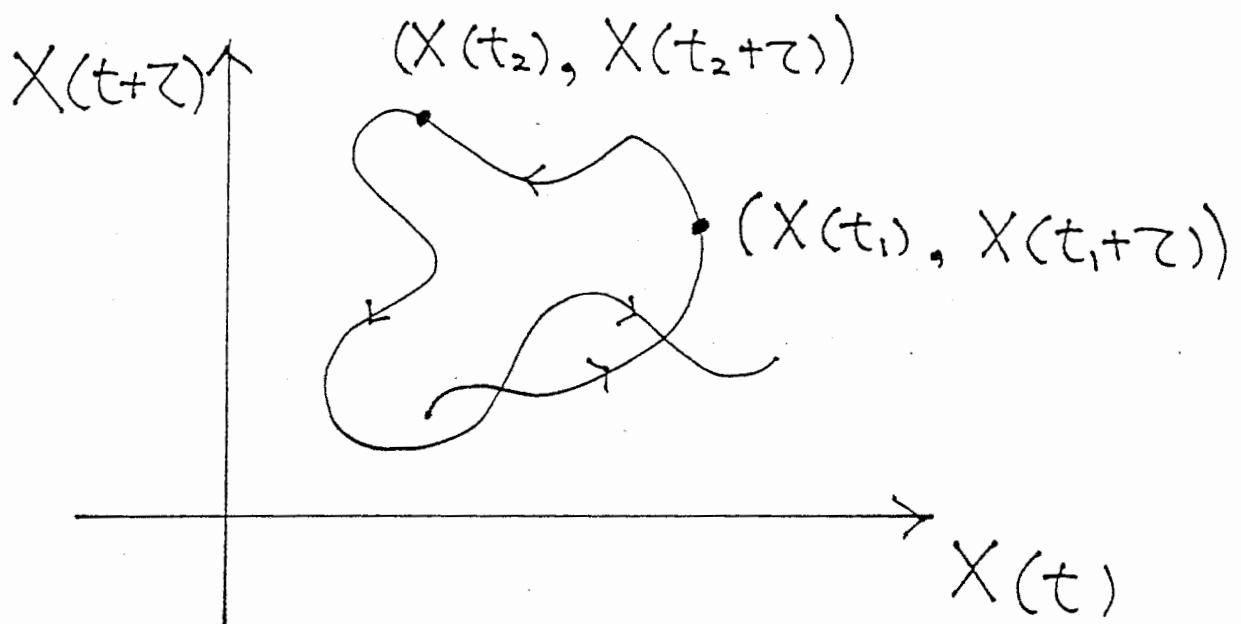




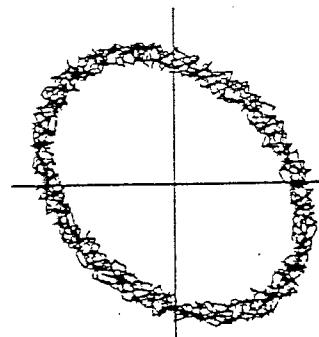
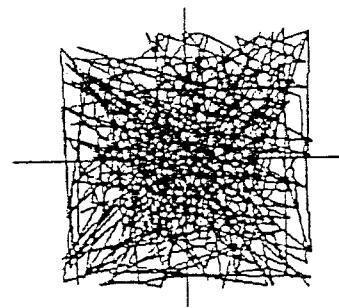
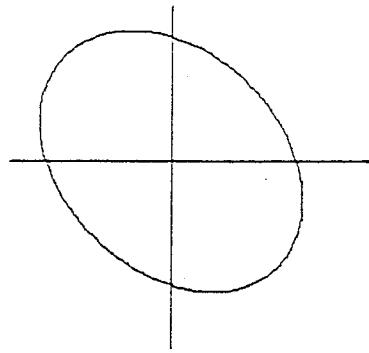
時向波形と位相図



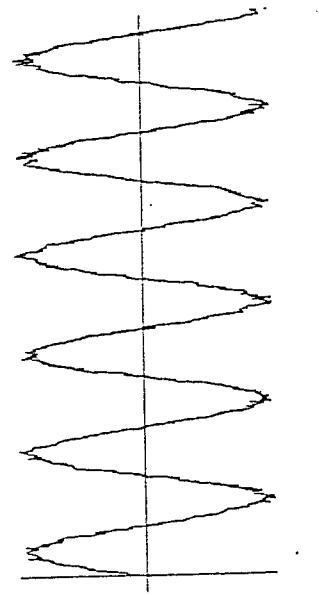
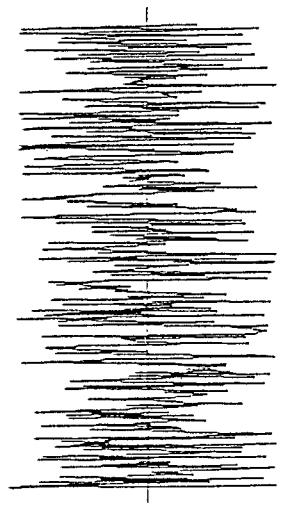
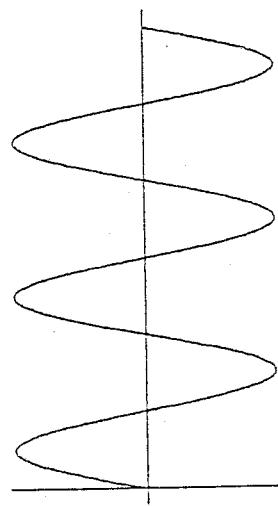
位相図



位相圖



周期波形



周期波形

ランダムノイズ

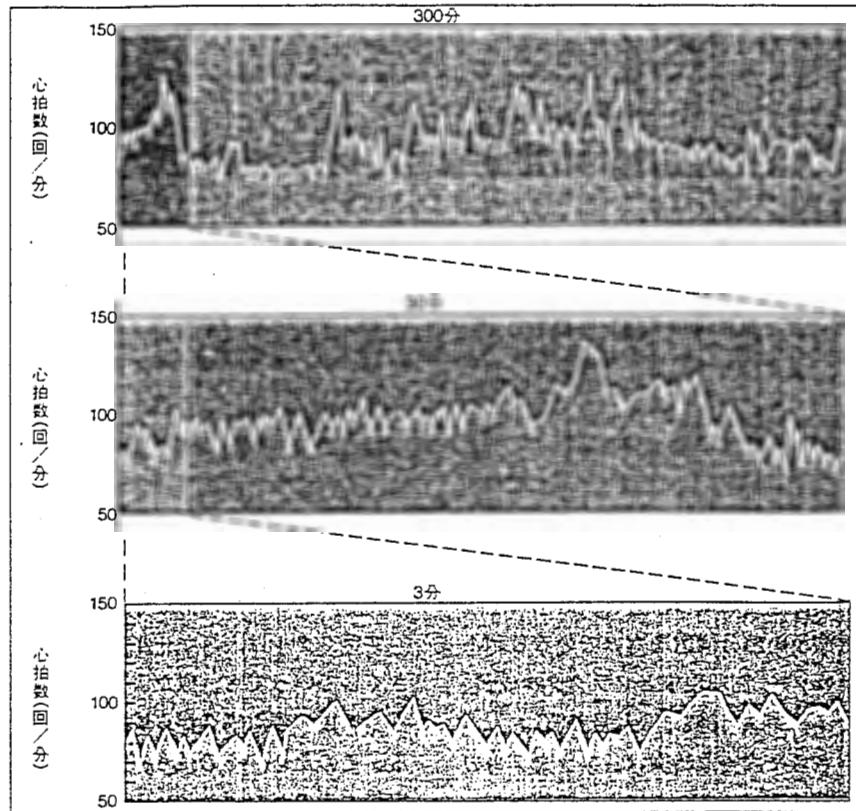
周期波形

+ ランダムノイズ

生物システムとカオス、ゆらぎ

○心臓の鼓動のゆらぎ

・自己相似性



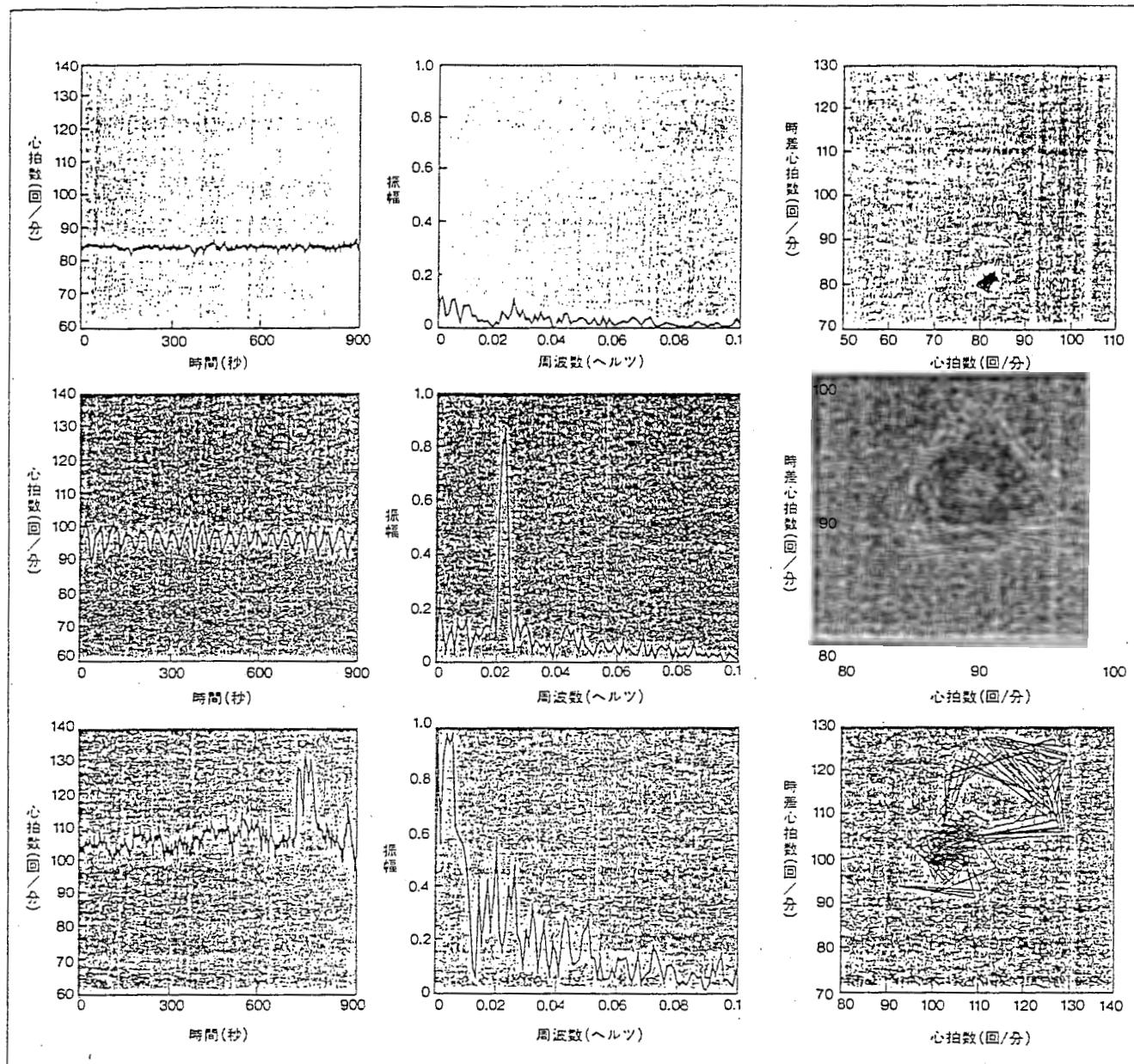
(A.L. Goldberger et.al.)

心拍数の

時間波形

フーリエ
スペクトラム

位相図



・ 健康な人の心拍数のゆらぎ

「非線形ダイナミクスから
生成される不規則性」(カオス)

2. リカレントネットのダイナミクスと学習則

Nonlinear Dynamic Systems

$$\frac{dX_i(t)}{dt} = F_i(X(t)) \quad F_i(X(t)) : \text{nonlinear function}$$

Nonlinear function = Linear sum of sigmoid function

Miyake & Irie
Funahashi

$$F_i(X) \approx \sum_a W A_{ia} \cdot G(\sum_j W B_{aj} \cdot X_j + W D_a)$$

Equation for Recurrent Network

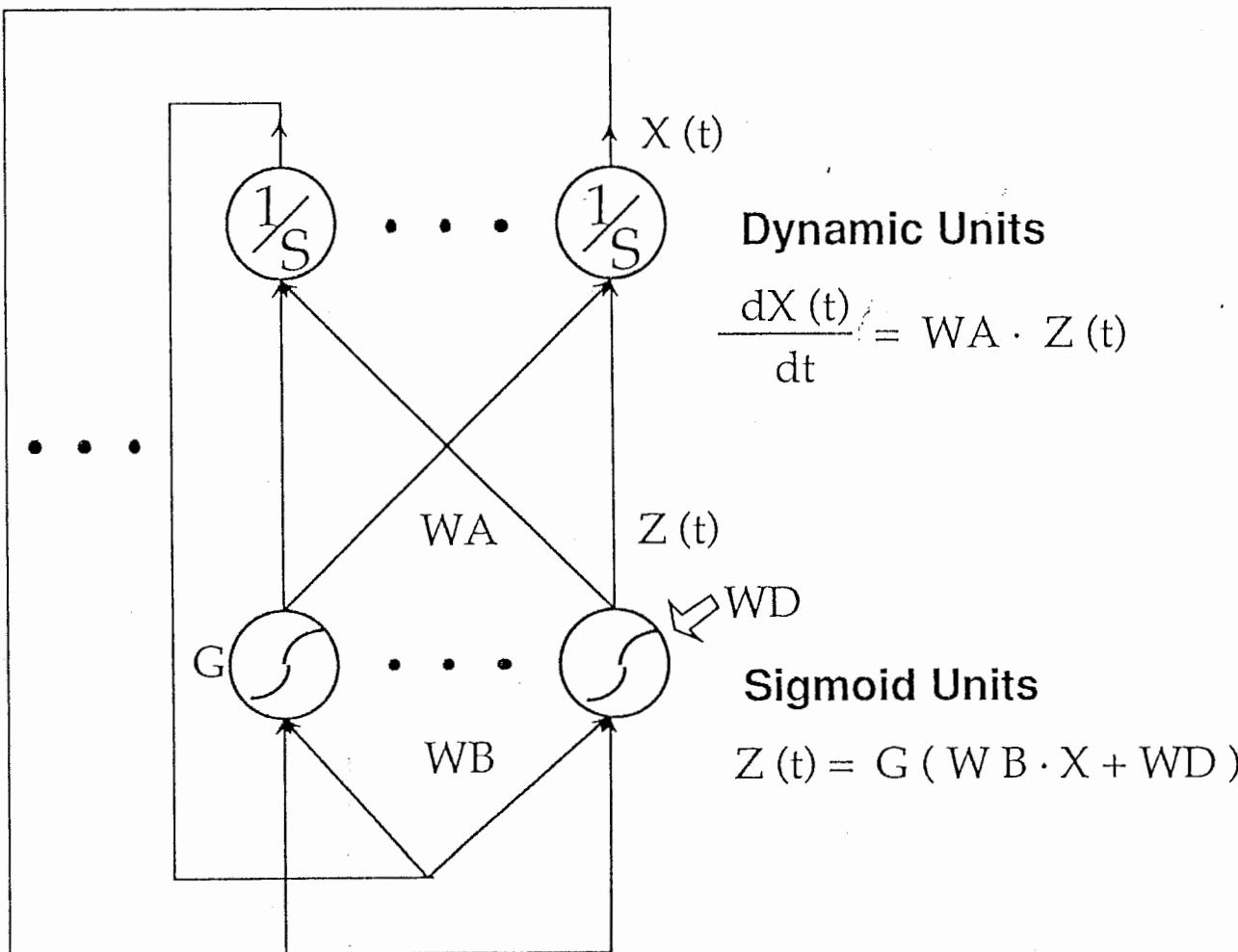
$$\frac{dX_i(t)}{dt} = \sum_a W A_{ia} \cdot Z_a(t)$$

$$Z_a(t) = G(\sum_j W B_{aj} \cdot X_j + W D_a)$$

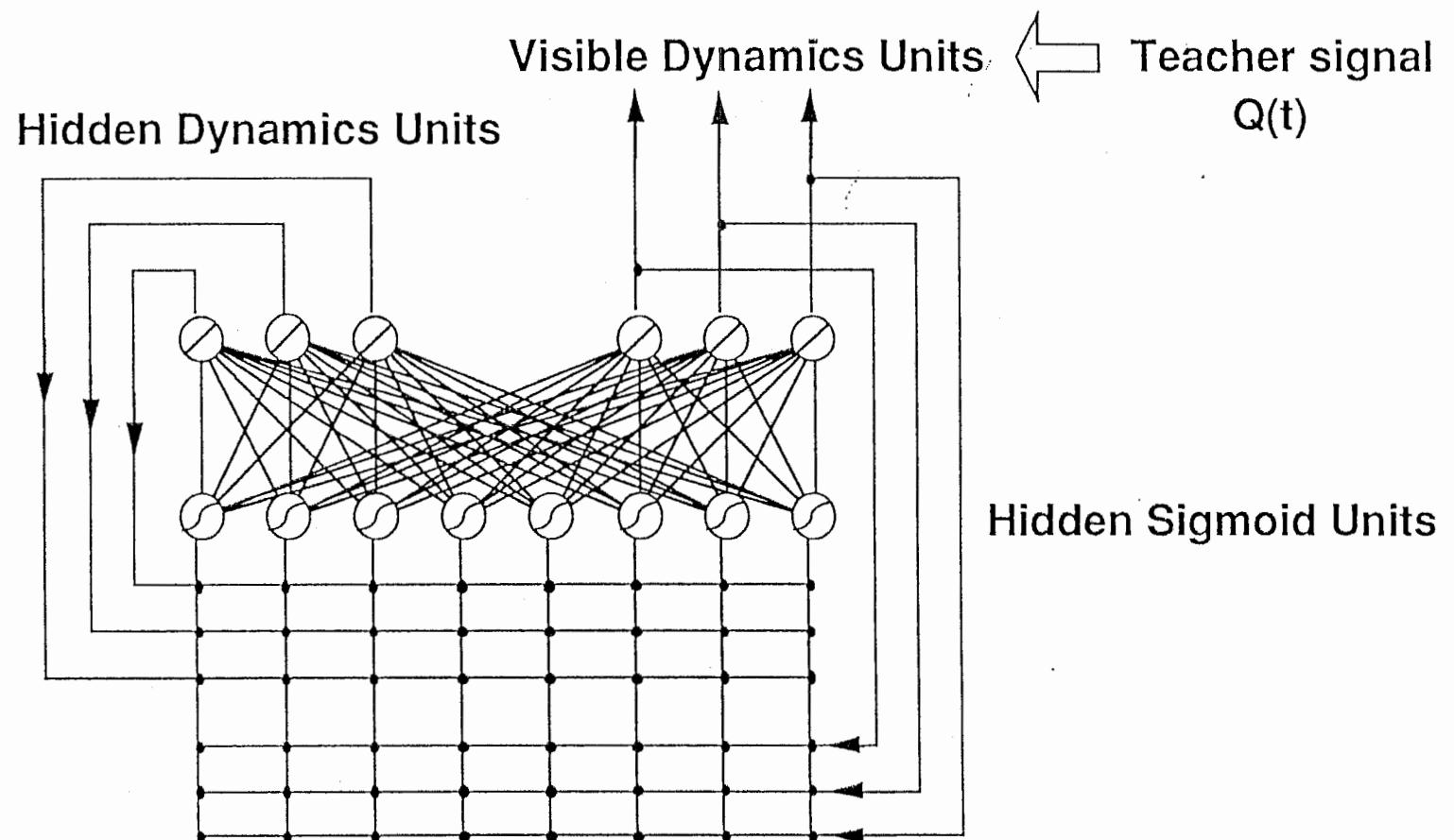


ATR Auditory and Visual Perception Research Laboratories

Recurrent Neural Network

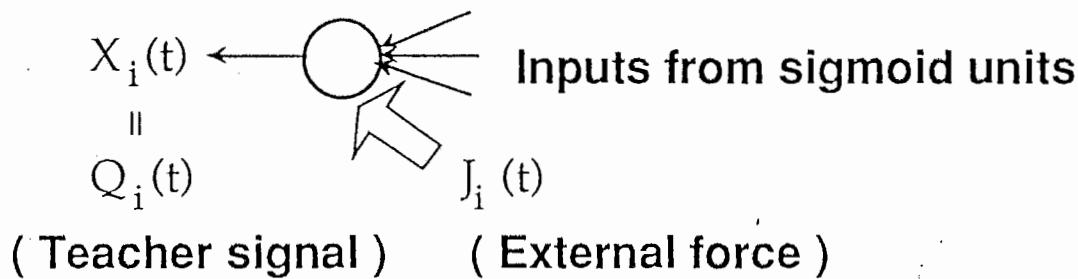


Supervised Learning for the Recurrent Network



ATR Auditory and Visual Perception Research Laboratories

Visible Dynamics Units



$$dX_i/dt = \sum_a W A_{ia} \cdot Z_a + J_i ; \text{ Imposing } X_i(t) = Q_i(t)$$

External Force

$$J_i = dQ_i/dt - \sum_a W A_{ia} \cdot Z_a$$

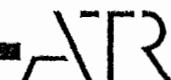
Error Function

$$E = \frac{1}{2} \int_T^{T+TF} dt \cdot \sum_{i \in V} J_i^2(t)$$

Weight Updates

$$\Delta W A_{im} \propto - \frac{\partial E}{\partial W A_{im}}$$

$$\Delta W B_{mi} \propto - \frac{\partial E}{\partial W B_{mi}}$$



② Calculation of Error Gradient

$$E = \frac{1}{2} \int_{t_i}^{t_f} dt \sum_{i \in V} J_i^2(t)$$

$$= \frac{1}{2} \int_{t_i}^{t_f} dt \sum_{i \in V} [\dot{Q}_i - \sum_a W_{ia} Z_a]$$

$$- \int_{t_i}^{t_f} dt \sum_{i \notin V} P X_i [\dot{X}_i - \sum_a W_{ia} Z_a]$$

$$- \int_{t_i}^{t_f} dt \sum_a P Z_a [Z_a - G(\sum_i T_{ai} X_i + B_a)]$$

- Weight Variation

$$W \rightarrow W + \delta W, T \rightarrow T + \delta T$$

$$\delta X_i(t) \equiv X_i(t; W + \delta W) - X_i(t; W)$$

- Variation of Error

$$\begin{aligned} \delta E = \int_{t_i}^{t_f} dt & [(\dots) \delta P X + (\dots) \delta P Z \\ & + (\dots) \delta X + (\dots) \delta Z \\ & + (\dots) \delta W + (\dots) \delta T] \end{aligned}$$

+ Boundary Terms

◦ Forward Run ($t_i \rightarrow t_f$)

$$\begin{cases} \dot{X}(t) = W \cdot Z(t) \\ Z(t) = G(T \cdot X(t) + B) \end{cases}$$

◦ Backward Run ($t_f \rightarrow t_i$)

$$\cdot P X(t_f) = 0$$

$$\cdot \dot{P} X(t) = - P Z(t)^T G'(T \cdot X(t) + B) \cdot T$$

$$\cdot P Z(t) = P(t)^T \cdot W$$

$$P(t) = \begin{cases} -J(t) & \text{for Visible Units} \\ P X(t) & \text{for Hidden Units} \end{cases}$$

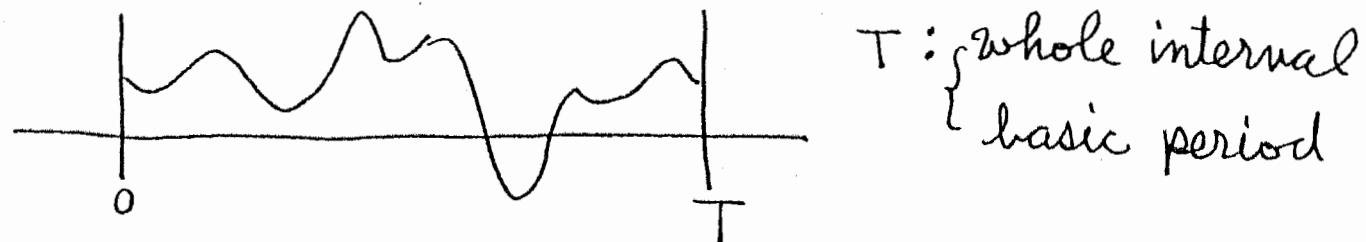
◦ Weight Change

$$\Delta W = -\eta \cdot \partial E / \partial W = \int_{t_i}^{t_f} dt P(t) Z(t)$$

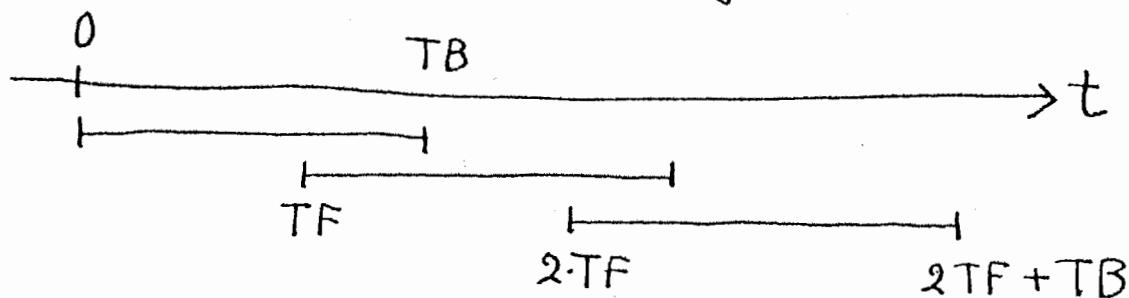
$$\Delta T = -\eta \cdot \partial E / \partial T = \int_{t_i}^{t_f} dt P Z(t) \cdot G'(T \cdot X(t) + B) X(t)$$

Learning Interval

- Periodic signal, finite length signal



- Unknown period, indefinite length signal, chaos
moving small learning interval



- arbitrary waveform NO
- signal generated by nonlinear dynamics ... OK
Time development rule is the same for each interval

Initial Value

- Visible unit's initial value : known



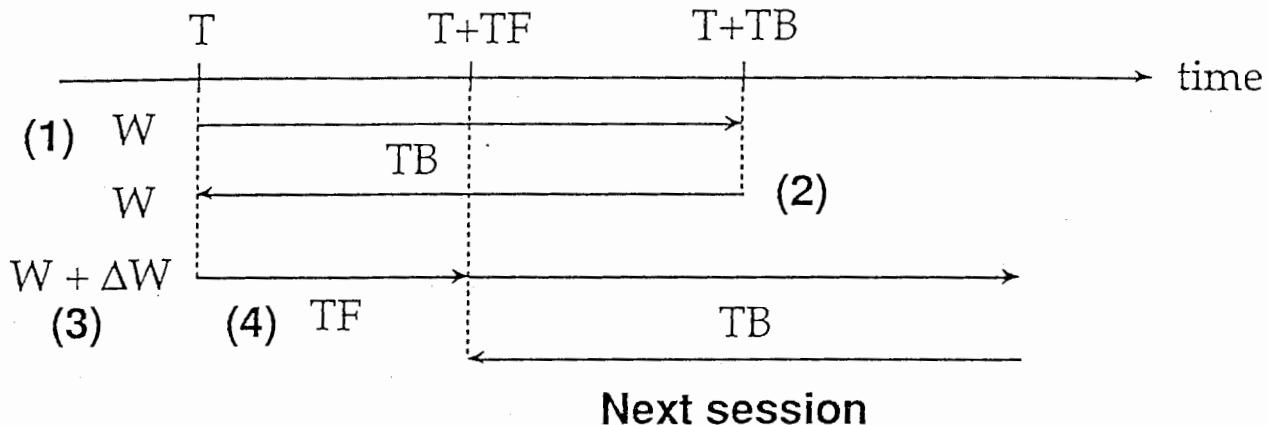
Initial value is clamped to the teacher signal.

- Hidden unit's initial value : Unknown



Initial value can be considered as learning parameters
and they are adjusted for each learning interval.

Learning Schedule



(1) Forward Run $t: T \rightarrow (T + TB)$

$$\frac{dX_i}{dt} = \sum_m WA_{im} \cdot G(Z_m) \quad (i \notin V) \quad X_i = Q_i \quad (i \in V)$$

(2) Error Calculation $t: (T + TB) \rightarrow T$

$$\frac{dPX_i}{dt} = \sum_m PZ_m \cdot WA_{im} \cdot G'(Z_m)$$

$$PZ_m = \sum_i WB_{mi} \cdot PX_i, \quad PX_i(T + TB) = 0 \quad (i \notin V)$$

$$PX_i = -J_i = \frac{dQ_i}{dt} + \sum_m WA_{im} \cdot G'(Z_m) \quad (i \in V)$$

(3) Weight Updates

$$\Delta WA_{im} = -\alpha \int_T^{T+TB} PX_i \cdot Z_m \quad \Delta WB_{mi} = -\alpha \int_T^{T+TB} PZ_m \cdot X_i \cdot G'(Z_m)$$

$$\Delta X_i(T) = -\beta \cdot PX_i(T)$$

(4) Forward Run with New Parameter

run for $t: T \rightarrow T + TF$ time update $T = T + TF$



ATR Auditory and Visual Perception Research Laboratories

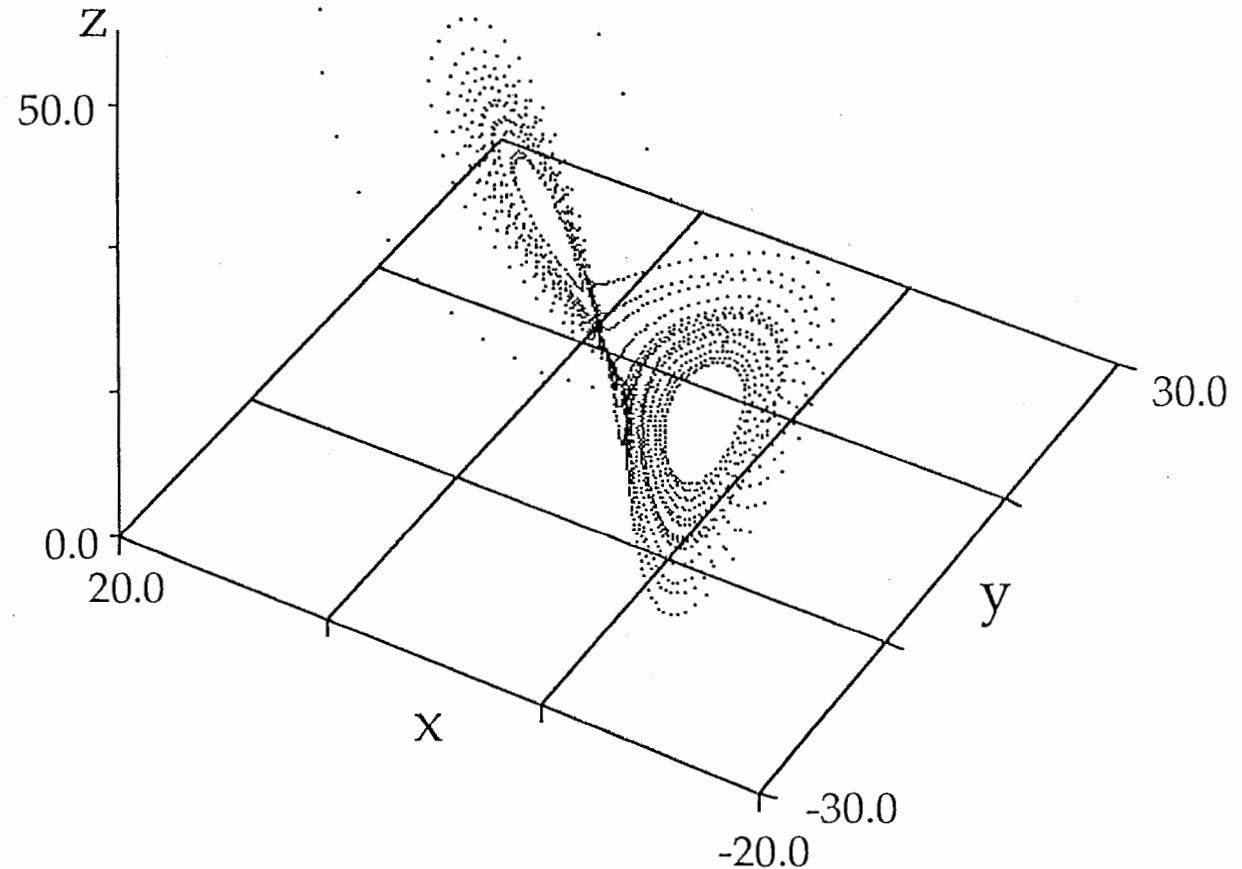
3. 力オースダイナミクスの学習

Lorentz Attractor

$$\frac{dx}{dt} = F_1(x, y, z) \\ = 10 \cdot (y - x)$$

$$\frac{dy}{dt} = F_2(x, y, z) \\ = -y + (28 - z) \cdot x$$

$$\frac{dz}{dt} = F_3(x, y, z) \\ = -\frac{8}{3} \cdot z + x \cdot y$$



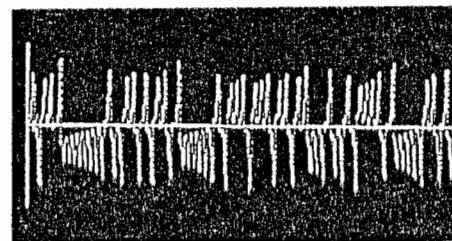
ATR

ATR Auditory and Visual Perception Research Laboratories

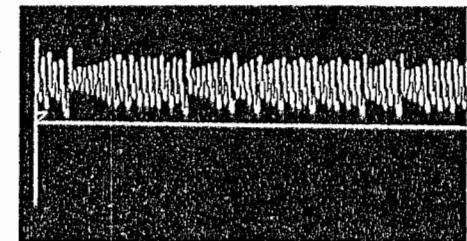
Temporal Behavior



$x(t)$

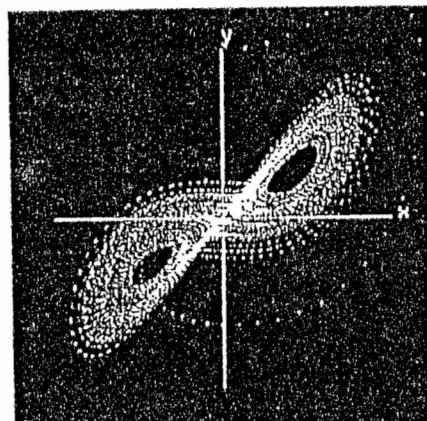


$y(t)$

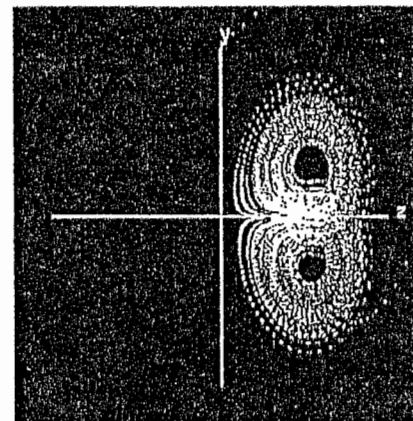


$z(t)$

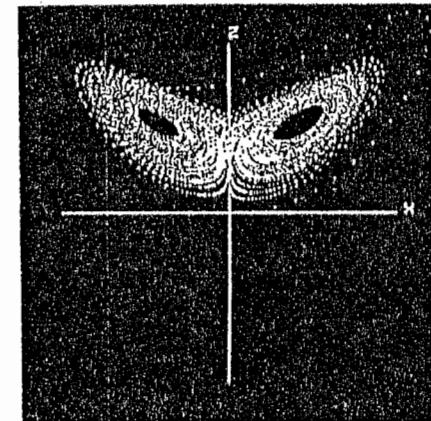
Phase Diagram



x - y plane

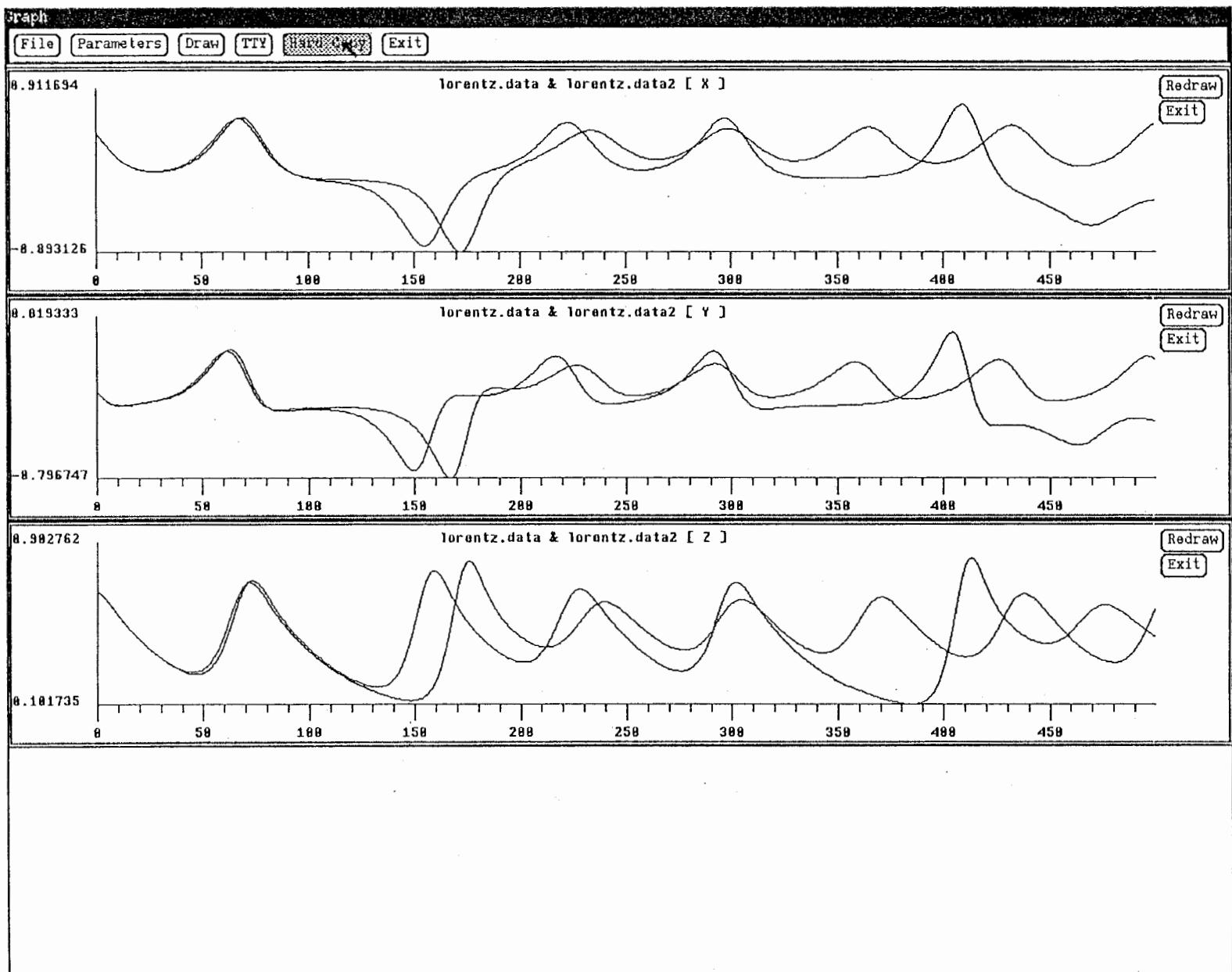


y - z plane



z - x plane





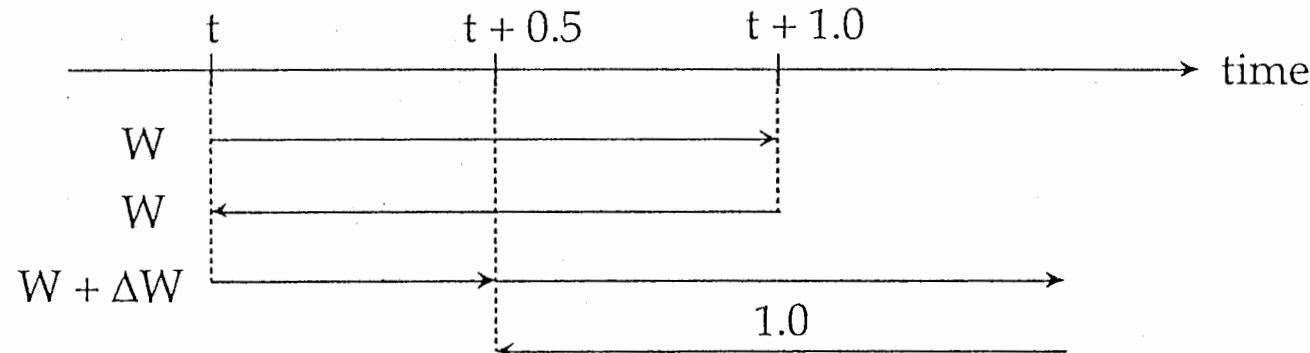
Simulation Situation

3 Dynamic Units

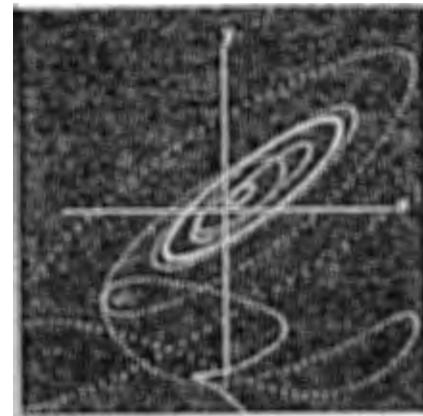
30 Sigmoid Units

2nd order Runge-Kutta method

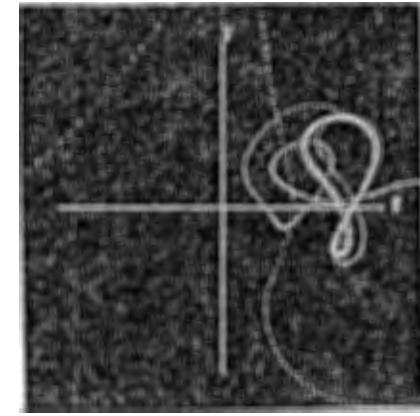
$\Delta t = 0.01$



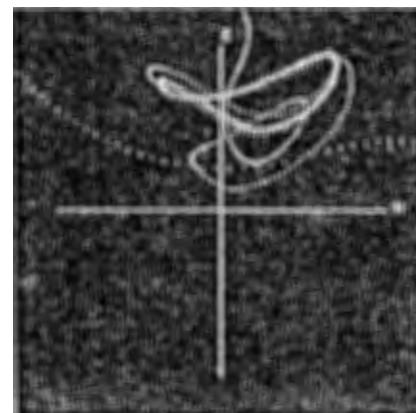
すべての変数に教師信号を与えた場合



x-y plane



y-z plane



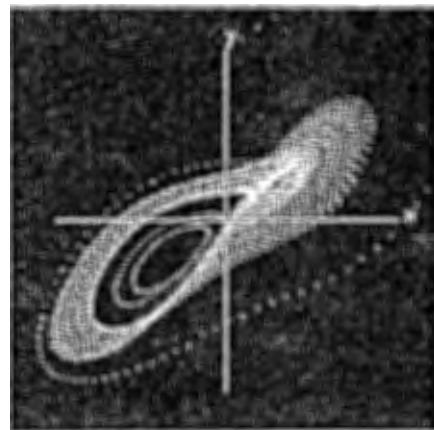
z-x plane

**$x(t)$, $y(t)$, and $z(t)$ are visible.
After 1,000 weight updates**

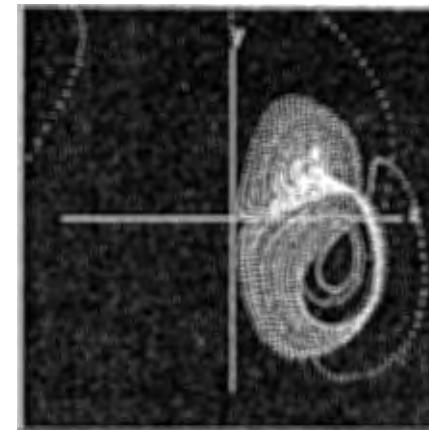


ATR Auditory and Visual Perception Research Laboratories

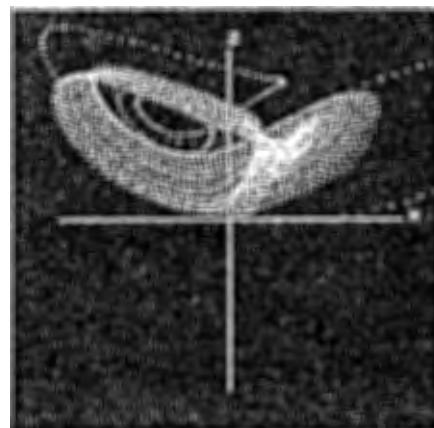
すべての変数に教師信号を与えた場合



x-y plane



y-z plane

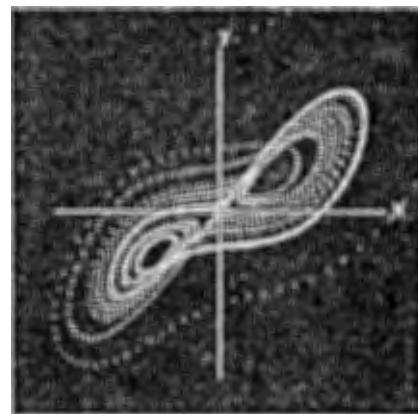


z-x plane

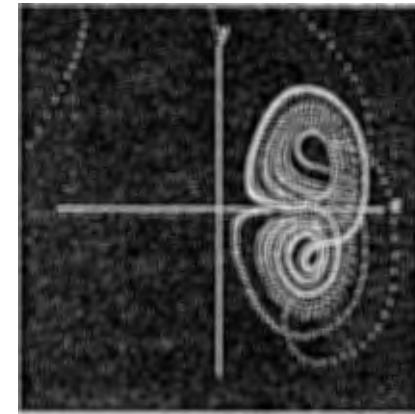
$x(t)$, $y(t)$, and $z(t)$ are visible.
After 2,000 weight updates



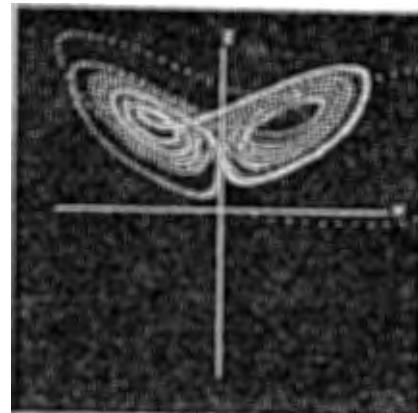
すべての変数に教師信号を与えた場合



x-y plane



y-z plane



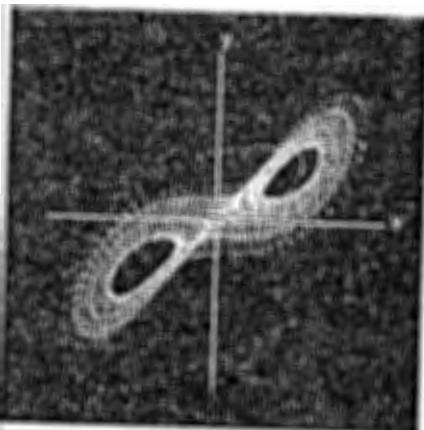
z-x plane

$x(t)$, $y(t)$, and $z(t)$ are visible.
After 4,000 weight updates

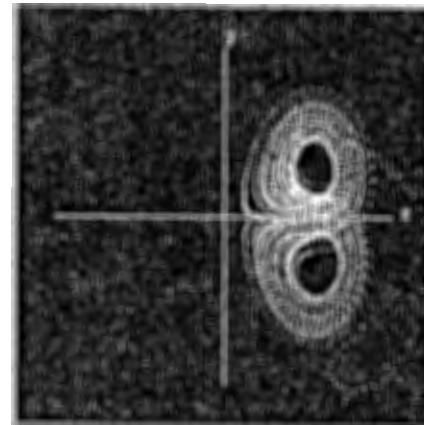


ATR Auditory and Visual Perception Research Laboratories

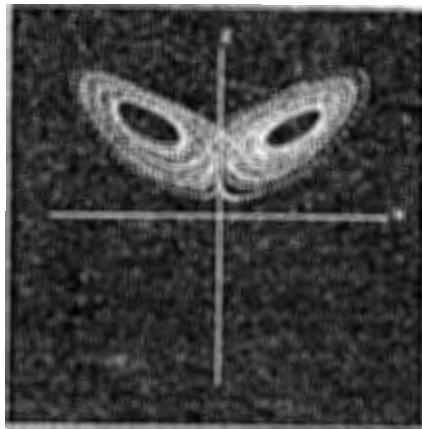
すべての変数に教師信号を与えた場合



x-y plane



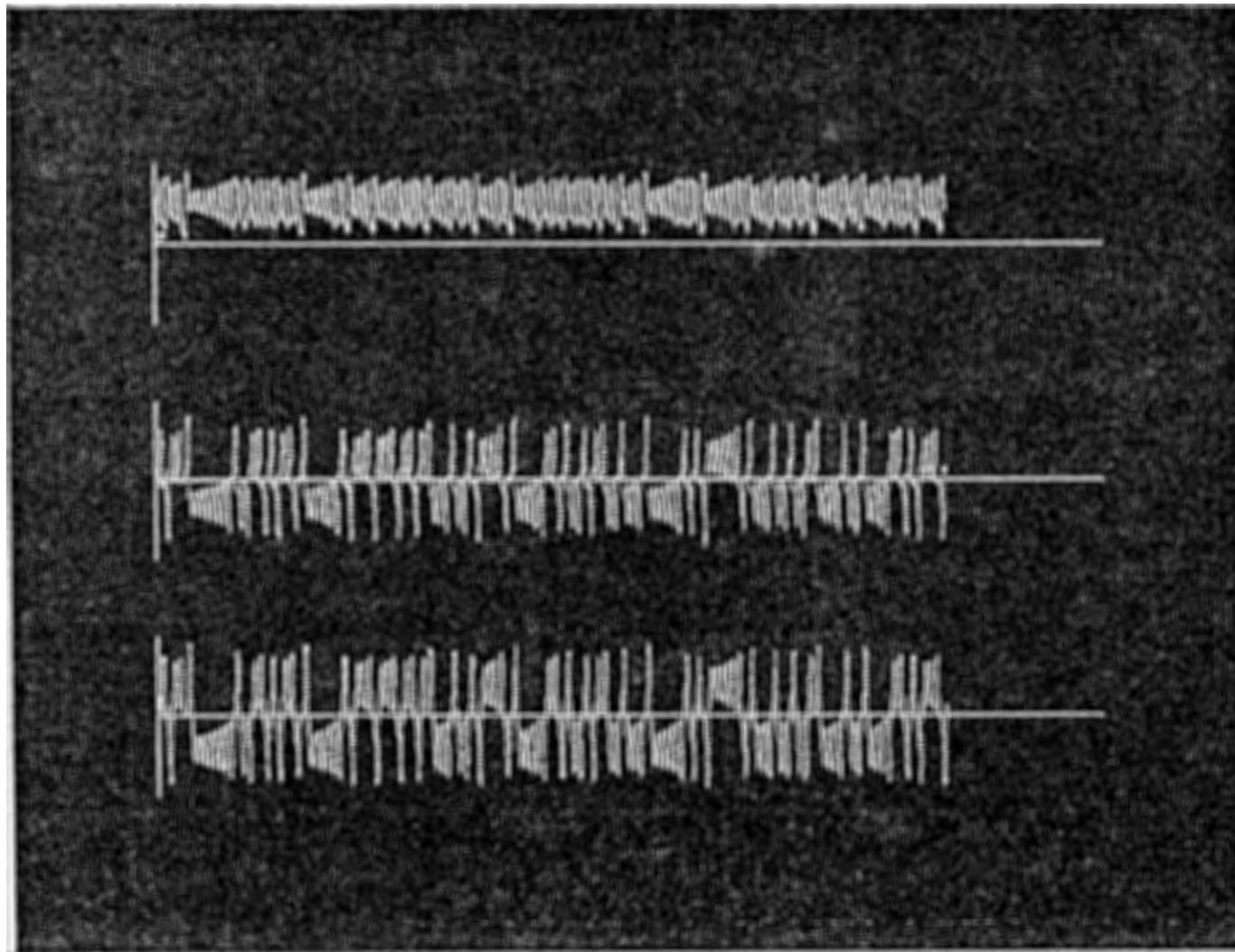
y-z plane



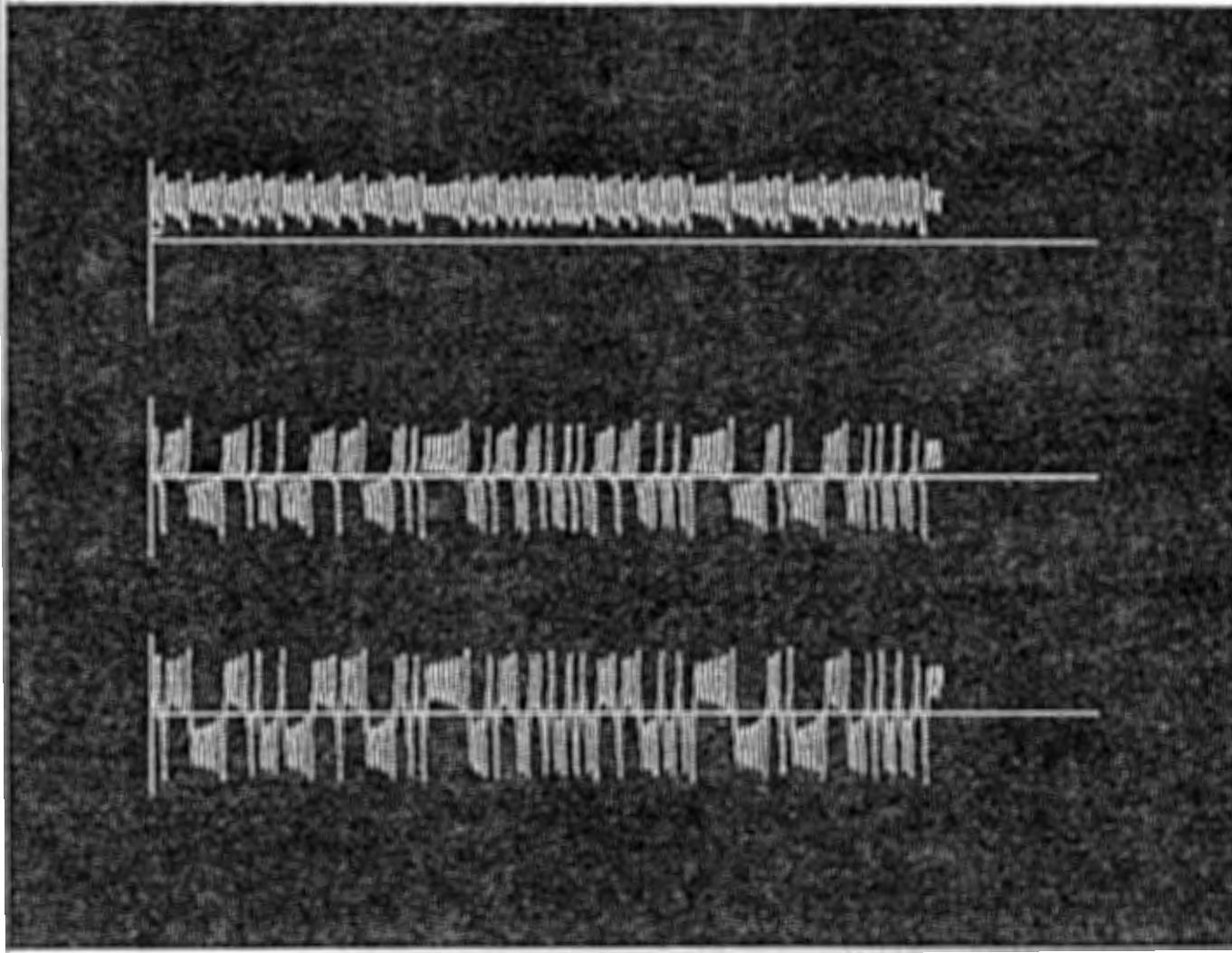
z-x plane

**$x(t)$, $y(t)$, and $z(t)$ are visible.
After 30,000 weight updates**

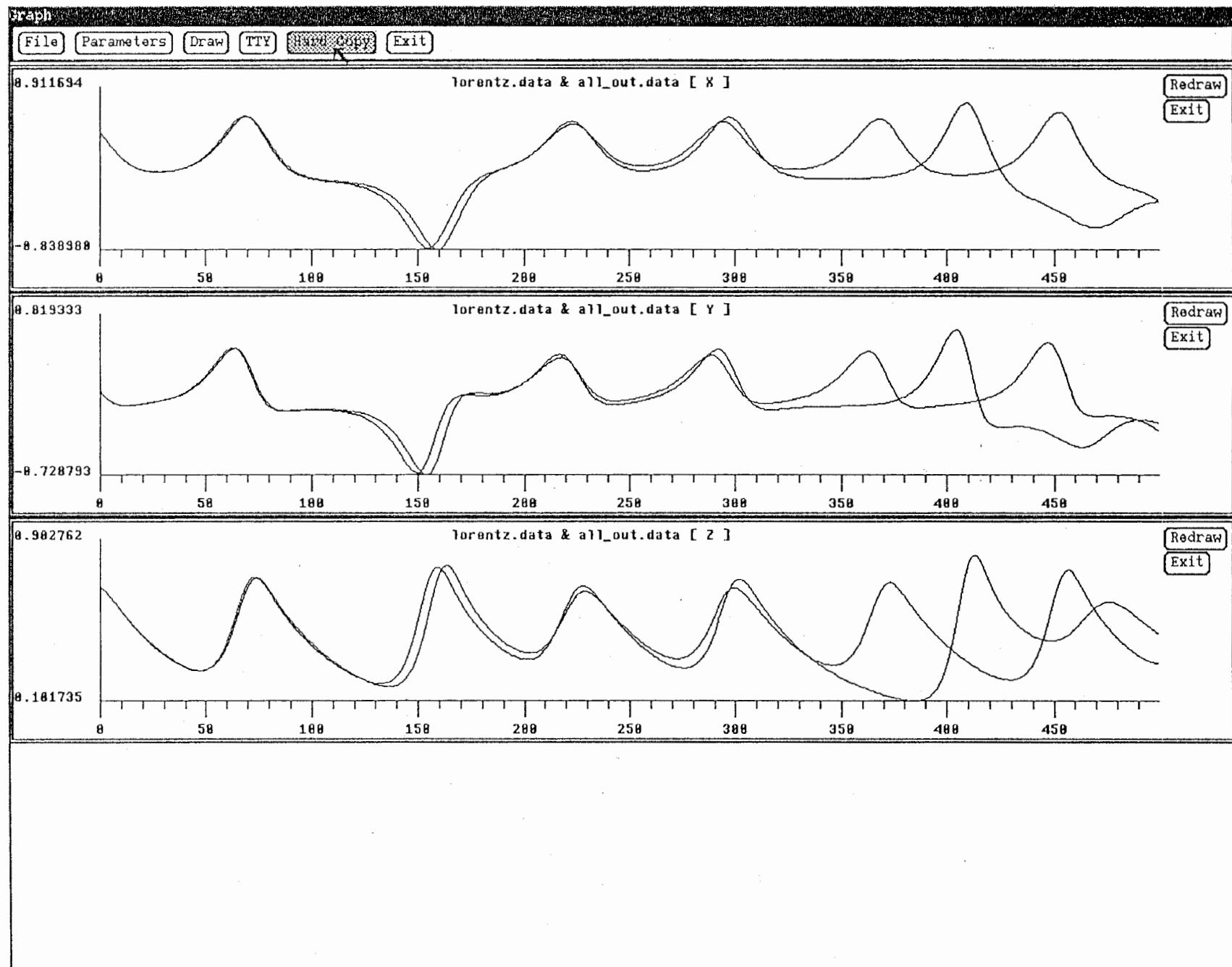




Teacher signal (Lorentz attractor)



Generated trajectory



Liapunov Exponents

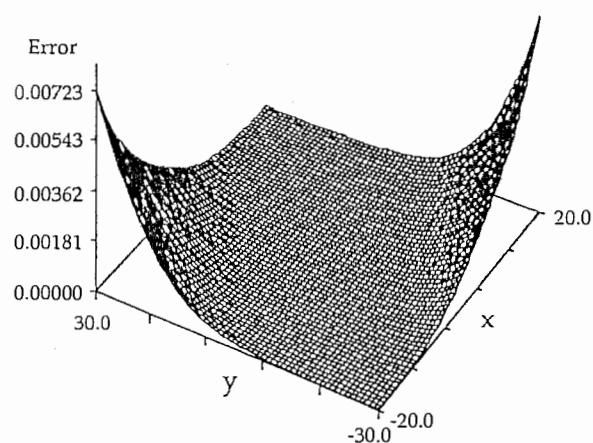
Lorentz Attractor	0.90
All Visible Dynamic Units	0.85
One hidden Dynamic Unit	0.76



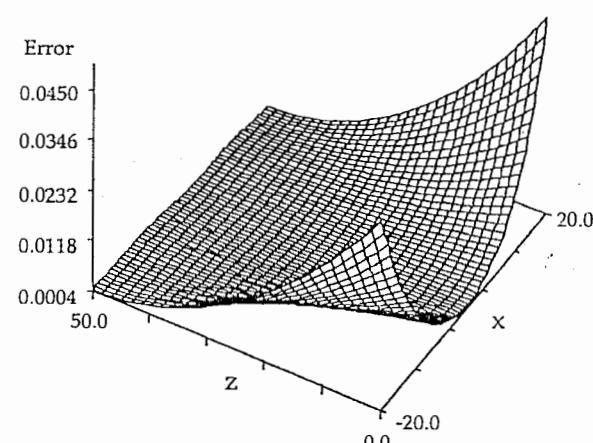
ATR Auditory and Visual Perception Research Laboratories

Error for the Vector Field

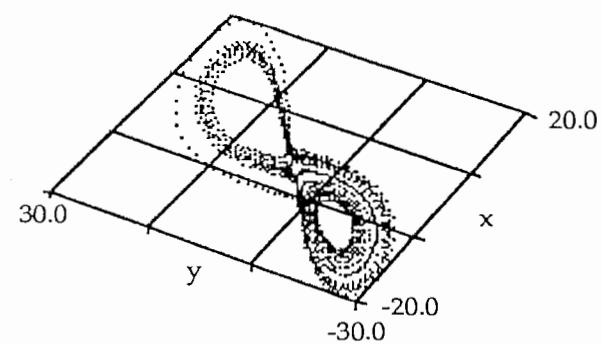
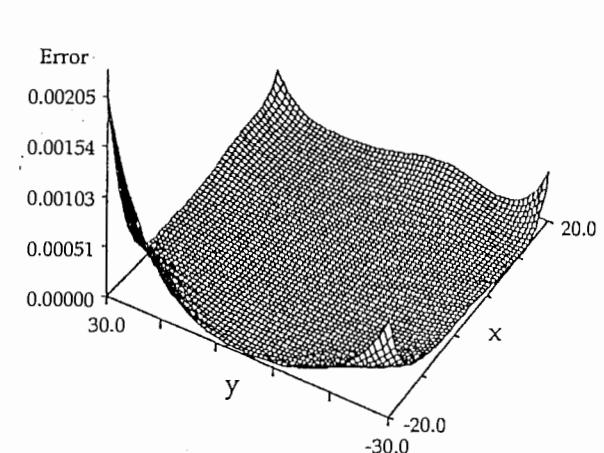
Error for $F_1(x, y, z)$



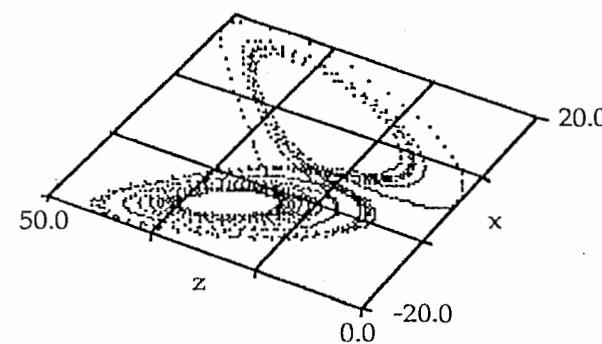
Error for $F_2(x, y, z)$



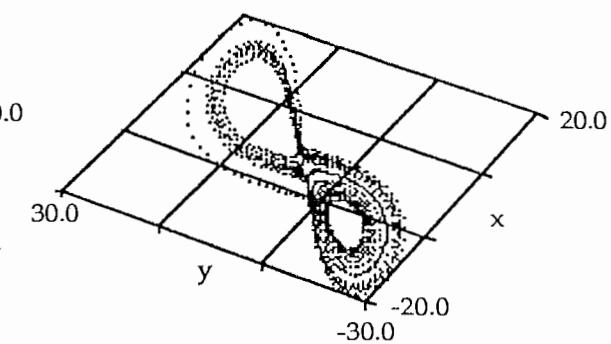
Error for $F_3(x, y, z)$



Error average over z



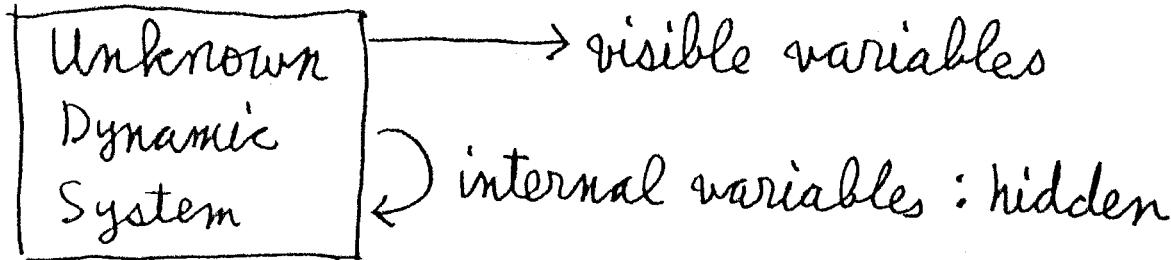
Error average over y



Error average over z



Learning Hidden Dynamics



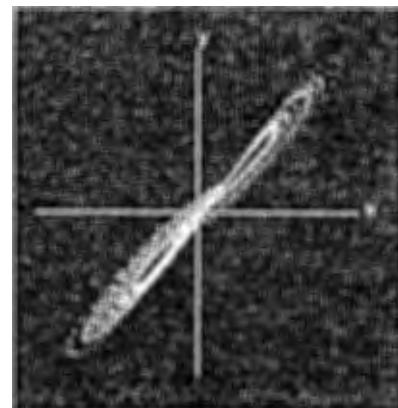
- Simulation { visible variables : y & z
hidden variable : x
- Ambiguity due to coordinate transformation
 - $\frac{dx}{dt} = F_1(x, y, z)$, $\frac{dy}{dt} = F_2(x, y, z)$, $\frac{dz}{dt} = F_3(x, y, z)$
 - coordinate trans. $x = h(x', y, z)$
: visible variable trajectories are unchanged

$$\frac{dx'}{dt} = \left[F_1(x, y, z) - \frac{\partial h}{\partial y} \cdot F_2(h(x', y, z), y, z) - \frac{\partial h}{\partial z} \cdot F_3(h(x', y, z), y, z) \right] / \frac{\partial h}{\partial x'}$$

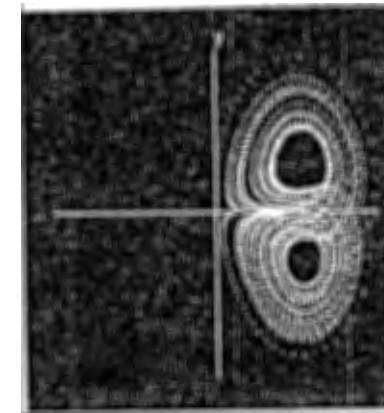
$$\frac{dy}{dt} = F_2(h(x', y, z), y, z)$$

$$\frac{dz}{dt} = F_3(h(x', y, z), y, z)$$

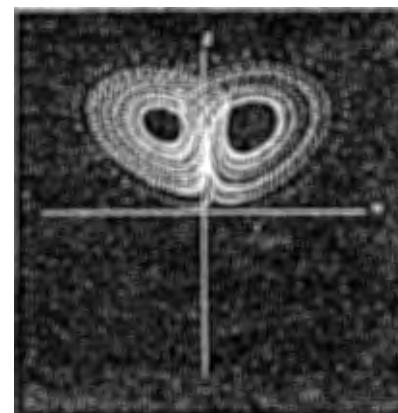
隠れたダイナミクスの学習



x-y plane



y-z plane



z-x plane

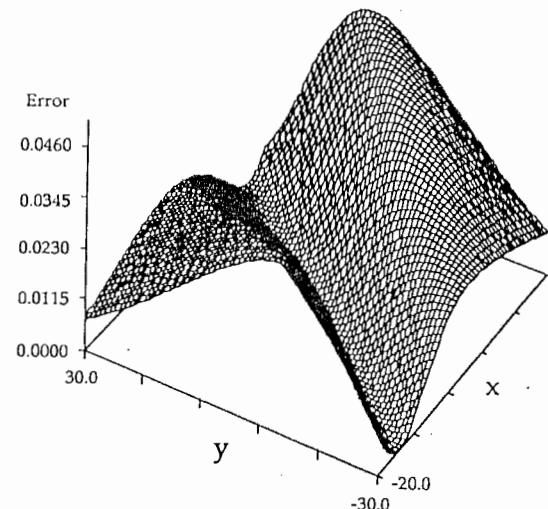
**y(t) and z(t) are visible.
x(t) is hidden.
50,000 weight updates**



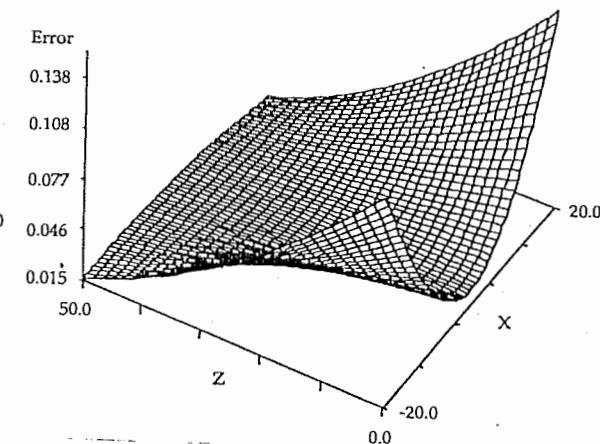
ATR Auditory and Visual Perception Research Laboratories

Error for the Vector Field

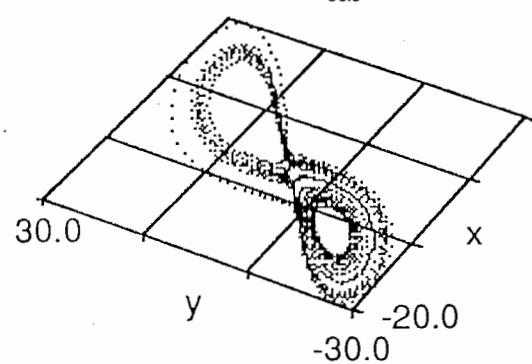
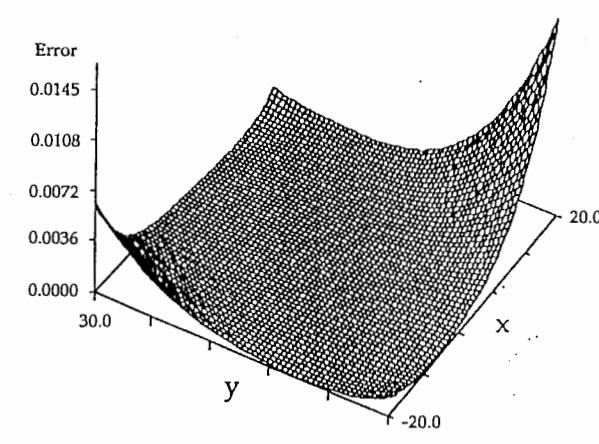
Error for $F_1(x, y, z)$



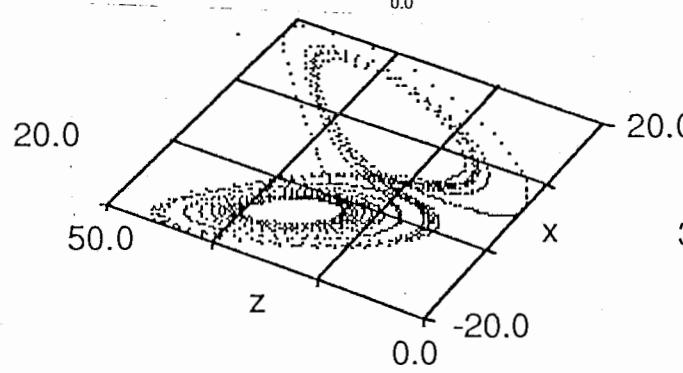
Error for $F_2(x, y, z)$



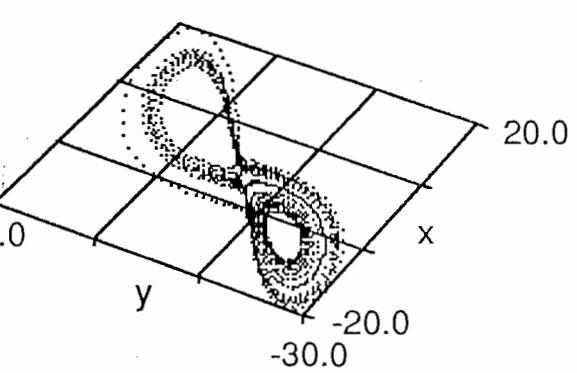
Error for $F_3(x, y, z)$



Error average over z



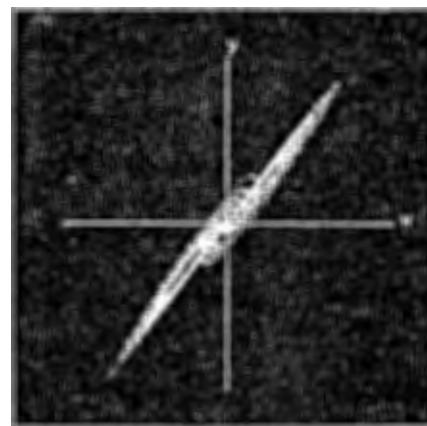
Error average over y



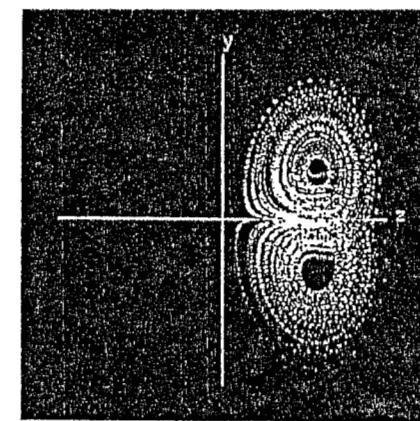
Error average over z



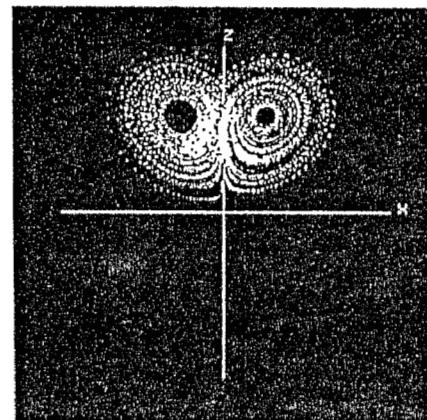
変数変換をしたローレンツアトラクタ



x'-y' plane



y'-z' plane



z'-x' plane

Coordinate Transformation

$$x' = x + 2y$$

$$y' = y$$

$$z' = z$$

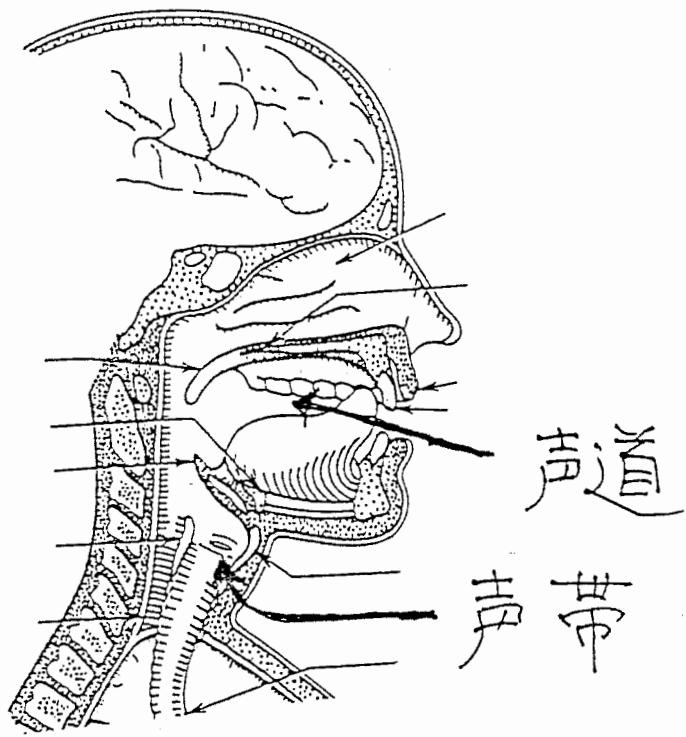


ATR Auditory and Visual Perception Research Laboratories

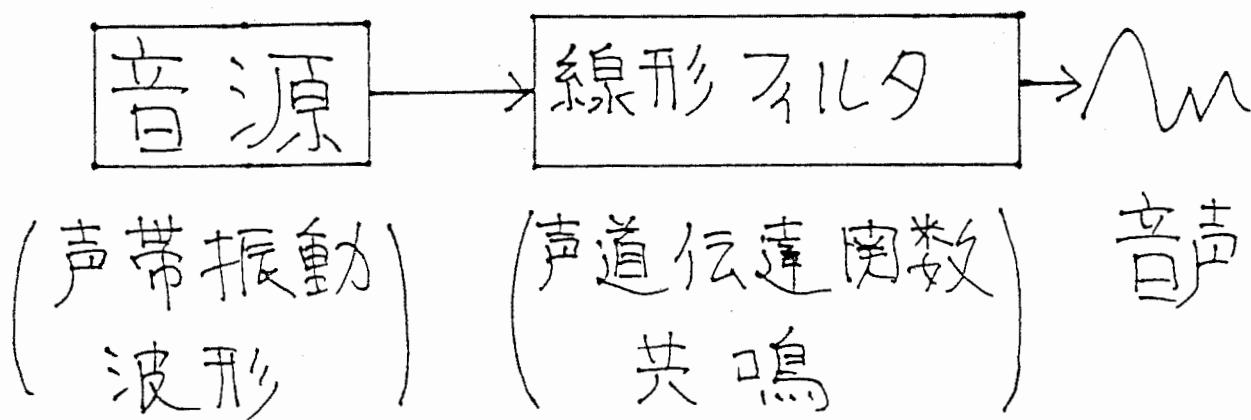
4. 音声ゆらぎの学習

音声ゆうき

・音声の発生機構

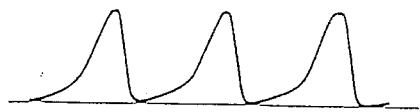


・人工音声合成（有聲音）



◦人工音源波形

周期波形



◦人間の音源波形

ゆらぎ



◦音源波形のゆらぎと音声の自然さ

(心理学実験: M. Hashiba et.al.)

1. 周期波形 → X

(人間の音源波形の1ビット分の繰返し)

2. 確率的ゆらぎ → X

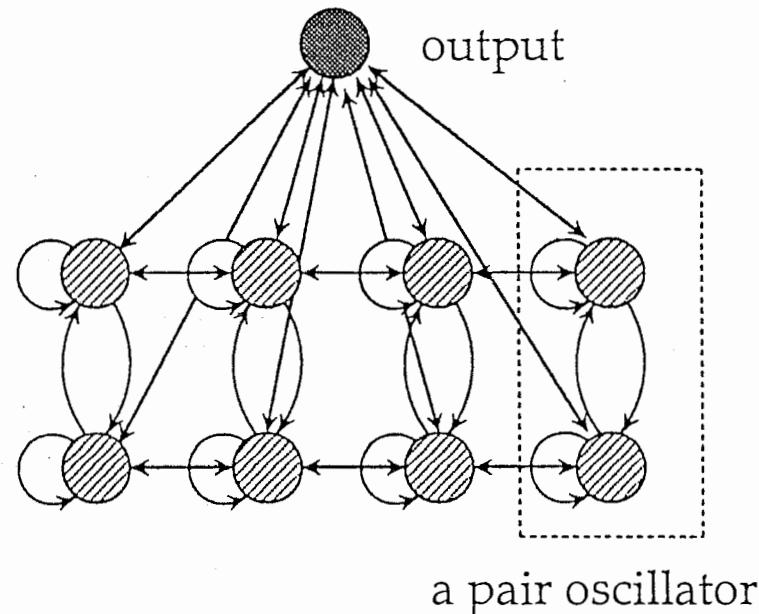
(人間の音源波形のゆらぎと
同じ統計的性質をもつ)

3. 自然なゆらぎ → ○

(少なくとも32ビット以上に
わたるゆらぎが必要)

APOLONN

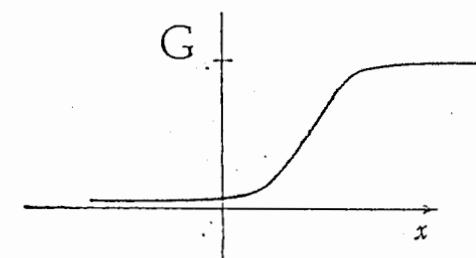
Adaptive nonlinear Pair Oscillators with Local coNNection



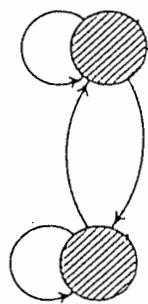
$$\tau_i \frac{dX_i}{dt} = -X_i + G \left(\sum_j W_{ij} \cdot X_j \right)$$

Each Pair has different time constant τ_i

Adjustable to different time scale components

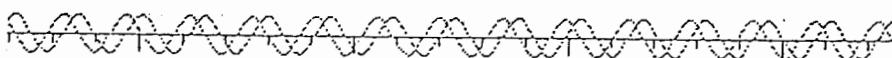


Pair oscillator

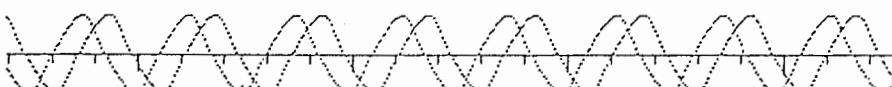


$$\tau_i \frac{dX_i}{dt} = -X_i + G \left(\sum_j W_{ij} \cdot X_j \right)$$

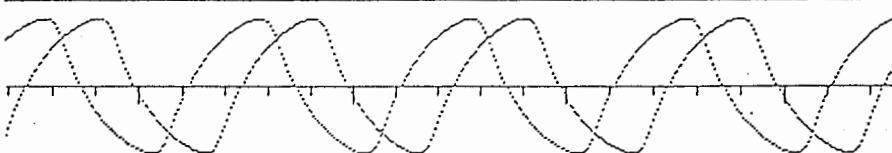
<<< WEIGHT VALUE >>>



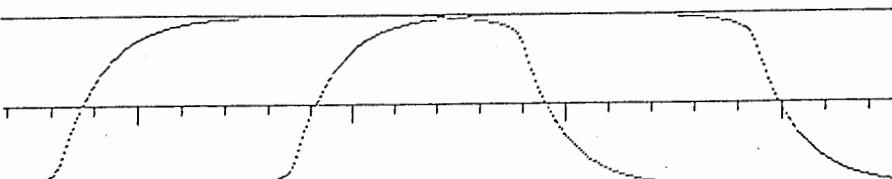
(0.4000)(3.1400)
(-3.1400)(0.4000)



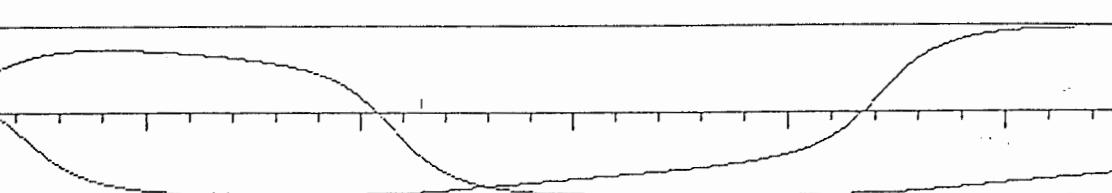
(1.0000)(3.1400)
(-3.1400)(1.0000)



(2.0000)(3.1400)
(-3.1400)(2.0000)



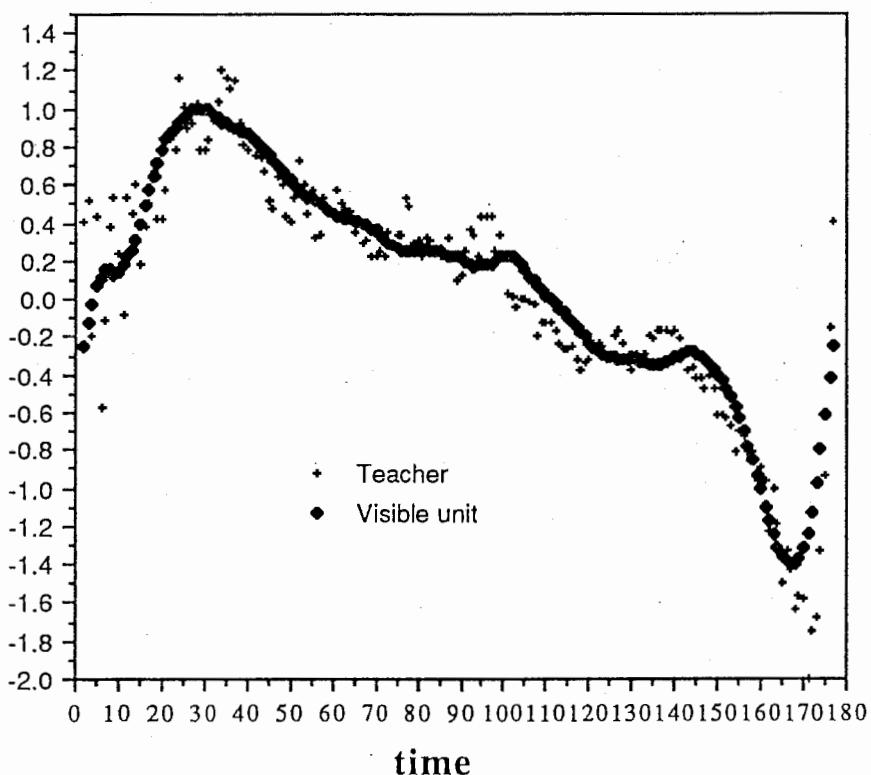
(4.0000)(3.1400)
(-3.1400)(4.0000)



(1.0000)(0.3140)
(-0.3140)(1.0000)

max[1.0000],min[-1.0000],gain[2.0000]delay[1.0000],tstep[0.0500]

Fig. 4 Trajectory generated by APOLO



(a) The trajectory of the visible unit and the teacher signal

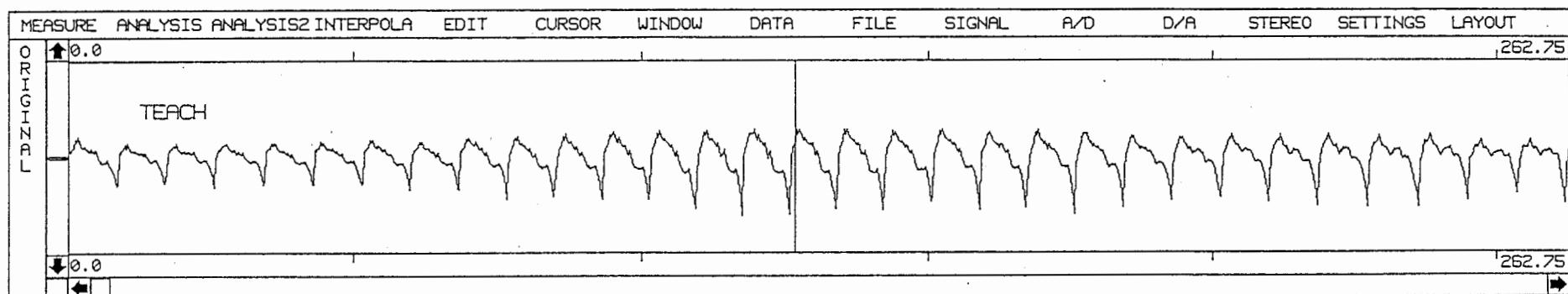
Simulation

Teacher Signal

32 pitch period waveform in natural voice / a /

(residual signal obtained by the 16th order inverse filtering)

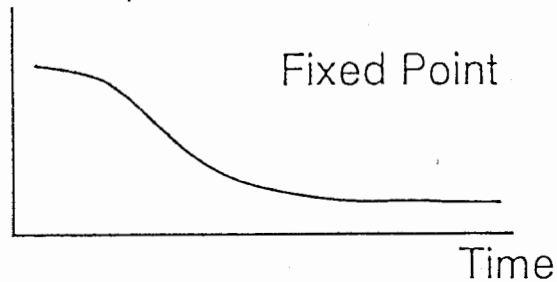
Sampling rate : 4 kHz 16 bits A.D.



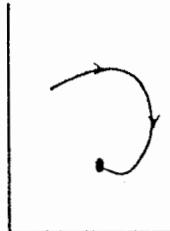
Natural voice source waveform

Bifurcations in Learning

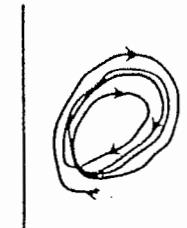
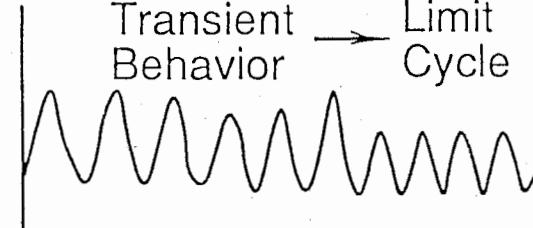
Unit Output



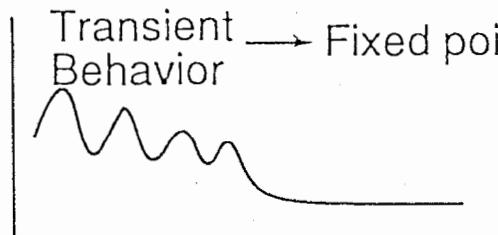
Phase Diagram



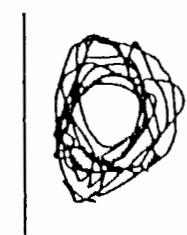
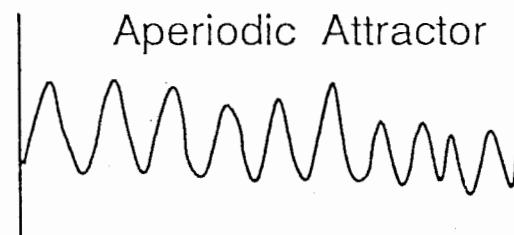
Transient Behavior → Limit Cycle



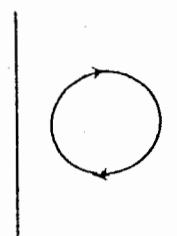
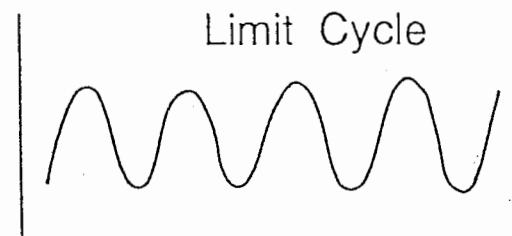
Transient Behavior → Fixed point



Aperiodic Attractor



Limit Cycle



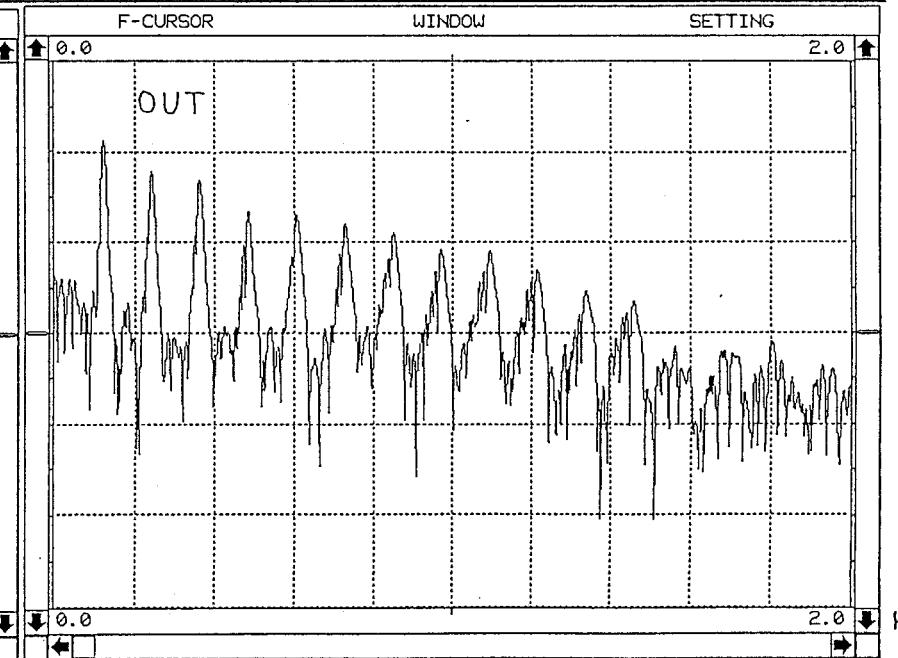
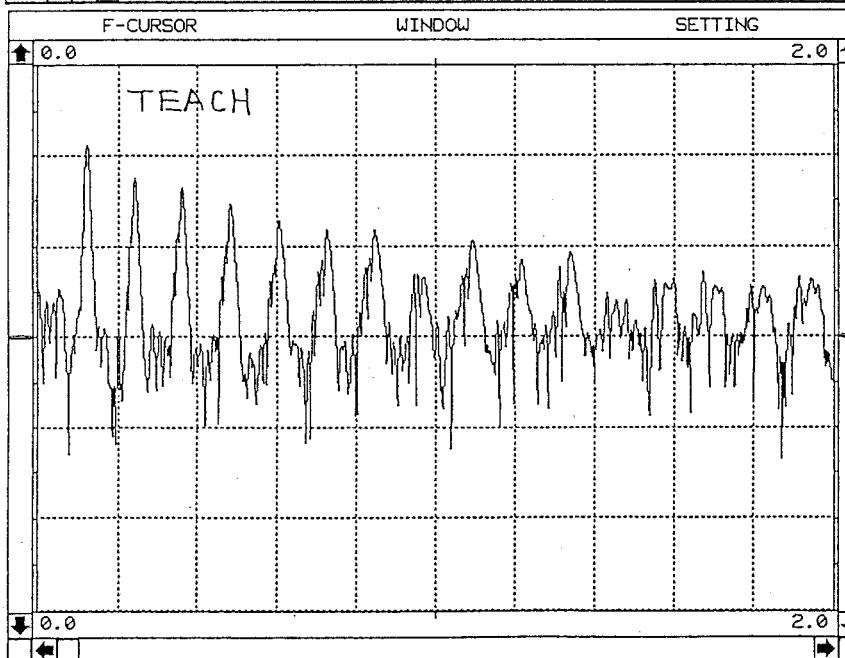
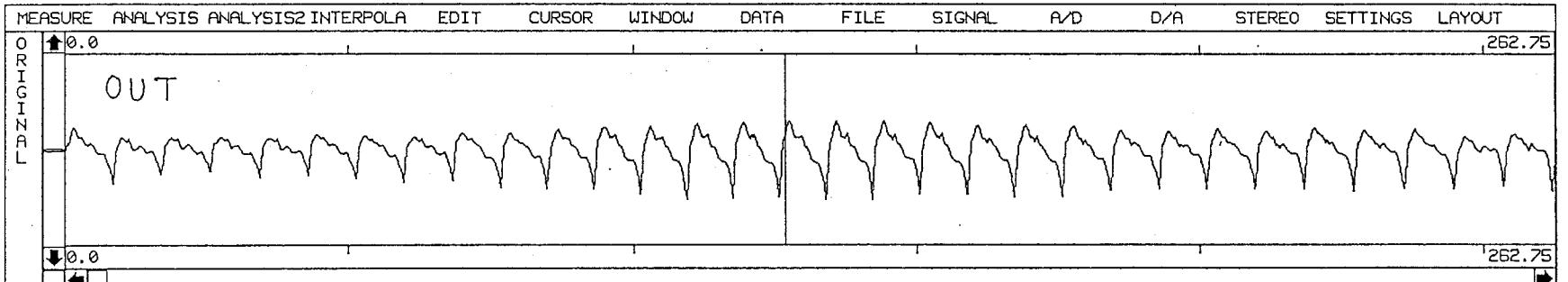
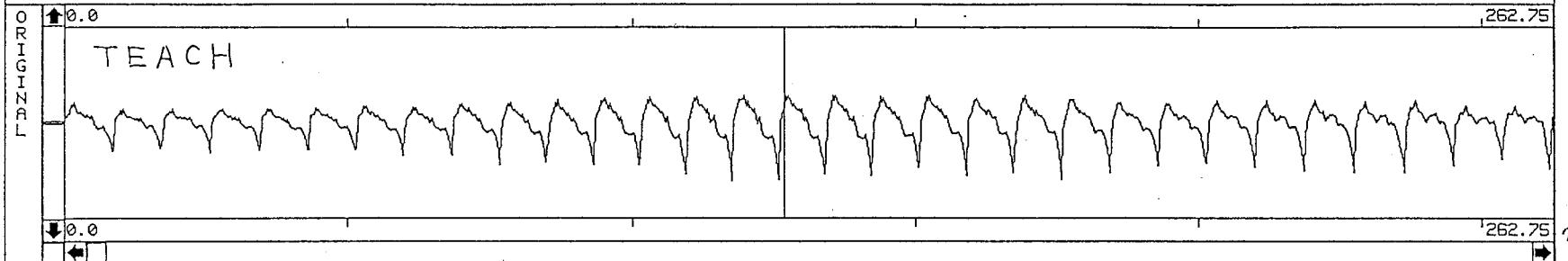
ATR

ATR Auditory and Visual Perception Research Laboratories

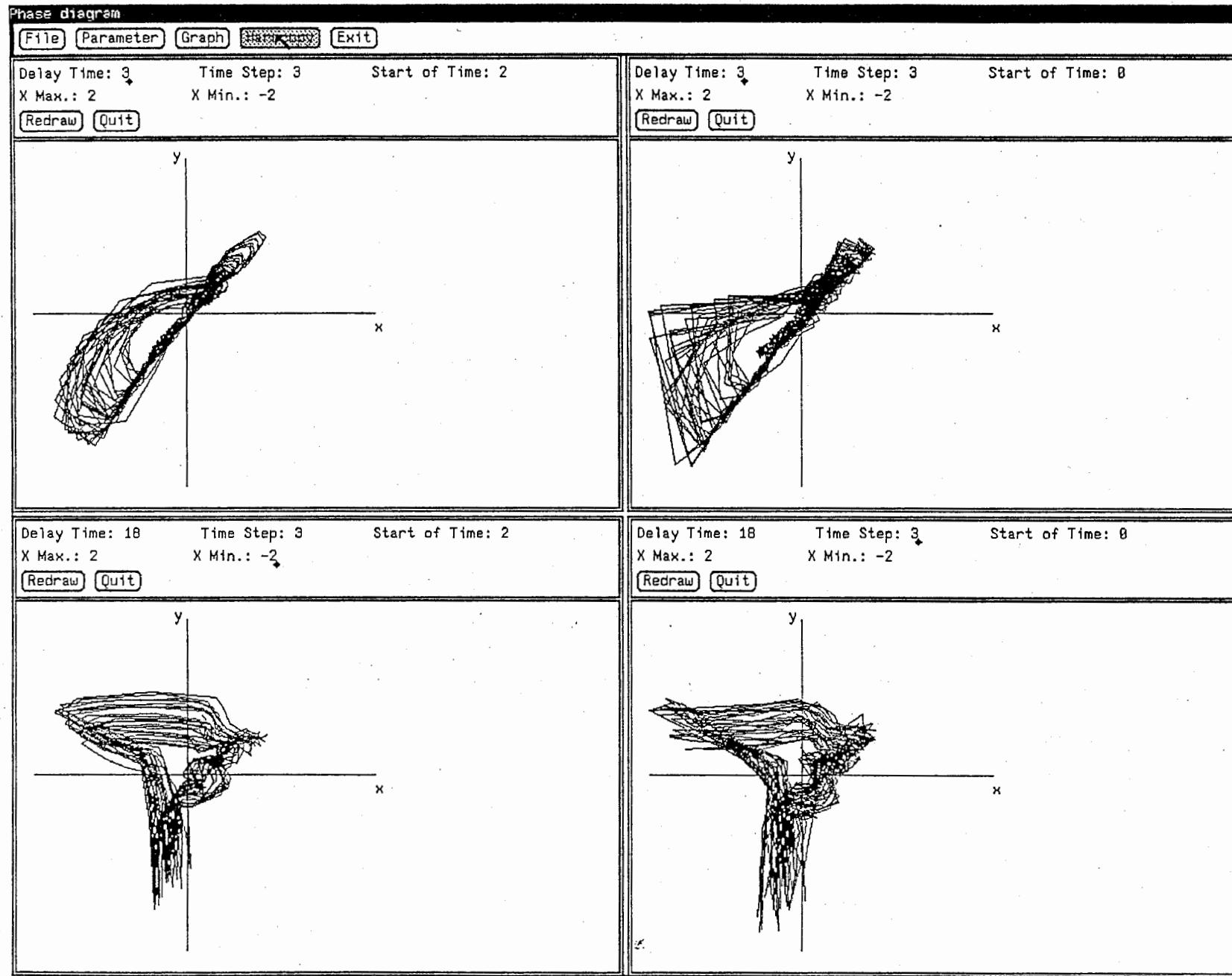
MANAGER SHELL LAYOUT HARDCOPY OPTIONS

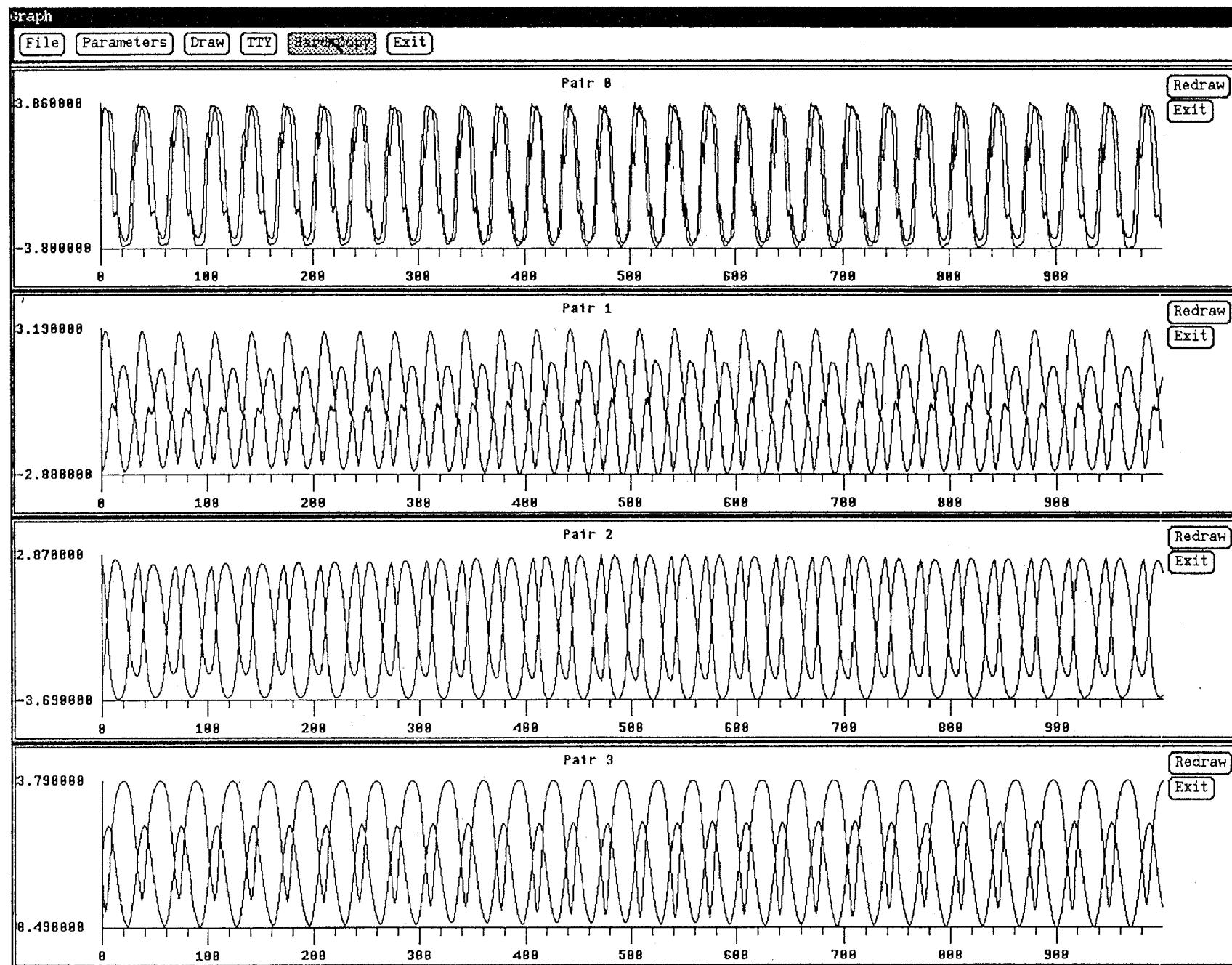
Thu Oct 4 17:36:26 1990

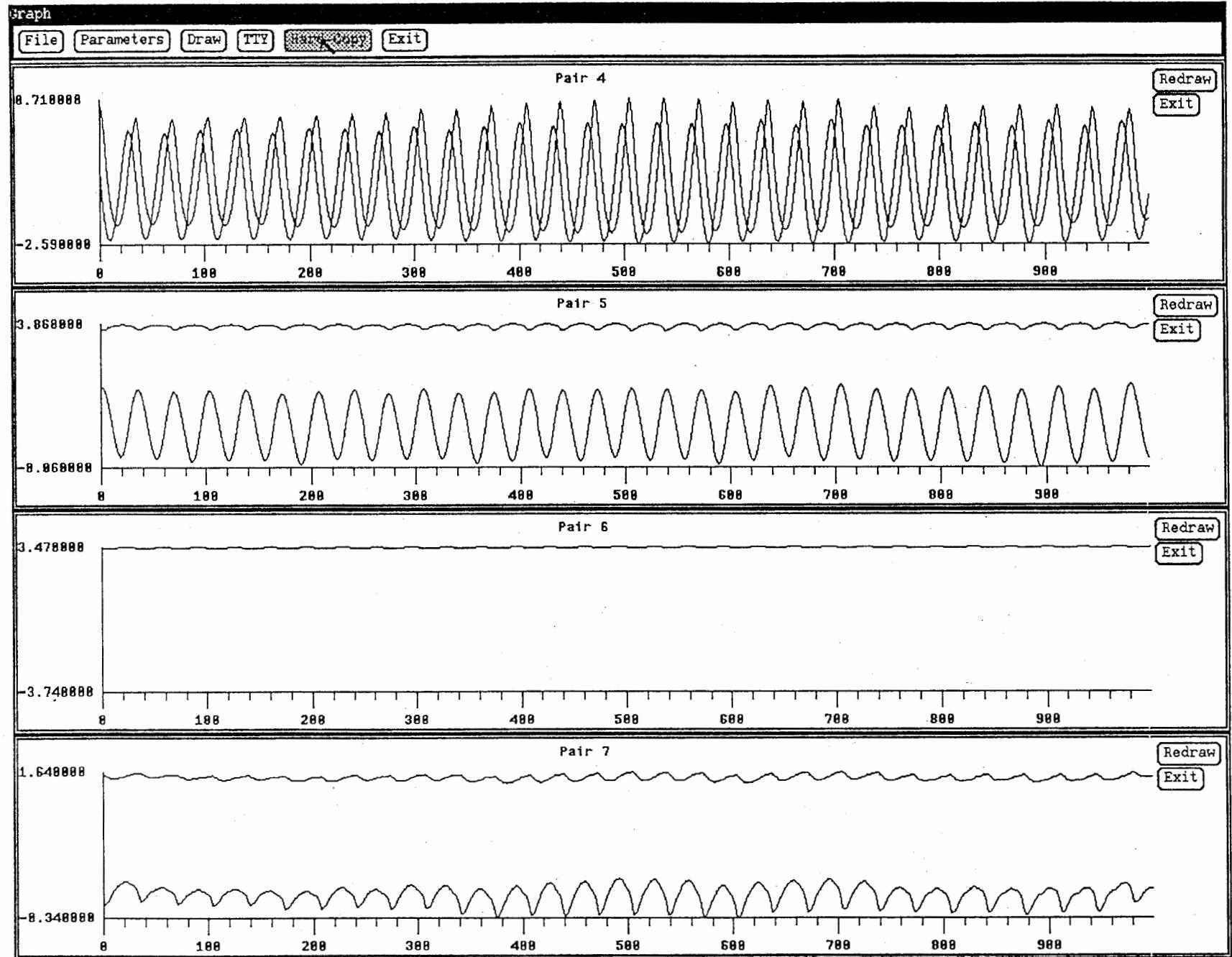
MEASURE ANALYSIS ANALYSIS2 INTERPOLA EDIT CURSOR WINDOW DATA FILE SIGNAL A/D D/A STEREO SETTINGS LAYOUT

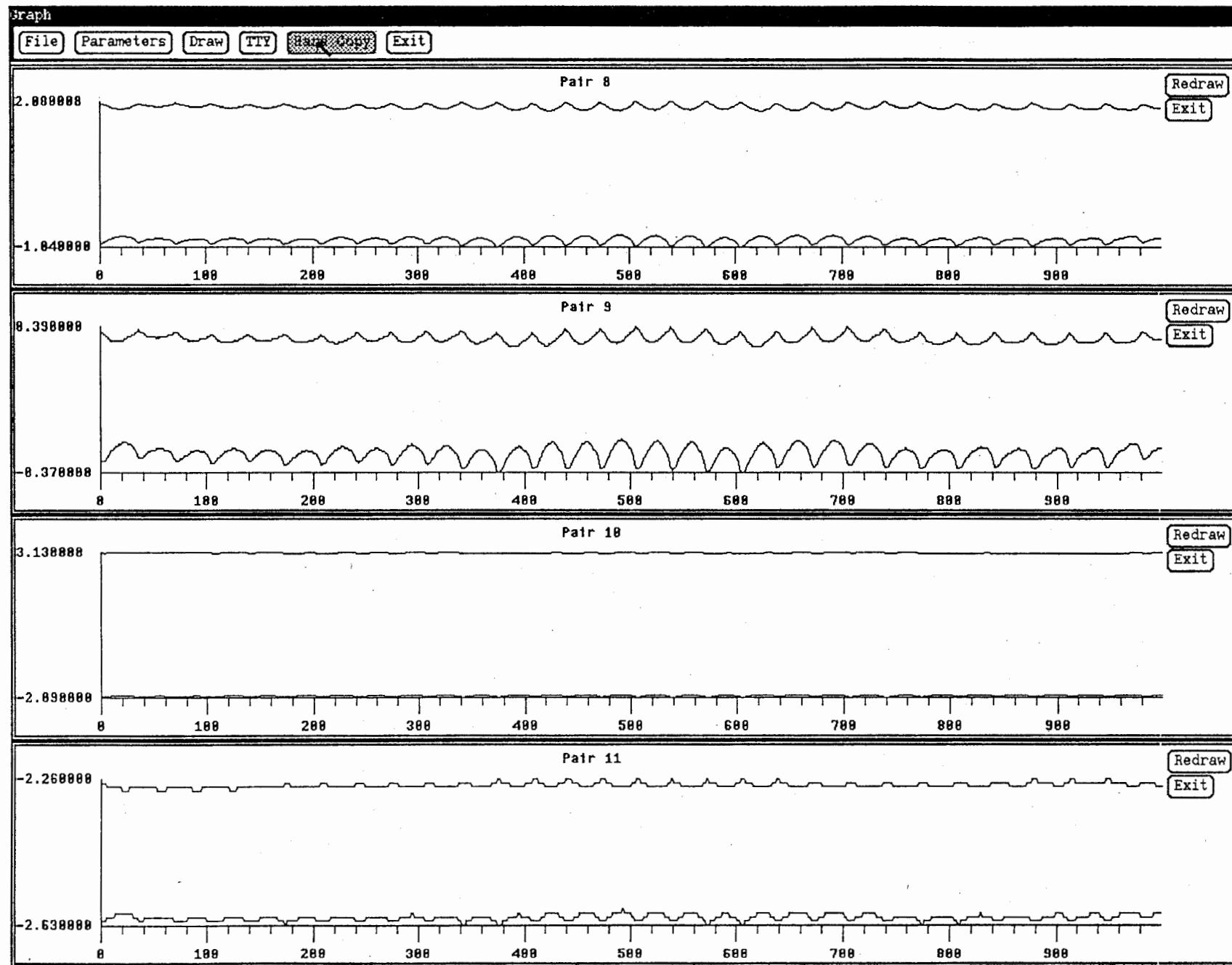


KHz









5. ウエーブレット

[1] ウェーブレット変換

13

- 基本ウェーブレット関数 : $\psi(t)$

- 2乗可積分 ($\int dt |\psi(t)|^2 = 1$)

- 時間一周波数空間で局在

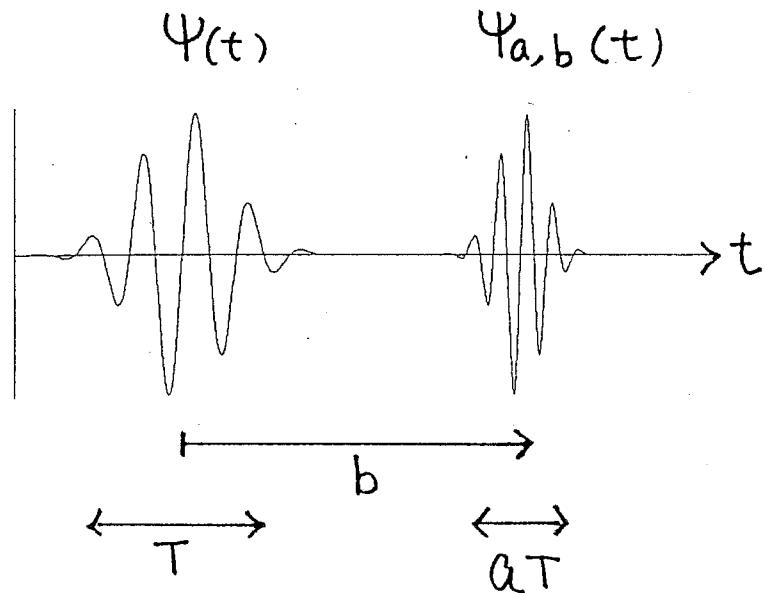
- アドミッシブル条件 : 直流成分 = 0

$$\hat{\psi}(0) = \frac{1}{\sqrt{2\pi}} \int dt \psi(t) = 0$$

- ウェーブレット関数 : $\psi_{a,b}(t)$

- a 倍スケール変換 & b 原点シフト

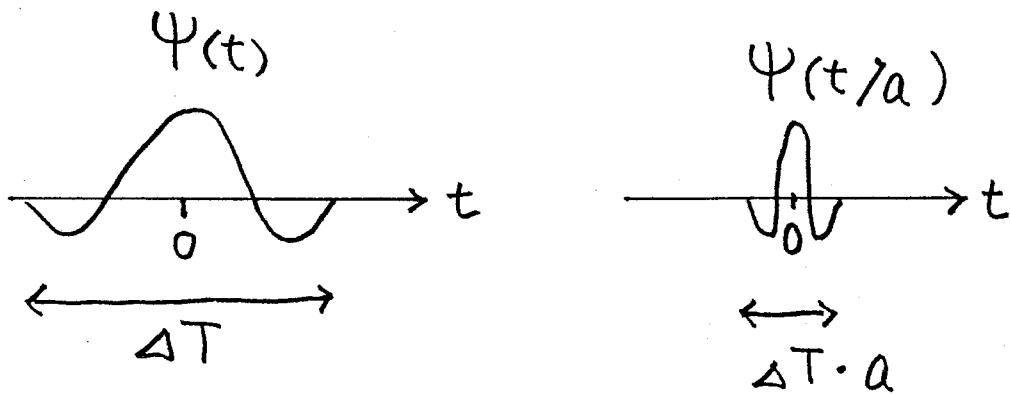
$$\psi_{a,b}(t) = a^{-1/2} \psi((t-b)/a)$$



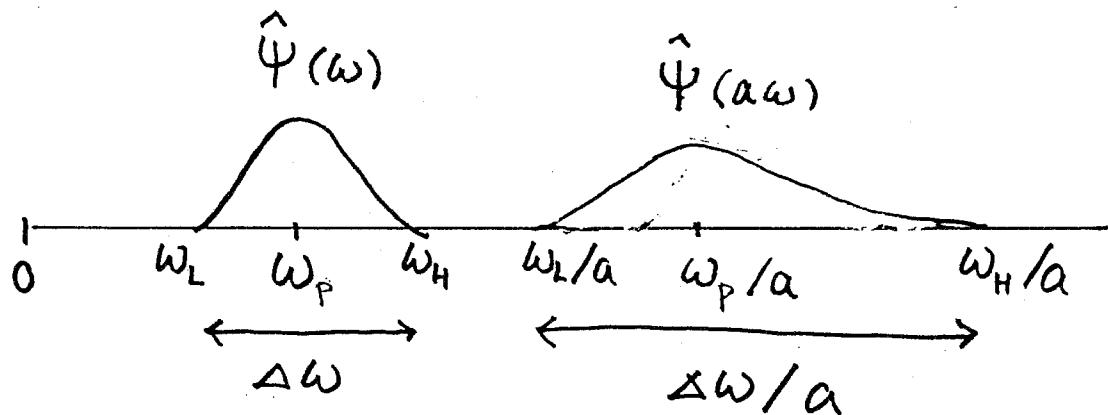
②スケール変換と時間-周波数分解能

14

- 時間的拡がり ($a < 1$)



- 周波数的拡がり ($\hat{\Psi}(\omega) = (2\pi)^{-1/2} \int dt e^{i\omega t} \Psi(t)$)



- a 倍スケール変換

$$\left\{ \begin{array}{l} \text{中心周波数: } \omega_p \rightarrow \omega_p/a \\ \text{時間的拡がり: } \Delta T \rightarrow a\Delta T \cdot a \\ \text{周波数的拡がり: } \Delta \omega \rightarrow \Delta \omega/a \end{array} \right.$$

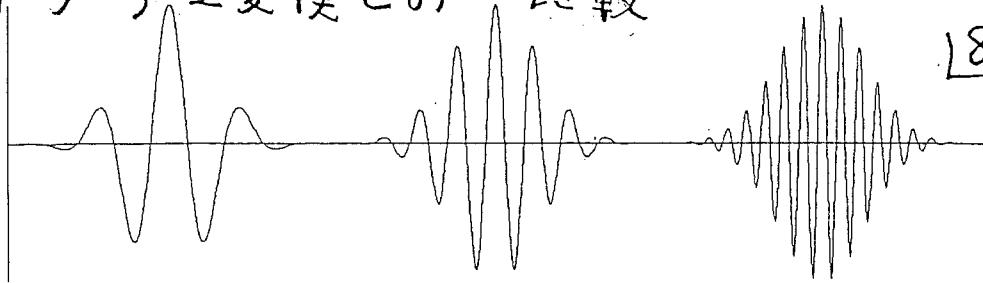
- 時間解像度 $\propto (\text{中心周波数})^{-1}$

- 周波数解像度 $\propto (\text{中心周波数})$

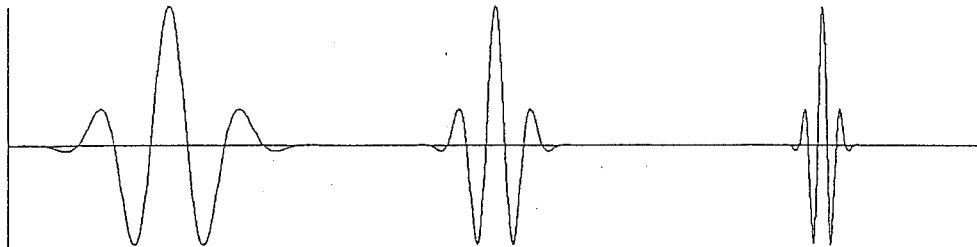
“高周波成分程、時間分解能が高い”

[3] 短時間 フーリエ変換との比較

短時間
フーリエ
変換の
基底周数



・ウェーブレット
変換の
基底周数



ウェーブレット変換

- スケール変換 & シフト変換

- ω -空間で等間隔サンプリング

- 特徴的スケールなし

- 時間分解能 $\propto (\text{中心周波数})^{-1}$

- 周波数 $\propto \propto (\text{中心周波数})$

- 非定常信号の分析

- 相似性

短時間 フーリエ 変換

- 周波数変調 & シフト変換

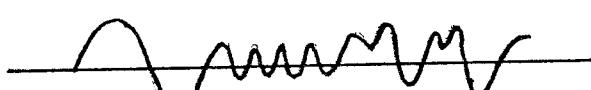
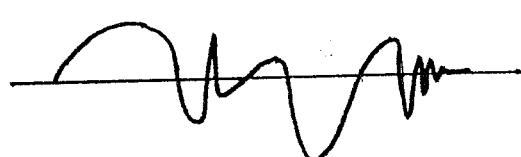
- ω -空間で等間隔サンプリング

- 窓周数の拡がり σ

$$\propto \sigma : \text{一定}$$

$$\propto \sigma^{-1} : \text{一定}$$

- 準定常信号の分析



[8] 自己相似性の角解析 124

・スケーリング指數

- $f_{x_0}(x) = f(x+x_0) - f(x_0)$

$$f_{x_0}(\gamma x) = \gamma^D f_{x_0}(x)$$

- $S_{x_0}(a, b) = \int dx \bar{\Psi}_{a,b}(x) f_{x_0}(x)$

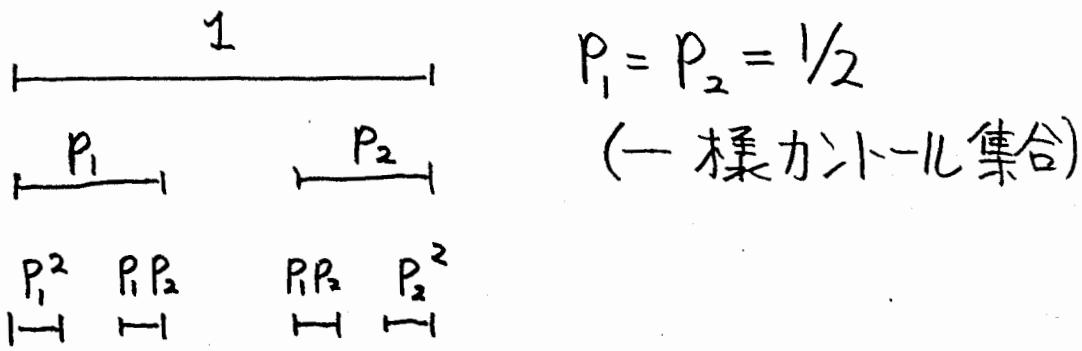
- $S(a, b) = \int dx \bar{\Psi}_{a,b}(x) f(x)$

$$S_{x_0}(\gamma a, \gamma b) = \gamma^{D+1/2} S_{x_0}(a, b)$$

$$S(\gamma a, x_0) = \gamma^{D+1/2} S(a, x_0)$$

◦ 非一様カントール集合

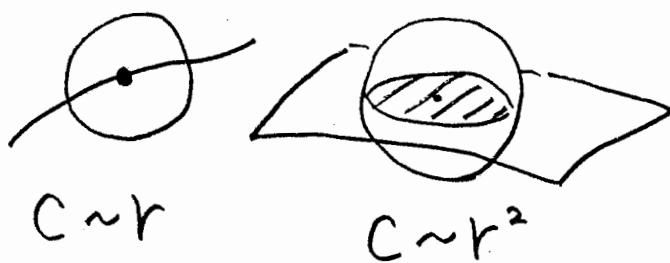
126



◦ 相対次元

$C(x, r) =$ 点 x を中心として半径 r の球

をかき、その中にいる点の総質量



$$C(x, r) \sim r^d$$

$$d_C(x) = \lim_{r \rightarrow 0} \ln C(x, r) / \ln r$$

tion of the construction process of the Cantor set. The positions of the local singularities of the measure are easily identified as spatial points b^* where the

(Arneodo, G. Grasseau

M. Holschneider 1988)

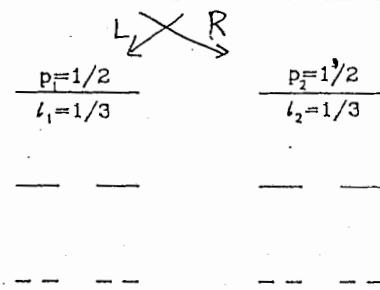
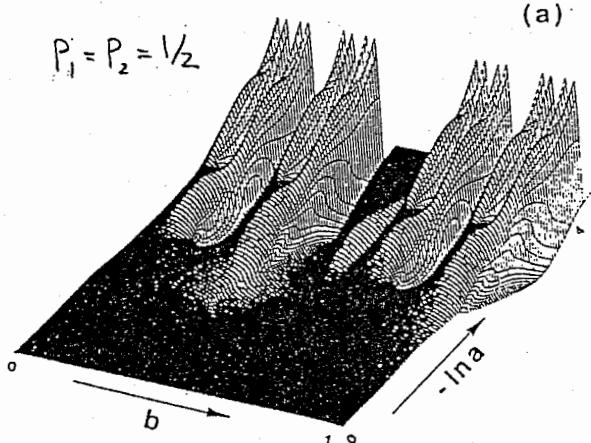


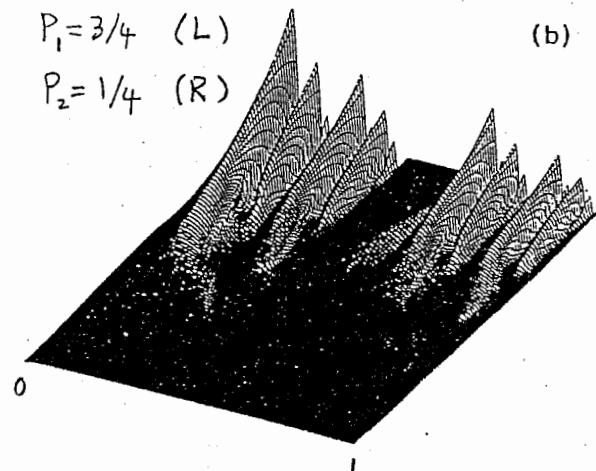
Fig. 1 At each stage of the construction of the standard Cantor set, the central third of each segment is removed from the set. Each segment has measure $p = 1/2$ and scale like $\ell = 1/3$ at the generation n .

Uniform Cantor Set

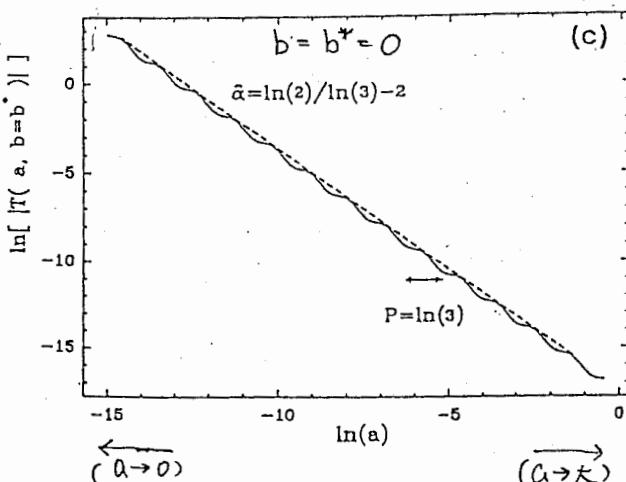


(a)

Nonuniform Cantor Set



(b)



($a \rightarrow \infty$)

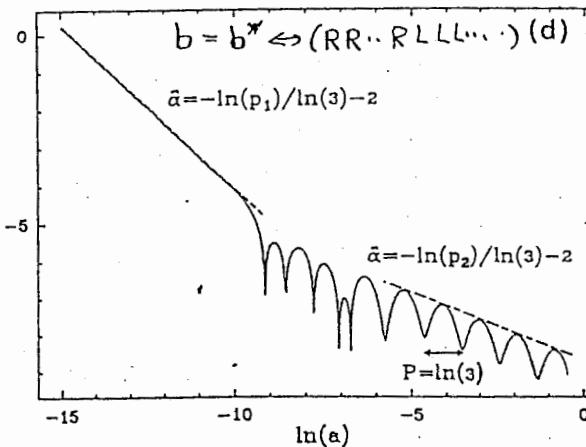


Fig. 2 The wavelet transform ($\text{sgn}(T)$, $|T(a, b)|^{1/2}$) of the triadic Cantor set with: (a) uniform measure $p_1 = p_2 = 1/2$; (b) two distinct measures $p_1 = 3/4$, $p_2 = 1/4$ (the scales in (a) and (b) are different). (c) $\ln|T(a, b=0)|$ (arbitrary scale) vs $\ln(a)$ for the uniform Cantor set. (d) $\ln|T(a, b=b^*)|$ (arbitrary scale) vs $\ln(a)$ for the nonuniform Cantor set where b^* corresponds to the kneading sequence $RRRRRRRRLLL..LL..g(x)$ is defined in Eq(6) and $n=2$ in Eqs (2) and (5).

Conclusion

- Recurrent Networks have capability of learning complex nonlinear dynamics such as chaos.
- Learning hidden dynamics is an open problem.
- Application for speech synthesis

Learning fluctuations of human voice

Recurrent network and chaos may open a new area of active and dynamic processing