

Copyright © 1991, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**NONLINEAR DYNAMICS OF A CLASS OF
ANALOG-TO-DIGITAL CONVERTERS**

by

Orla Feely and Leon O. Chua

Memorandum No. UCB/ERL M91/30

22 April 1991

**NONLINEAR DYNAMICS OF A CLASS OF
ANALOG-TO-DIGITAL CONVERTERS**

by

Orla Feely and Leon O. Chua

Memorandum No. UCB/ERL M91/30

22 April 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

Nonlinear Dynamics of a Class of Analog-to-Digital Converters

Orla Feely and Leon O Chua*

April 22, 1991

Abstract

Oversampled sigma-delta modulators are finding widespread use in audio and other signal processing applications, due to their simple structure and robustness to circuit imperfections. Exact analyses of the system are complicated by the presence of a discontinuous non-linear element — a one-bit quantizer. In this paper we study the dynamics of the one-dimensional mapping which models the behavior of the single-loop modulator. This mapping has a discontinuity at the origin and constant slope at all other points. With slope one, the dynamics in the region of interest reduce to those of the rotation of the circle. With slope less than one, almost all system inputs give rise to globally asymptotically stable periodic orbits. We emphasize the case with slope greater than one, and explain the structure of the resultant bifurcation diagram. A symbolic dynamics based study allows us to explain the self-similarity of the dynamics and the nature of chaos in the system.

*The authors are with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA

1 Introduction

Oversampled Sigma-Delta ($\Sigma - \Delta$) modulation as a method of analog-to-digital conversion in electronic circuits has attracted much interest since it was first studied in [Inose & Yasuda, 1963] and [Candy, 1974]. The technique is now finding widespread use in such applications as digital signal processing systems, voiceband telecommunication systems and commercial compact disc players [Goedhart et al, 1982; Misawa et al, 1981]. The simple structure of the $\Sigma - \Delta$ modulator, together with its robustness towards circuit imperfections and component matching inaccuracy, make it especially attractive for integrated circuit implementation.

Understanding of the operation of the $\Sigma - \Delta$ modulator is far from complete, due to the presence of a nonlinear element — a quantizer — in the system. Most researchers who study the problem begin by linearizing the nonlinearity, thus allowing standard linear theory to be applied, while others avoid the analysis altogether and confine their efforts to simulation. These studies provide little insight into the operation of the system, and in many cases do not yield correct quantitative or even qualitative results.

In this paper we study the dynamics of the nonlinear difference equation which models the behavior of the single-loop modulator. The nature of the mapping makes it especially amenable to a symbolic dynamics based study [Hao, 1989]. The model contains a parameter p which accounts for component nonideality in the circuit implementation. We study the dependence of system behavior on this parameter.

In the ideal single-loop modulator $p = 1$. In this case the dynamics of our model in the region of interest are identical to those of the well-known rotation of the circle. To account for integrator leak, a common circuit im-

perfection, p is decreased below 1. In this case the region of admissibility of each observed limit cycle widens out from a point (in the $p = 1$ case) to an interval, giving rise to a tongue in the p -input plane. Global asymptotic stability of all limit cycles guarantees that the tongues cannot overlap. Various extensions to the basic modulator topology have been proposed — to examine the behavior of certain of these extensions it is necessary to set $p > 1$. This is the case which will be emphasized in this paper. In this case all periodic orbits are unstable and the tongues overlap. We use symbolic dynamics to study the overlap of the tongues and the chaotic nature of the motion.

Section 2 contains a description of $\Sigma-\Delta$ modulation for the benefit of readers who may not be familiar with the technique. Section 3 describes the difference equation used to represent the single-loop $\Sigma-\Delta$ system, and Section 4 explains the dynamics of this equation for $p < 1$. Section 5 considers the location and overlap of the tongues for $p > 1$. Section 6 discusses several bifurcation diagrams for the $p > 1$ system and contains a symbolic dynamics based proof of the nature of the chaos in the system. In Section 7 the main findings of the paper are summarized.

2 Sigma-Delta Modulation

With recent rapid advances in very large scale integrated circuit (VLSI) technology, digital means of transmitting, processing and storing data are becoming increasingly prevalent over conventional analog techniques. Telephone networks, audio systems and the new field of high definition television are just some of the areas in which digital signal processing is finding its way into everyday life. Digital systems offer several advantages over analog,

including smaller size, lower sensitivity to noise, greater reliability and lower cost.

The increasing use of digital techniques has led to significant research interest in the analog-to-digital and digital-to-analog converters which serve as interfaces between digital processing systems and real-world analog signals. Ideally, these interfaces should be implemented in VLSI technology so as to maximize reliability and minimize cost of the complete system. While modern VLSI techniques produce very high-speed and high-density digital circuits, they restrict the dynamic range and precision of the analog stages. Conventional analog-to-digital conversion techniques require high precision components and often do not take advantage of the very high speeds permitted by the VLSI technology.

Oversampled analog-to-digital converters overcome both of these disadvantages. The structure is simple and is tolerant of circuit imperfections and component matching inaccuracy. The quantization can be coarse — in the basic implementation the quantizer has just two levels. To permit accurate signal reconstruction with such a quantizer the signal is sampled at a rate much higher than the usual Nyquist rate ¹, and a large number of the resultant coarse representations of the signal are used to generate a single high resolution representation.

Sigma-Delta ($\Sigma-\Delta$) modulation is the most popular method of oversampled analog-to-digital conversion. In the simplest $\Sigma-\Delta$ modulator — the single-loop system — a one-bit quantizer is used together with a discrete time integrator inside a feedback loop. This basic structure can be modi-

¹The Nyquist rate is twice the largest frequency component of the signal — it is the smallest sampling rate at which a continuous-time signal is uniquely represented by its discrete-time samples. See [Oppenheim *et al.*, 1983].

fied by adding more feedback loops, increasing the number of quantization levels or changing the forward path transfer function. Since such modifications increase circuit complexity and often give rise to instability, the most commonly used Σ - Δ structures are the single- and double-(feedback) loop modulators.

One important feature of Σ - Δ modulation is the appearance of limit cycles in the output bit stream. As a result, the quantization noise of the single-loop system is not white, but rather contains discrete spikes at frequencies depending on the input. This “pattern noise” can be particularly objectionable in audio applications. Higher order systems suffer from this problem to a lesser extent than does the single-loop system.

3 Single-Loop Σ - Δ Modulator with Leaky Integration

The structure of the ideal single-loop Σ - Δ modulator is as shown in Fig. 1.

Figure 1

It consists of a discrete-time integrator together with a quantizer in a feedback loop. The only nonlinear element in the modulator is the one-bit quantizer whose output is 1 when its input is ≥ 0 ; -1 when its input is negative. We will assume throughout this paper that the input to the modulator is constant. Under this assumption, the ideal single-loop system is described by the first order difference equation

$$u_{n+1} = u_n + x - \text{sgn}(u_n) \tag{1}$$

where x is the input to the modulator and u_n is the quantizer input. Both are discrete-time signals — since the input is constant we drop the subscript on x .

This Σ - Δ model has been studied in [Friedman, 1988] and [Gray, 1987,1989]. If the initial state u_0 lies in the interval $[x - 1, x + 1)$ the dynamics are equivalent to those of the rotation of the circle

$$\theta_{n+1} = (\theta_n + k)_{\text{mod } 2\pi} \quad (2)$$

which is covered in any elementary textbook on nonlinear dynamical systems [Hao, 1989; Devaney, 1989]. It follows that for rational input x the output bit stream is periodic, the average over a complete period being equal to x . With an irrational input to the modulator the output is quasiperiodic.

One major assumption made in modeling the single-loop system by (1) is that the integrator is ideal. In any practical implementation of the modulator, circuit nonidealities will result in leaky integration [Gregorian and Temes, 1986]. giving the more complete system description

$$u_{n+1} = p u_n + g.(x - \text{sgn}(u_n)) \quad (3)$$

p and g represent the effect of circuit nonidealities. g is always positive, so it can be removed by a scale change. As long as $p = 1$, therefore, the dynamics of (3) on the interval $[g(x - 1), g(x + 1))$ are still those of the rotation of the circle, so the qualitative features mentioned earlier still hold. In a practical implementation, circuit nonidealities will cause p to be less than 1. In order to quantify the effect of leaky integration on the behavior of the Σ - Δ modulator, it is necessary to study the dynamics of (3) for the case $p < 1$. This investigation is described, and the implications for the operation of the modulator discussed, in [Feely & Chua, 1990]. The

results are summarized in the following section — for further detail readers are referred to [Feely & Chua, 1990], [Veerman, 1987] and [Ding & Hemmer, 1987].

4 $p < 1$

With the leaky $\Sigma - \Delta$ modulator of Sec. 3 as our motivation, we study equation (3) for $0 < p < 1$. If there exists a limit cycle of period N , (3) can be summed over the limit cycle to yield the condition

$$u_k = \frac{g}{1-p} x - \frac{g}{1-p^N} \sum_{i=k}^{N+k-1} p^{N+k-1-i} \text{sgn}(u_i) \quad (4)$$

for $1 \leq k \leq N$. To check for admissibility of a given period N limit cycle we follow the procedure:

1. substitute the assumed bit sequence for the sgn terms in (4);
2. calculate the resulting sequence of states u_i in terms of x , g , and p ;
3. impose the conditions that these N u_i [quantizer inputs] must be of such polarities as to give the assumed bit sequence at the quantizer output;
4. calculate the range of possible values for x from the N inequalities in step 3.

In the ideal single-loop $\Sigma - \Delta$ system each limit cycle could exist for a fixed value of x only; the average output over the complete limit cycle being equal to x . In the $p < 1$ case each limit cycle can exist over a range of x values, limiting the resolution of the modulator. All limit cycles are globally

asymptotically stable, so for fixed x and p there can exist at most one limit cycle.

The limit cycles (consisting of strings of ones and zeros representing quantizer outputs of $+1$ and -1 respectively) which appear at the output of the ideal single-loop $\Sigma - \Delta$ modulator with constant rational input x are those produced by the Euclid algorithm of number theory [Friedman, 1988]. We will use the following special form of the Euclid algorithm:

Algorithm: To find the limit cycle with a ones and $(b - a)$ zeros (a and b coprime)

(i) Form the continued fraction expansion of a/b

$$\frac{a}{b} = \frac{1}{\alpha_1 + \frac{1}{\alpha_2 + \frac{1}{\alpha_3 + \frac{1}{\dots + \frac{1}{\alpha_n}}}}}$$

To guarantee uniqueness we require that the final coefficient $\alpha_n \neq 1$.

(ii) Define

$$\begin{aligned} S_0 &= 0 \\ S_1 &= 1(0)^{\alpha_1 - 1} \\ &\vdots \\ S_k &= S_{k-2}(S_{k-1})^{\alpha_k} \\ &\vdots \\ S_n &= S_{n-2}(S_{n-1})^{\alpha_n} \end{aligned}$$

where $(S_j)^{\alpha_k}$ consists of the block S_j repeated α_k times, and the α_i are the coefficients of the continued fraction expansion.

S_n is the limit cycle at the output of the ideal single-loop system with constant input $x = 2a/b - 1$. For n even we term S_n the *R-sequence* corresponding to S_n . The first two bits of S_n in this case are 01. Interchanging these two bits gives the *L-sequence* of S_n . For n odd S_n begins with the bits 10 — we term this the L-sequence of S_n and obtain the R-sequence by interchanging the first two bits. It can easily be shown that the L-sequence is a shifted version of the R-sequence.

To determine whether the limit cycle represented by S_n persists as p is decreased below 1, substitute into (4) to find the condition

$$\frac{\sum_{i=k_1}^{N+k_1-1} p^{N+k_1-1-i} v_i}{p^{N-1} + p^{N-2} + \dots + 1} > x \geq \frac{\sum_{i=k_2}^{N+k_2-1} p^{N+k_2-1-i} v_i}{p^{N-1} + p^{N-2} + \dots + 1} \quad (5)$$

Here $S_n = (s_1, \dots, s_N)$, $s_i \in \{0, 1\}$, $v_i = 1$ (resp. -1) if $s_i = 1$ (resp. 0) and k_1 and k_2 (both $\leq N$) are chosen subject to the constraint $v_{k_1} = -1$ and $v_{k_2} = 1$. To find the greatest lower bound on x it is necessary to find the shift k_2 of S_n which maximizes the polynomial

$$\sum_{i=k_2}^{N+k_2-1} p^{N+k_2-1-i} v_i$$

subject to the constraint $v_{k_2} = 1$. In fact the appropriate shift of S_n is just the L-sequence of S_n . Similarly the R-sequence is the shift of S_n which produces the least upper bound on x .

A limit cycle is admissible iff the corresponding greatest lower bound on x is lower than the corresponding least upper bound. For the “ideal” limit cycles (by which we mean those that appear in the $p = 1$ system) this is

always the case, the difference between the bounds being

$$\frac{2p^{N-2}(-p+1)}{p^{N-1} + \dots + p + 1} > 0 \quad \text{for } p \in (0, 1) \quad (6)$$

Thus any limit cycle which can exist at the output of the ideal $\Sigma - \Delta$ system can also exist at the output of the leaky system. Any limit cycle which is inadmissible in the ideal ($p = 1$) system is also inadmissible for $p < 1$. This is a consequence of the fact that for a given p the complement in $(-1, 1)$ of the set of input intervals which give rise to "ideal" limit cycles has measure zero.

Fig. 2 shows the dependence of the average output over a limit cycle on the dc input x for $p = 0.8$.

Figure 2

The plot was obtained by choosing 20000 dc input values uniformly spaced in the interval $[-1, 1]$. The form of the graph is that of the well known devil's staircase, the qualitative form being replicated at varying levels of resolution. The staircase contains a step at average output q , where q is any rational number in the range $(-1, 1)$. From (6) it is clear that the width of the steps corresponding to limit cycles with period N decreases with N . The widest step is that corresponding to the limit cycle 01 (average output 0) and the next widest are those corresponding to limit cycles 101 (average value $\frac{1}{3}$) and 100 (average value $-\frac{1}{3}$). Fig. 3 shows the 27 widest steps predicted by the analysis for $p = 0.8$.

Figure 3

The correspondence between theory and simulation is clear. Fig. 4 shows the locations of the 27 widest steps for varying x and p .

Figure 4

At $p = 1$, as expected, the widths of all steps shrink to zero, and the “steps” are just the rational numbers.

5 $p > 1$

In the leaky Σ - Δ system which motivated our study of equation (3), physical constraints require that p never exceed 1. There exist other Σ - Δ topologies, however, which attempt to overcome the disadvantages of the basic modulator by changing the forward path transfer function. [Chao *et al.*,1990], for example, describes an interpolative Σ - Δ modulator, the first order version of which is described by equation (3) with p no longer constrained to be less than 1. For this reason, in this section and for the remainder of the paper we consider the dynamics of (3) for the case $p > 1$. Once again we assume a constant input x with absolute value less than 1. Since g affects only the scaling of the state variable, we shall without loss of generality set $g = 1$, giving the difference equation

$$u_{n+1} = pu_n + x - \text{sgn}(u_n) \tag{7}$$

We can immediately deduce some important facts:

1. All fixed points and periodic orbits are unstable.
2. If $|x| \leq -1 + 2/p$ the interval $[\frac{x+1}{1-p}, \frac{x-1}{1-p}]$ is invariant under the map. In the context of Σ - Δ modulation, this means that if $|x| \leq -1 + 2/p$, all initial integrator states in the range $[\frac{x+1}{1-p}, \frac{x-1}{1-p}]$ give rise to a bounded series of quantizer inputs.

As outlined in Sec. 4, the range of inputs x which give rise to a particular admissible period N orbit for a fixed $p < 1$ is found to be of the form

$$\frac{l(p)}{1_N(p)} \leq x < \frac{r(p)}{1_N(p)} \quad (8)$$

Here $l(\cdot)$, $r(\cdot)$ and $1_N(\cdot)$ are $(N - 1)$ -th order polynomials. $1_N(\cdot)$ has all coefficients equal to 1, while the coefficients of $l(\cdot)$ and $r(\cdot)$ are obtained by applying the particular form of the Euclid algorithm given in Sec. 4.

When $p > 1$ the analysis is the same but for the reversal of the directions of the inequalities. The range of inputs x which give rise to a particular admissible period N orbit is now found to be of the form

$$\frac{r(p)}{1_N(p)} < x \leq \frac{l(p)}{1_N(p)} \quad (9)$$

Since

$$\frac{l(p) - r(p)}{1_N(p)} = \frac{2p^{N-2}(p-1)}{p^{N-1} + \dots + p + 1} > 0 \quad \text{for } p > 1 \quad (10)$$

we find that all admissible periodic orbits from the $p \leq 1$ system are again admissible for $p > 1$. Fig. 5 shows the tongues or regions of admissibility in the $p - x$ plane of these particular periodic orbits.

Figure 5

Whereas for $p \leq 1$ the tongues could not overlap, we see that for $p > 1$ there is overlap. The upper boundaries of two distinct tongues never intersect — this is a consequence of the fact that the upper boundaries correspond to periodic orbits containing the point $x - 1$. Similarly the lower boundaries of two distinct tongues never intersect. This fact will be used in later analysis.

Global asymptotic stability was used for $p < 1$ to rule out the possibility of periodic orbits other than those produced by the Euclid algorithm.

Since with $p > 1$ periodic orbits are no longer globally asymptotically stable, other periodic orbits may be admissible. The technique of Sec. 4 is again applicable. To test whether a given sequence $V = (v_1, \dots, v_N)$ of 1s and -1 s represents an admissible periodic orbit we substitute into

$$\frac{\sum_{i=k_1}^{N+k_1-1} p^{N+k_1-1-i} v_i}{p^{N-1} + p^{N-2} + \dots + 1} \leq x < \frac{\sum_{i=k_2}^{N+k_2-1} p^{N+k_2-1-i} v_i}{p^{N-1} + p^{N-2} + \dots + 1} \quad (11)$$

where $v_{k_1} = -1$ and $v_{k_2} = 1$. This is just (5) with the direction of the inequality signs reversed. V is shifted to find the tightest bounds on x — if the greatest lower bound is less than the least upper bound for some x and p the orbit is admissible for those parameter values. This investigation proves that orbits represented by *any* periodic sequence of zeros and ones are admissible for some values of $p < 2$ and $x \in (-1, 1)$. This follows from the inequality

$$p^N > p^{N-1} + p^{N-2} + \dots + 1 \quad \text{for } p \geq 2, \quad (12)$$

which implies that the fact that the upper bound polynomial has first coefficient 1 and the lower has first coefficient -1 is in itself enough to guarantee admissibility for some values of x and $p < 2$.

Fig. 6 shows the regions of admissibility of several orbits with average value 0.

Figure 6

Note that these regions all lie inside the region of admissibility of the “ideal” orbit with average 0 — the 10 orbit. In fact the region of admissibility for any orbit lies within that of the “ideal” orbit with the same average value. This is proven in the following theorem.

Theorem 1: Let S_k be the symbolic representation of a periodic orbit produced by the Euclid algorithm and let F be the symbolic representation of a second periodic orbit which has the same average value as S_k . The region of admissibility of F lies within that of S_k .

Proof: We will show that the upper boundary of the region of admissibility of F lies below the upper boundary of the region of admissibility of S_k . The corresponding proof for the lower boundaries is analogous.

An upper bound on the region of admissibility of F is the rational function

$$\frac{f(p)}{1_{mN}(p)}$$

where mN is the length of F and $f(\cdot)$ is a polynomial of order $mN - 1$ with coefficients given by any shift of F which leaves a 1 in the first position.

The least upper bound on the region of admissibility of S_k is

$$\frac{l(p)}{1_N(p)} = \frac{l(p)(1 + p^N + \dots + p^{(m-1)N})}{1_{mN}(p)}$$

Our goal is to show that there is some shift of F which leaves a 1 in the first position and produces

$$f(p) < l(p)(1 + p^N + \dots + p^{(m-1)N}) \quad \text{for } p > 1$$

Define $L = L_k L_k \dots L_k$ (repeated m times) where L_k is the L-sequence of S_k and so gives the coefficients of $l(\cdot)$. If there is a shift \hat{F} of F with first element 1 such that the cumulative total of $(L - \hat{F})$ is everywhere non-negative, the theorem is proven.

Suppose this is not the case — suppose that for every such shift of F there is a point q where

$$\sum_{i=1}^j L(i) - F(i) \geq 0 \quad \forall j < q \quad \text{and} \quad \sum_{i=1}^q L(i) - F(i) < 0$$

Take the shift with largest such q : call it F^* . (If there are several such shifts, choose any of them.) By producing another shift which gives rise to a cumulative total function which is non-negative until $q^* > q$, we will prove the theorem by contradiction.

Choose the largest \hat{q} such that (i) $F^*(\hat{q}) = 1$ and (ii) $\sum_{i=1}^{\hat{q}} L(i) - F^*(i) < 0$. We know such a \hat{q} exists, since q itself satisfies conditions (i) and (ii). Define \hat{F} (resp. \hat{L}) to be the shift of F (resp. L) with \hat{q} in first position.

Case 1: If $L(\hat{q}) = 1$,

$$\sum_{i=1}^j \hat{L}(i) - \hat{F}(i) \geq 0 \quad \forall j \leq q + mN - \hat{q}$$

and in fact over this range of j , $\sum_{i=1}^j \hat{L}(i) - \hat{F}(i) = 0$ only for $1 \leq j < \hat{j}$, where $\hat{F}(i) = 0 \quad \forall 2 \leq i \leq \hat{j}$

Applying Lemma 11 of [Feely & Chua, 1990] yields that

$$\sum_{i=1}^j L(i) - \hat{F}(i) \geq 0 \quad \forall j \leq q + mN - \hat{q} = q^*$$

Case 2: If $L(\hat{q}) = 0$,

$$\sum_{i=1}^j \hat{L}(i) - \hat{F}(i) \geq -1 \quad \forall j \leq q + mN - \hat{q}$$

and in fact over this range of j , $\sum_{i=1}^j \hat{L}(i) - \hat{F}(i) = -1$ only for $1 \leq j < \hat{j}$, where $\hat{F}(i) = 0 \quad \forall 2 \leq i \leq \hat{j}$

Applying Lemma 11 of [Feely & Chua, 1990] yields that

$$\sum_{i=1}^j L(i) - \hat{F}(i) \geq 0 \quad \forall j \leq q + mN - \hat{q} = q^*$$

The theorem is therefore proven by contradiction. \square

We can derive some measures of the extent of tongue overlap. For $x = 0$ it appears from Fig. 5 that for $1 \leq p < p^* \approx 1.4$ the only orbits present are those with average value zero. A quick analysis confirms this intuition. We wish to find the infimum over all periodic orbits with average greater than zero of the set

$$\left\{ p \mid \frac{r(p)}{1_N(p)} = 0 \right\} \quad (13)$$

Since lower boundaries of distinct tongues cannot intersect, it suffices to consider the lower boundaries of periodic orbits with average value $1/(2N+1)$ for integer $N \rightarrow \infty$. These boundaries are given by

$$x = \frac{(-p+1)(1+p^2+\dots+p^{2N-4})+p^{2N-2}+p^{2N-1}-p^{2N}}{1+p+\dots+p^{2N}} \quad (14)$$

As $N \rightarrow \infty$, these expressions converge to

$$\frac{(1-p)(p^2-2)}{p^3+p^2} \quad (15)$$

which is positive for $1 < p < \sqrt{2}$. For $x = 0$ and $1 < p < \sqrt{2}$, therefore, all admissible periodic orbits have average value zero.

For $x = 0.02$, solving the inequality

$$\frac{(1-p)(p^2-2)}{p^3+p^2} > 0.02 \quad (16)$$

yields that for $1.05048\dots < p < 1.32211\dots$ all admissible periodic orbits have average zero. In Sec. 6 these values will be derived in a slightly different manner.

The same procedure can be repeated for any tongue. To study the tongue with average value $\frac{1}{3}$, we consider the lower boundaries of tongues ($N +$

1)/(3N + 1) and the upper boundaries of tongues N/(3N + 2) as N → ∞.

The loci converge to

$$\frac{-p^4 + p^3 + p^2 + 2p - 2}{p^4 + p^3 + p^2} \quad \text{and} \quad \frac{p^4 - p^3 + p^2 - 2p + 2}{p^4 + p^3 + p^2}$$

respectively. These loci, together with the boundary of the region of admissibility of the 101 orbit tongue, are plotted in Fig. 7.

Figure 7

In the shaded region all admissible periodic orbits have average value $\frac{1}{3}$.

For $x = 0.3$, for example, all periodic orbits will have average $\frac{1}{3}$ for p such that

$$\frac{-p^4 + p^3 + p^2 + 2p - 2}{p^4 + p^3 + p^2} > 0.3 > \frac{p^4 - p^3 + p^2 - 2p + 2}{p^4 + p^3 + p^2}$$

i.e. for $p \in (1.06123\dots, 1.16043\dots)$. Again, Sec. 6 provides a different interpretation of this analysis.

6 Bifurcation Diagrams

In this section we study the bifurcation structure of the map

$$u \rightarrow f(u, p, x) \doteq pu + x - \text{sgn}(u) \quad (17)$$

with bifurcation parameter p . We begin with the case $x = 0$. The bifurcation diagram corresponding to this case is plotted in Fig. 8.

Figure 8

(This diagram was obtained using the “brute-force” method described in [Parker & Chua, 1989].)

To aid us in our study of this diagram we define the following functions

$$\begin{aligned}
 F_1^+(p) &= f(1, p, 0) & F_1^-(p) &= f(-1, p, 0) \\
 F_2^+(p) &= f^2(1, p, 0) & F_2^-(p) &= f^2(-1, p, 0) \\
 F_3^+(p) &= f^3(1, p, 0) & F_3^-(p) &= f^3(-1, p, 0) \\
 & & & \vdots
 \end{aligned}$$

These functions are plotted in Fig. 9.

Figure 9

The correspondence between Figs. 8 and 9 is clear:

- (i) The band boundaries of the bifurcation diagram are given by

$$\begin{aligned}
 &\pm 1 && \text{for } p > p_1 \approx 1.4 \\
 &\pm 1, F_1^\pm(p), F_2^\pm(p) && \text{for } p_1 > p > p_2 \approx 1.2 \\
 &\pm 1, F_1^\pm(p), \dots, F_6^\pm(p) && \text{for } p_2 > p > p_3
 \end{aligned}$$

- (ii) The bands of the bifurcation diagram merge at points p^* where

$$F_i^+(p^*) = F_j^+(p^*) \text{ for } i \neq j$$

If two of the functions F_i intersect at a point, infinitely many of the F_i intersect at that point.

- (iii) Within the bands of the bifurcation diagram can be seen darker “shadowing” bounded by the functions F_i .

These phenomena are easily explained by the fact that 1 and -1 are extrema of the map (17), leading to grouping of points on one side of the iterates of these values. Interested readers are referred to [Hao, 1989], where this approach is applied to a number of systems.

From (ii) above it is clear that at the band-merging points the extrema ± 1 are homoclinic to unstable periodic orbits of (17). Homoclinicity in maps does not in general play as important a role as it does in flows, but it often provides useful insights. See [Block, 1978] and [Grebogi *et al.*, 1982] for further details. The merging of three bands of the bifurcation diagram into one occurs at p_1 , where the locus of the unstable symmetric period-2 orbit intersects the bands. At this point the orbit of $+1$ has the symbolic form $1(10)^\infty$. Substituting into (17) to find the condition for such homoclinicity yields the value $p_1 = \sqrt{2}$. Similarly, the merging of seven bands into three occurs at p_2 , where the orbit of $+1$ has the symbolic form $110(1001)^\infty$. Again, we can substitute into (17) to find $p_2 = \sqrt[4]{2}$. At $p_0 = 2$, the extrema ± 1 are fixed points of the mapping.

We will now use the self-similarity of the bifurcation structure to generalize these observations and to define precisely the sense in which (17) is chaotic for $p > 1$ and $x = 0$. For $p = 2$, the dynamics of (17) are equivalent to those of the circle map

$$\theta \rightarrow (2\theta)_{\text{mod } 2\pi} \quad \theta \in [0, 2\pi) \quad (18)$$

which is known to be chaotic [Devaney, 1989]. The chaotic nature is also clear from the proof in Sec. V that for $p = 2$ and $x = 0$ all possible blocks of ones and zeros correspond to admissible periodic orbits. (See [Hao, 1989] for a discussion of this feature in the context of the logistic map.)

We now state and prove the central theorem concerning the self-similarity of the bifurcation structure for $x = 0$.

Theorem 2: Given the symbolic representation S of a periodic orbit which is admissible for $x = 0$, $p = \hat{p} > 1$. Define T by performing the following operations on each element of S : $0 \rightarrow 01$ and $1 \rightarrow 10$. The periodic orbit represented symbolically by T is admissible for $x = 0$, $p = \sqrt{\hat{p}}$.

Proof: We will show that the least upper bound on the region of admissibility of T at $\sqrt{\hat{p}}$ has the same sign as the least upper bound on the region of admissibility of S at \hat{p} . The corresponding proof for the greatest lower bounds is analogous.

Consider any shift of T which leaves a 1 in the first position. We will show that either (i) this shift corresponds exactly to the least upper bound shift of S or (ii) we can find another shift of T with a 1 in first position which produces a lower polynomial.

Case 1: The 1 in the first position of T comes from a 10 block obtained by transforming a 1 from S (say the 1 in position q of S). This sequence leads to a polynomial $(p - 1)s(p^2)$, where $s(\cdot)$ is the polynomial with coefficients given by the shift of S which starts in position q . Evaluating this polynomial at $\sqrt{\hat{p}}$ gives $(\sqrt{\hat{p}} - 1)s(\hat{p})$. To minimize this over all such shifts we must choose the shift of S which minimizes $s(\hat{p})$ — i.e. we must ensure that our shifted T is a direct transformation of the least upper bound shift of S . Clearly the polynomial at $\sqrt{\hat{p}}$ corresponding to this shift of T will have the same sign as the least upper bound polynomial of S at \hat{p} .

Case 2: The 1 in the first position of T comes from a 01 block (obtained by transforming a 0 from S) preceded by a 01 block. Consider \hat{T} , obtained by shifting each element of T two places to the right. Consider $T - \hat{T}$. Over each 10 (resp. 01) block of T preceded by another 10 (resp. 01) block, $T - \hat{T} = 0$. Corresponding to each 10 (resp. 01) block of T preceded by a 01 (resp. 10) block, $T - \hat{T} = 1-1$ (resp, $-1 1$). $T - \hat{T}$ is of the form

$$(0)^{a_1}1-1(0)^{a_2}-1 1(0)^{a_3}1-1 \dots -1 1(0)^{a_n}$$

which implies that \hat{T} gives a lower polynomial than does T .

Case 3: The 1 in the first position of T comes from a 01 block (obtained by transforming a 0 from S) preceded by a 10 block. Consider \hat{T} , obtained by shifting each element of T three places to the right. $T - \hat{T}$ is then of the form

$$(0)^{a_1}1(0)^{a_2}-1(0)^{a_3}1(0)^{a_4}-1 \dots -1(0)^{a_n}$$

To prove this, one can write out all possible subblocks of T which would give a $-1(0)^{a_i}-1$ block in $T - \hat{T}$ — each has three adjacent elements of the same type, which is not allowable. Similarly a $1(0)^{a_i}1$ block is not allowable so, since the first non-zero element of $T - \hat{T}$ is 1, we get the above sequence. This implies that \hat{T} gives a lower polynomial than does T . □

We are now in a position to discuss the chaotic nature of (17) for $p > 1$ and $x = 0$. For $p \geq 2$ all symbolic blocks of zeros and ones correspond to admissible periodic orbits at $x = 0$, and the map is chaotic in the sense described in [Hao, 1989]. For $p \geq \sqrt{2}$, so, all symbolic blocks obtained by the transformation $0 \rightarrow 01$ and $1 \rightarrow 10$ correspond to admissible periodic

orbits. The motion here is as chaotic as it is for $p \geq 2$, since coarse-grained over the scale (10, 01) the orbits at $p \geq \sqrt{2}$ are identical to those at $p^2 \geq 2$. Successive applications of the same reasoning yield that the motion is as chaotic for $p > 2$ as it is for any $p > \sqrt[4]{2}, \sqrt[8]{2}, \sqrt[16]{2} \dots \sqrt[2^n]{2}$ — for any $p > 1$, in other words. Note also that this analysis yields the values of the band-merging points — $p_0 = 2, p_1 = \sqrt{2}, p_2 = \sqrt[4]{2} \dots$

We now generalize this analysis still further by showing that the (coarse-grained) motion along the center curve of the region of admissibility of *any* of the periodic orbits produced by the Euclid algorithm corresponds to the motion along the $x = 0$ line.

Theorem 3: Given a sequence B of zeros and ones such that the corresponding periodic orbit is admissible for $x = 0$ and $p = \hat{p} > 1$. Given any sequence S formed by the Euclid algorithm. Let N denote the length of S ; R the R-sequence of S ; and L the L-sequence. Apply the following transformation to each element of B to form a new sequence C :

$$1 \rightarrow L \qquad 0 \rightarrow R$$

Then the sequence C corresponds to an admissible periodic orbit for $p = \sqrt[N]{\hat{p}}$ and $x = (r(p) + l(p))/(2.1_N(p))$.

Proof: The proof will consist of two main steps. We will first show that the least upper bound and greatest lower bound shifts of C are the exact transformations of the least upper bound and greatest lower bound shifts of B . Then we will show that the least upper bound polynomial of C lies above the line $x = (r(p) + l(p))/(2.1_N(p))$ at $p = \sqrt[N]{\hat{p}}$ if the least upper bound polynomial of B is greater than zero at $p = \hat{p}$. The corresponding proof for the greatest lower bound shift is analogous.

Given any shift \hat{C} of C with first element 1 — say \hat{C} is the shift such that the k th element of C moves to first position.

- Case 1: $(k)_{\text{mod } N} = 1$. In this case, \hat{C} is the direct transformation of a shift of B . As in Theorem 2, therefore, the least upper bound shift of this type is the direct transformation of the least upper bound shift of B .
- Case 2: $(k)_{\text{mod } N} > 2$. Marking \hat{C} off in blocks of N from the start, we see that each such block is a shifted version of L beginning with a 1. It follows, using Lemma 11 of [Feely & Chua, 1990], that the direct transformation of any shift of B will produce a lower bound polynomial than does \hat{C} , so \hat{C} in this case can not be the least upper bound shift of C .
- Case 3: $(k)_{\text{mod } N} = 2$. Again, we mark \hat{C} off in blocks of n from the start. Each such block is a shifted version of L , except for those blocks taken from an LR or an RL block of C . In these cases the last element of the block is changed from a 0 to a 1 (in the case of the RL block) or from a 1 to a 0 (in the case of the LR block). Note also that the first element of a block is 1 for blocks taken from an RL or RR block of C and 0 for blocks taken from an LR or LL block of C . Combining these facts with Lemma 11 of [Feely & Chua, 1990], we find that the direct transformation of any shift of B will produce a lower bound polynomial than does \hat{C} , so \hat{C} in this case can not be the least upper bound shift of C .

The corresponding proof for the greatest lower bound shift is analogous, so the first step of the theorem is proven. Now we show that the least upper bound polynomial of C lies above the line $x = (r(p) + l(p))/(2.1_N(p))$ at

$p = \sqrt[\alpha]{\hat{p}}$ if the least upper bound polynomial of B is greater than zero at $p = \hat{p}$. Let $b(\cdot)/1_\alpha(\cdot)$ be the least upper bound polynomial of B ; $c(\cdot)/1_{\alpha N}(\cdot)$ the least upper bound polynomial of C .

$$\frac{c(p)}{1_{\alpha N}(p)} - \frac{r(p)}{2 \cdot 1_N(p)} - \frac{l(p)}{2 \cdot 1_N(p)} \geq 0$$

$$\Leftrightarrow \frac{c(p)}{1_N(p)(1 + p^N + \dots + p^{(\alpha-1)N})} - \frac{r(p)}{2 \cdot 1_N(p)} - \frac{l(p)}{2 \cdot 1_N(p)} \geq 0$$

$$\Leftrightarrow 2c(p) - r(p)(1 + p^N + \dots + p^{(\alpha-1)N}) - l(p)(1 + p^N + \dots + p^{(\alpha-1)N}) \geq 0$$

$$\Leftrightarrow 2(p-1)p^{N-2}b(p^N) \geq 0$$

$$\Leftrightarrow b(p^N) \geq 0$$

Once again, the corresponding proof for the greatest lower bound polynomial is analogous, so the theorem is proven. \square

Figure 10 shows the bifurcation diagram of (17) along the curve $x = 1/(p^2 + p + 1)$ — i.e. the center curve of the 101 tongue.

To conclude this section we briefly examine the bifurcation diagrams of (17) for $x = 0.02$ and 0.3 . The bifurcation diagram for $x = 0.02$ is plotted in Fig. 11.

Figure 11

Note the change in the structure of the bands at $p = 1.05048\dots$ and $1.32211\dots$. These values were obtained in Sec. 5 as the points where other tongues impinge on the 01 tongue for $x = 0.02$. We now recognize them as the points where the locus of the unstable period-2 orbit collides with the

bands — the points where one or other of the extrema -0.98 and 1.02 are homoclinic to an unstable period-2 orbit. The zeroth band-merging point p_0 , where the locus of the unstable period-1 orbit (fixed point) collides with the band, is at $p = 2/(x + 1) = 1.9607\dots$, as predicted in Sec. 5.

The bifurcation diagram for (17) with $x = 0.3$ is plotted in Fig. 12.

Figure 12

Notice again the zeroth band-merging at $p = 1.538\dots$ and the collision with the unstable 110 orbit at $p = 1.06123\dots$ and $1.16043\dots$ — the values obtained in Sec. 5 from a study of the 110 tongue and its neighbors. The 110 tongue is the only one whose influence is plainly visible in Fig. 12, but this is merely because it is the widest tongue of influence here. As explained in [Feely & Chua, 1990], given two tongues corresponding to periodic orbits where the fraction of ones present is f_1 and f_2 , the widest tongue between the two is found by taking the Farey composition of f_1 and f_2 [Hao, 1989]. Applying this operation here yields that the next widest tongue of influence is that with average $\frac{7}{23}$. In Fig. 13 we plot a magnified segment of the bifurcation diagram for $x = 0.3$.

Figure 13

The $\frac{7}{23}$ tongue is clearly visible between the homoclinic collisions of the 10101101101101101101 orbit.

7 Summary

With the $\Sigma-\Delta$ modulator as our motivation, we have studied the dynamics of (17) for varying $x \in (-1, 1)$ and $p > 0$. With $p = 1$ (17) reduces to the

simple rotation of the circle, so all rational inputs lead to periodic orbits, all irrational inputs to quasiperiodic orbits. As p is decreased below 1 the regions of admissibility of the periodic orbits widen out from points (at $p = 1$) to intervals, forming tongues in the $p - x$ plane. The bounds of these tongues are rational functions of p such that (i) the tongues never overlap, and (ii) almost all inputs give rise to periodic outputs. Note that for $p = 1$ the set of inputs which give periodic outputs has measure zero, while for $p < 1$ the set of inputs which *do not* give periodic orbits has measure zero.

Our emphasis in this paper has been on the $p > 1$ case. For $p > 1$ all periodic orbits are unstable, and for $|x| < -1 + 2/p$ the interval $[\frac{x+1}{1-p}, \frac{x-1}{1-p}]$ is invariant under (17). Again, the region of admissibility of each periodic orbit observed in the $p = 1$ system widens out to form a tongue in the $p - x$ plane, but in this case the tongues overlap. New periodic orbits now become admissible — in fact any periodic sequence of zeros and ones is admissible for some values of $p \leq 2$ and $x \in (-1, 1)$. For $x = 0$ and $p = 2$ *all* such orbits are admissible. There is a element of order in that the region of admissibility of each periodic orbit with average value q lies within the region of admissibility of the “ideal” periodic orbit with the same average value. The bifurcation diagrams of (17) were plotted for various values of x , and were discussed in the context of such concepts as homoclinic collisions and iterates of extrema. The self-similarity of the bifurcation structure for $x = 0$ was proven and used to explain the chaotic nature of the map along the line $x = 0$. This analysis was then generalized to discuss the self-similarity along the center curve of any of the tongues corresponding to “ideal” periodic orbits.

8 Acknowledgements

This work is supported in part by the Office of Naval Research under Grant N00014-89-J-1402

9 References

Block L [1978] "Homoclinic Points of Mappings of the Interval". *Proc. AMS*, 72, 576-580.

Candy J C [1974] "A use of Limit Cycle Oscillations to Obtain Robust Analog-to-Digital Converters". *IEEE Trans. Comm*, COM-22 298-305.

Chao K C H, Nadeem S, Lee, W L & Sodini C G [1990] "A Higher Order Topology for Interpolative Modulators for Oversampling A/D Converters". *IEEE Trans. Circuits and Systems*, CAS-37, 309-318.

Devaney Robert L [1989] *An Introduction to Chaotic Dynamical Systems*. (Redwood City, Calif.: Addison-Wesley)

Ding E-J & Hemmer P C [1987] "Exact Treatment of Mode Locking for a Piecewise Linear Map" *J. Stat. Phys*, 46, 99-110.

Feely Orla & Chua L O [1990] "The Effect of Integrator Leak in $\Sigma - \Delta$ Modulation" Univ. of California, Berkeley, Electronics Research Laboratory Memorandum No. M90/116. Also submitted to *IEEE Trans. Circuits and Systems*.

Friedman Vladimir [1988] "The Structure of the Limit Cycles in Sigma Delta Modulation". *IEEE Trans Comm.*, COM-36, 972-979.

Goedhart D, van de Plassche R & Stikvoort E [1982] "Digital-to-Analog Conversion in Playing a Compact Disc". *Philips Tech. Rev.*, 40 174-179.

Gray R M [1987] "Oversampled Sigma-Delta Modulation". *IEEE Trans. Comm.*, COM-35, 481-489.

Gray R M [1989] "Spectral Analysis of Quantization Noise in a Single Loop Sigma Delta Modulator with dc Input". *IEEE Trans. Comm.*, COM-37, 588-599.

Grebogi C, Ott E & Yorke J A [1982] "Crises, Sudden Changes in Chaotic Attractors, and Transient Chaos". *Physica*, 7D, 181.

Gregorian Roubik & Temes Gabor C [1986] *Analog MOS Integrated Circuits for Signal Processing*. (New York: Wiley), p. 485.

Hao Bai-Lin [1989] *Elementary Symbolic Dynamics and Chaos in Dissipative Systems*. (Singapore: World Scientific)

Inose H & Yasuda Y [1963] "A Unity bit Coding Method by Negative Feedback", *Proc. IEEE* 51 1524-1535

Misawa T, Iwerson J, Loporcario L & Ruch J [1981] "Single-chip per Channel Codec with Filters utilizing $\Sigma-\Delta$ Modulation". *IEEE Journ. of Solid State*

Circuits, SC-16 333–341.

Oppenheim A, Willsky A & Young I [1983] *Signals and Systems* (New Jersey: Prentice–Hall)

Parker T S & Chua L O [1989] *Practical Numerical Algorithms for Chaotic Systems*. (New York: Springer–Verlag) Chap. 8.

Veerman P [1987] “Symbolic Dynamics of Order-Preserving Orbits” *Physica*, **29D**, 191–201.

List of Figures

Figure 1: Block diagram of ideal single-loop Σ - Δ system. The modulator consists of a discrete-time integrator and a one-bit quantizer inside a feedback loop. x is the input to the modulator and u the input to the quantizer.

Figure 2: Average output over a limit cycle plotted versus dc input x . This graph corresponds to the leaky single-loop system with $p = 0.8$.

Figure 3: Analytically predicted intervals of admissibility of the 27 shortest limit cycles for $p = 0.8$. Compare to Fig. 2.

Figure 4: Regions of admissibility of the 27 shortest limit cycles — i.e. those with period ≤ 9 — for $0 < p < 1$. Tongues of the same color correspond to limit cycles of the same period: turquoise — period 2; red — period 3; dark green — period 4; dark blue — period 5; blue-grey — period 6; orange — period 7; light green — period 8; purple — period 9.

Figure 5: Regions of admissibility of the 27 shortest limit cycles for $0 < p < 2$. The region of admissibility of the 10 orbit is highlighted in blue and that of the 101 orbit in red. Note the overlap in purple.

Figure 6(a): Regions of admissibility of several orbits with average zero — symmetric.

Figure 6(b): Regions of admissibility of several orbits with average zero — non-symmetric.

Figure 7: 101 tongue — in the shaded region all admissible orbits have average $\frac{1}{3}$.

Figure 8: Bifurcation diagram of Eq.(17) for $1 < p < 2$ with $x = 0$. The horizontal axis shows u .

Figure 9: Skeleton of bifurcation diagram of Eq. (17) for $1 < p < 2$ with $x = 0$ — i.e. $F_1^\pm(p), \dots, F_8^\pm(p)$ plotted versus p .

Figure 10: Bifurcation diagram of Eq. (17) for $1 < p < 2$ along the curve $x = 1/(p^2 + p + 1)$ (center of the 101 tongue). The horizontal axis shows u .

Figure 11: Bifurcation diagram of Eq. (17) for $1 < p < 2$ with $x = 0.02$. The horizontal axis shows u .

Figure 12: Bifurcation diagram of Eq. (17) for $1 < p < 2$ with $x = 0.3$. The horizontal axis shows u .

Figure 13: Bifurcation diagram of Eq. (17) for $1 < p < 1.05$ with $x = 0.3$. The horizontal axis shows u .

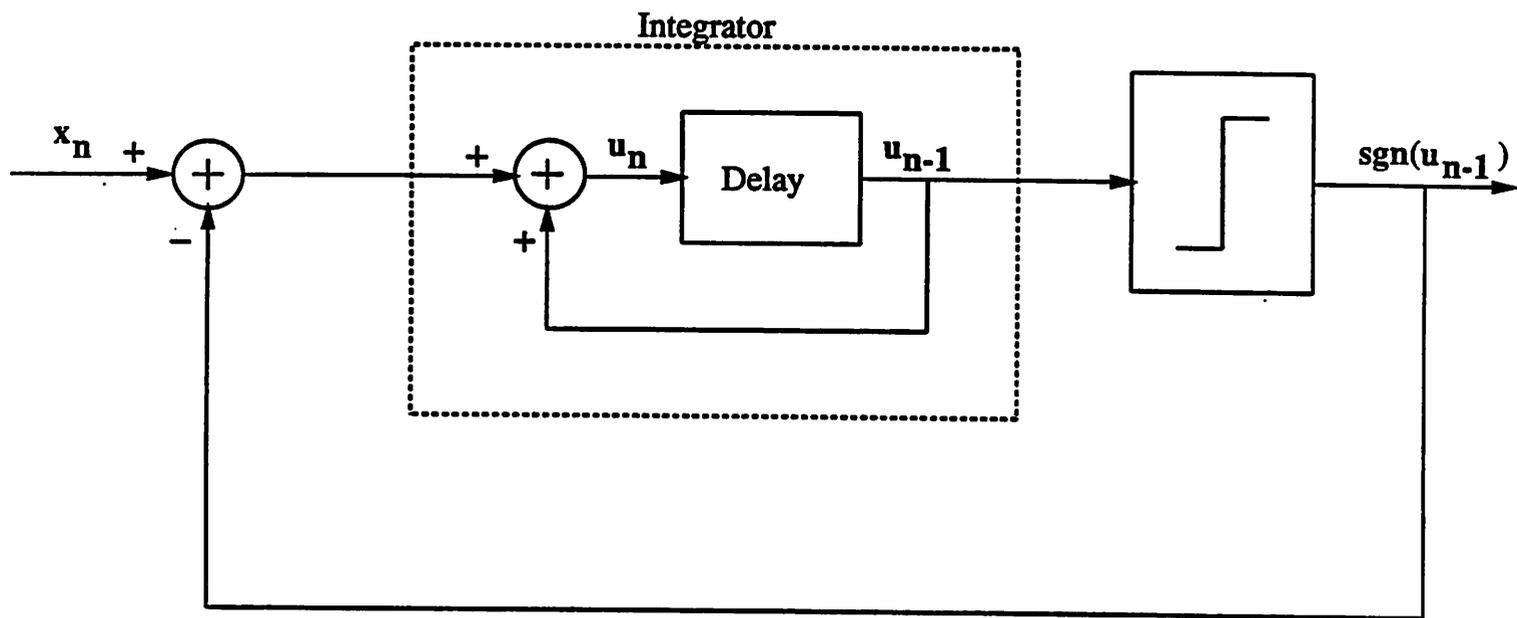


Figure 1

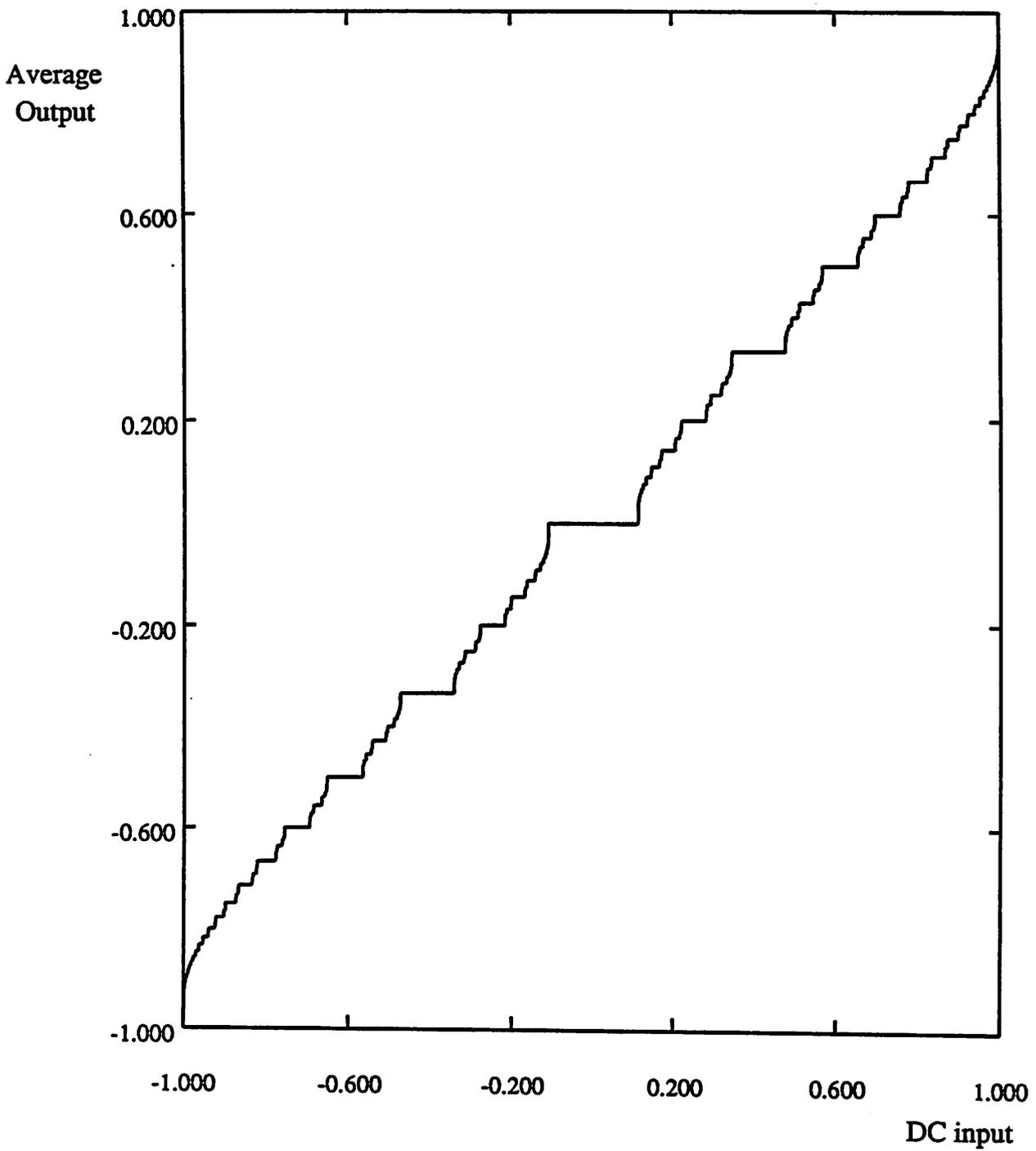


Figure 2

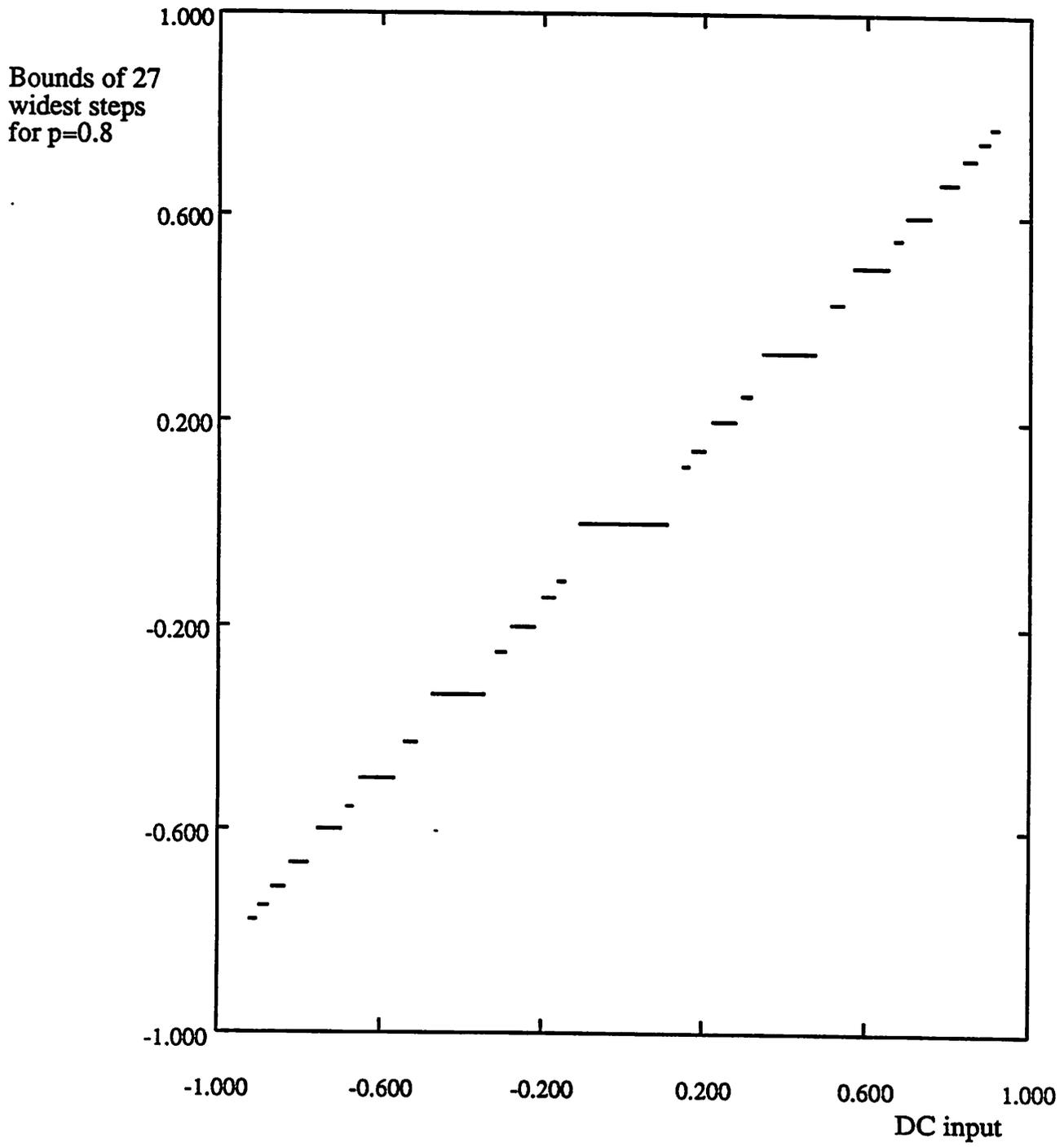


Figure 3

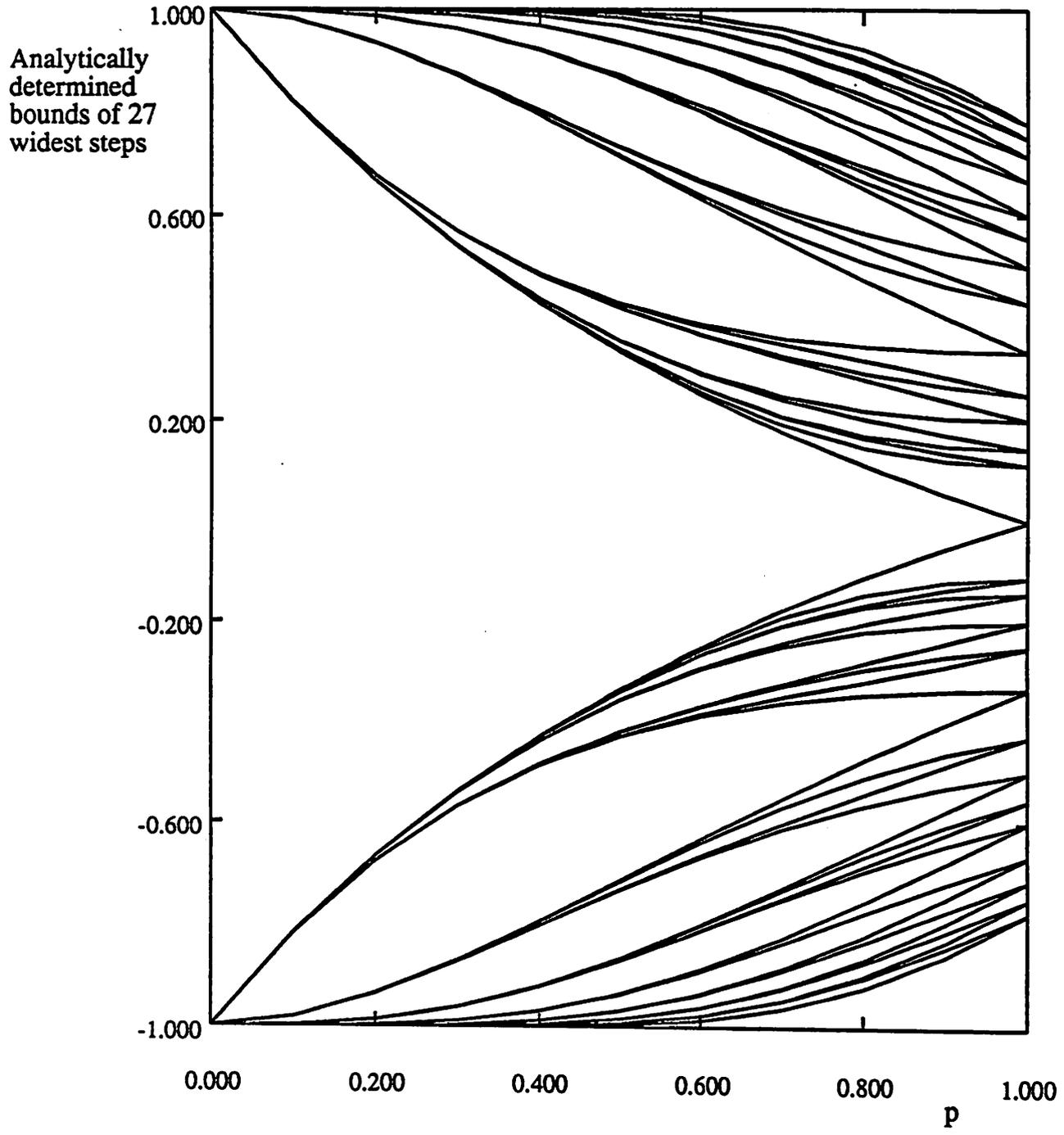
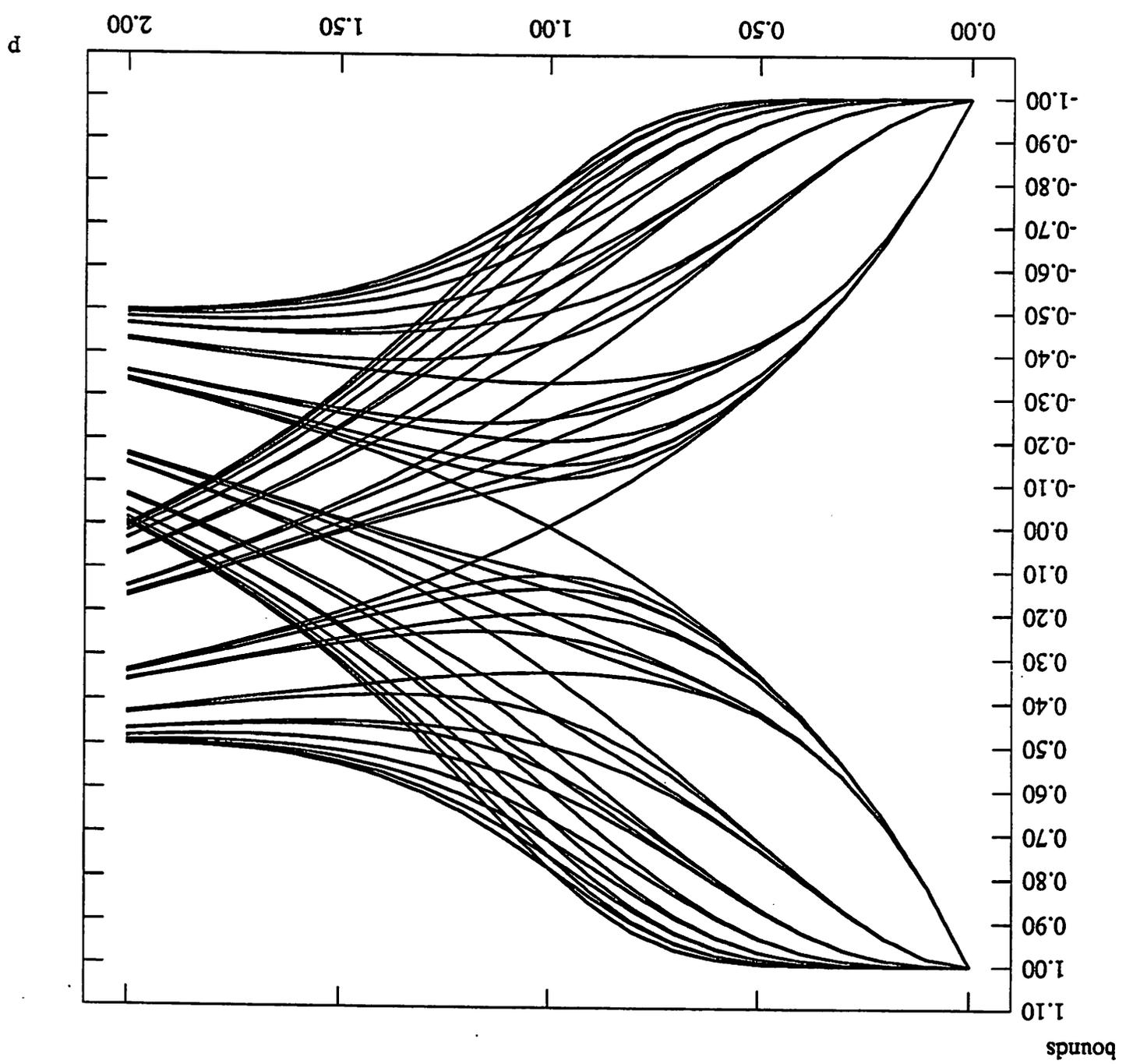


Figure 4

Figure 5



$\times 10^{-3}$

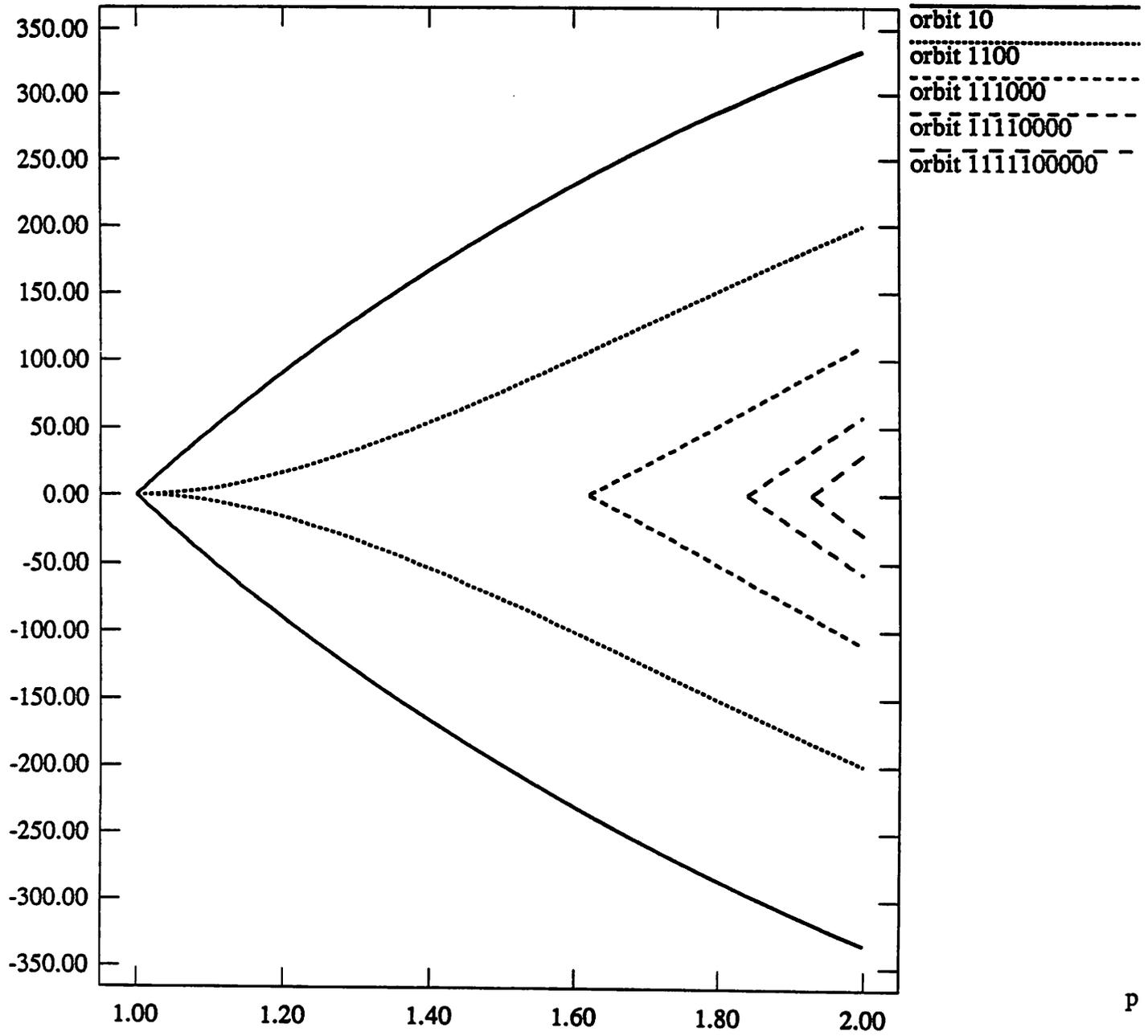
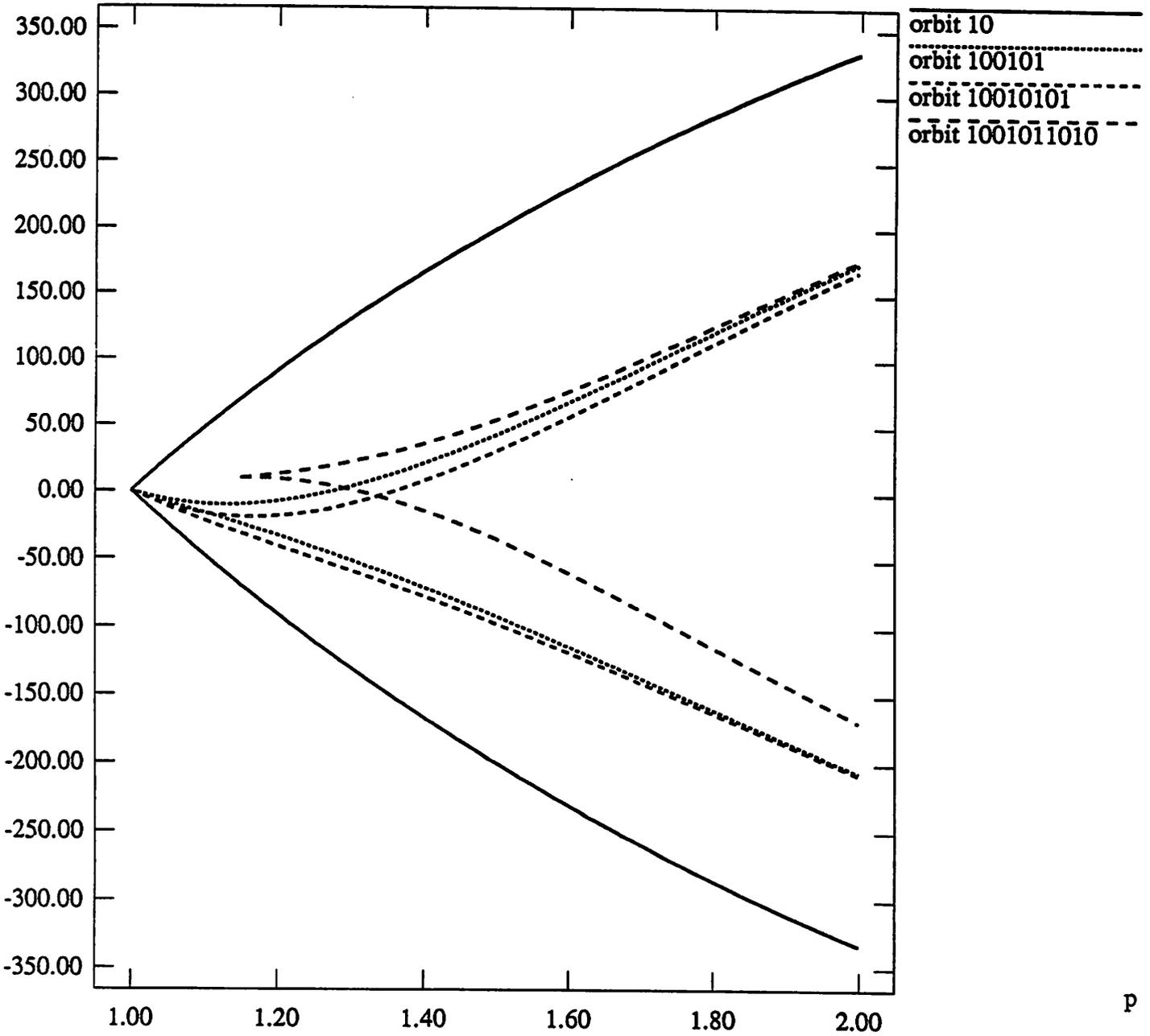


Figure 6 (a)

$x \times 10^{-3}$



P

Figure 6 (b)

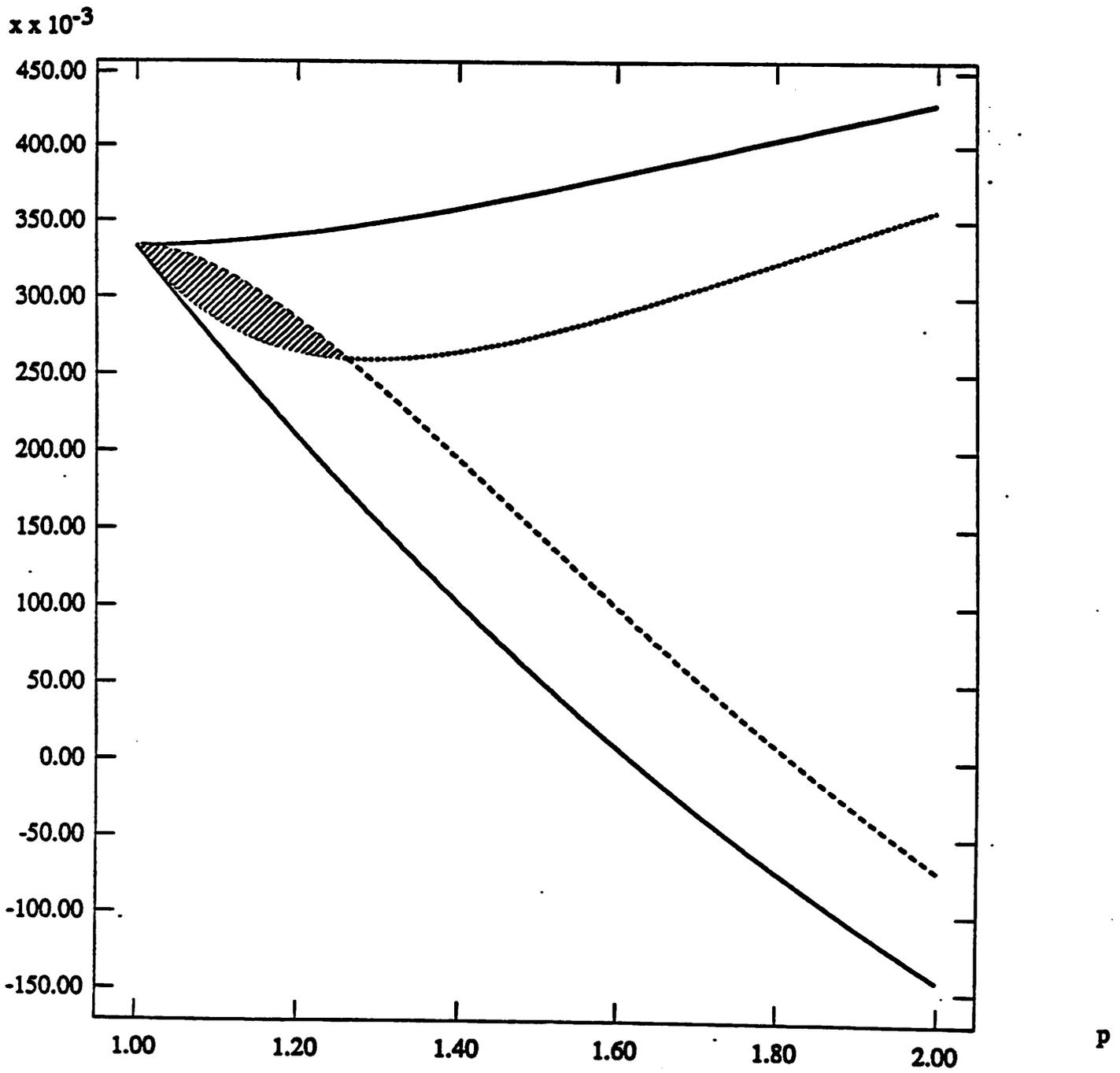


Figure 7

$x = 0$

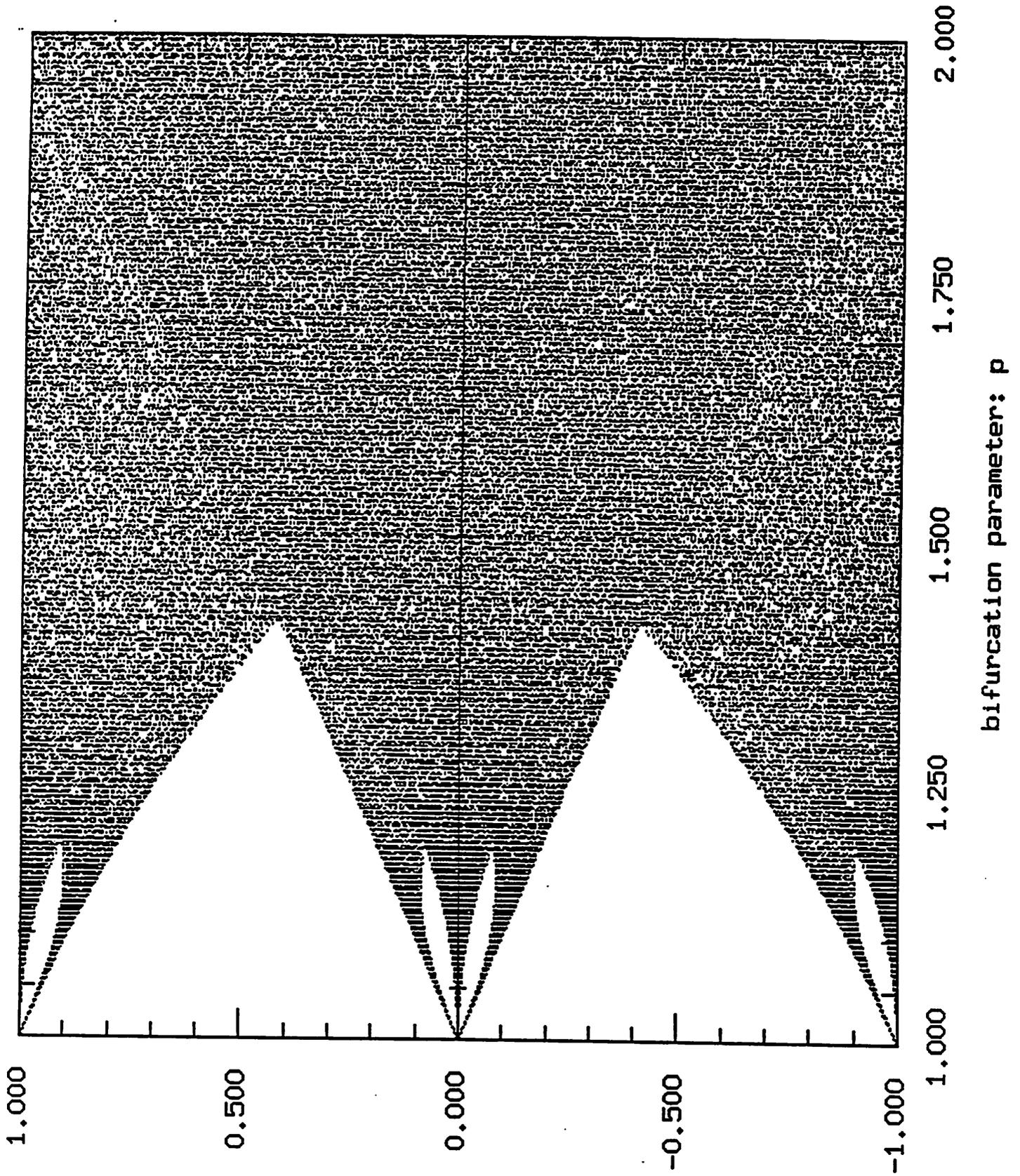
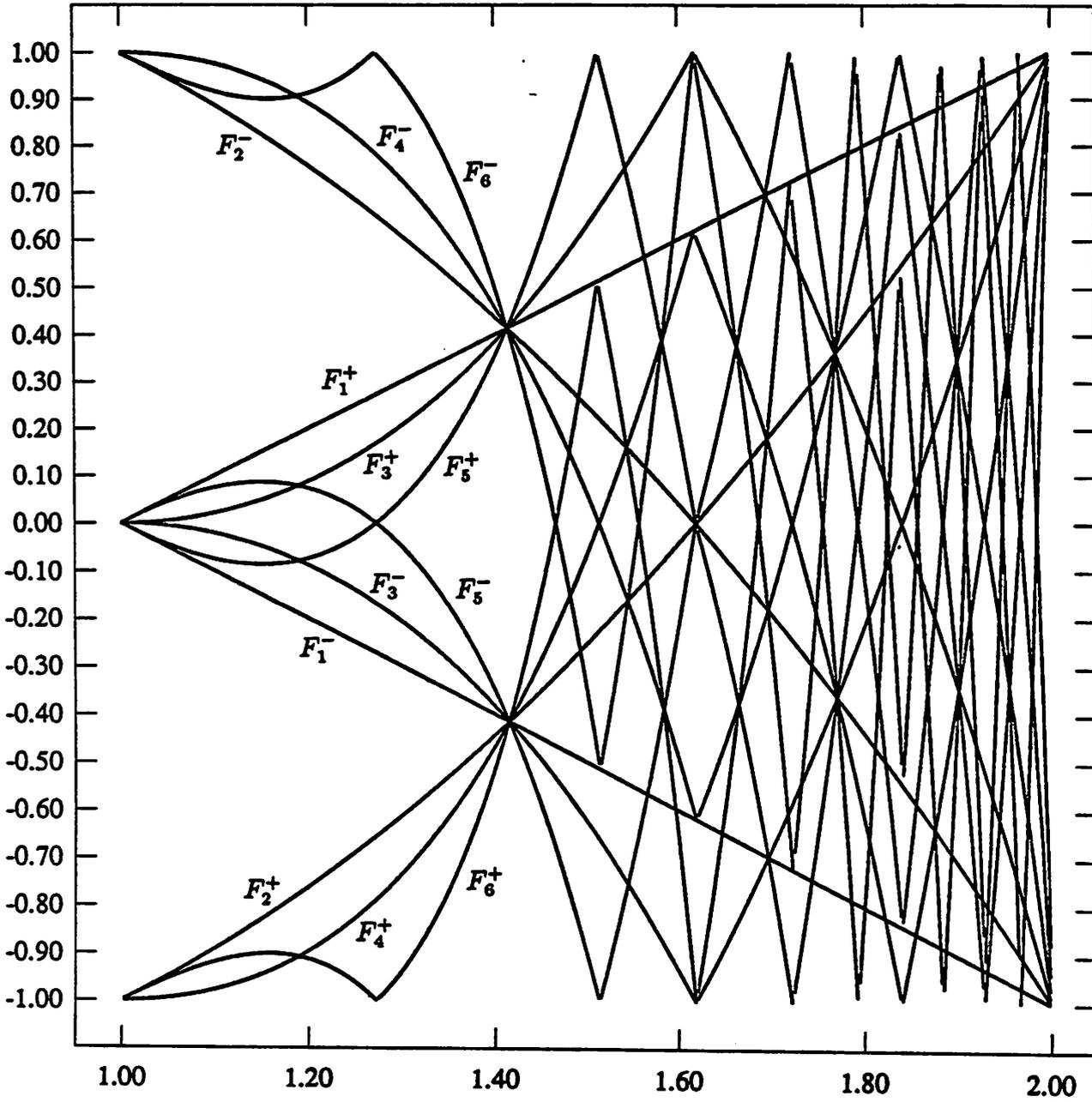


Figure 8

state



P

Figure 9

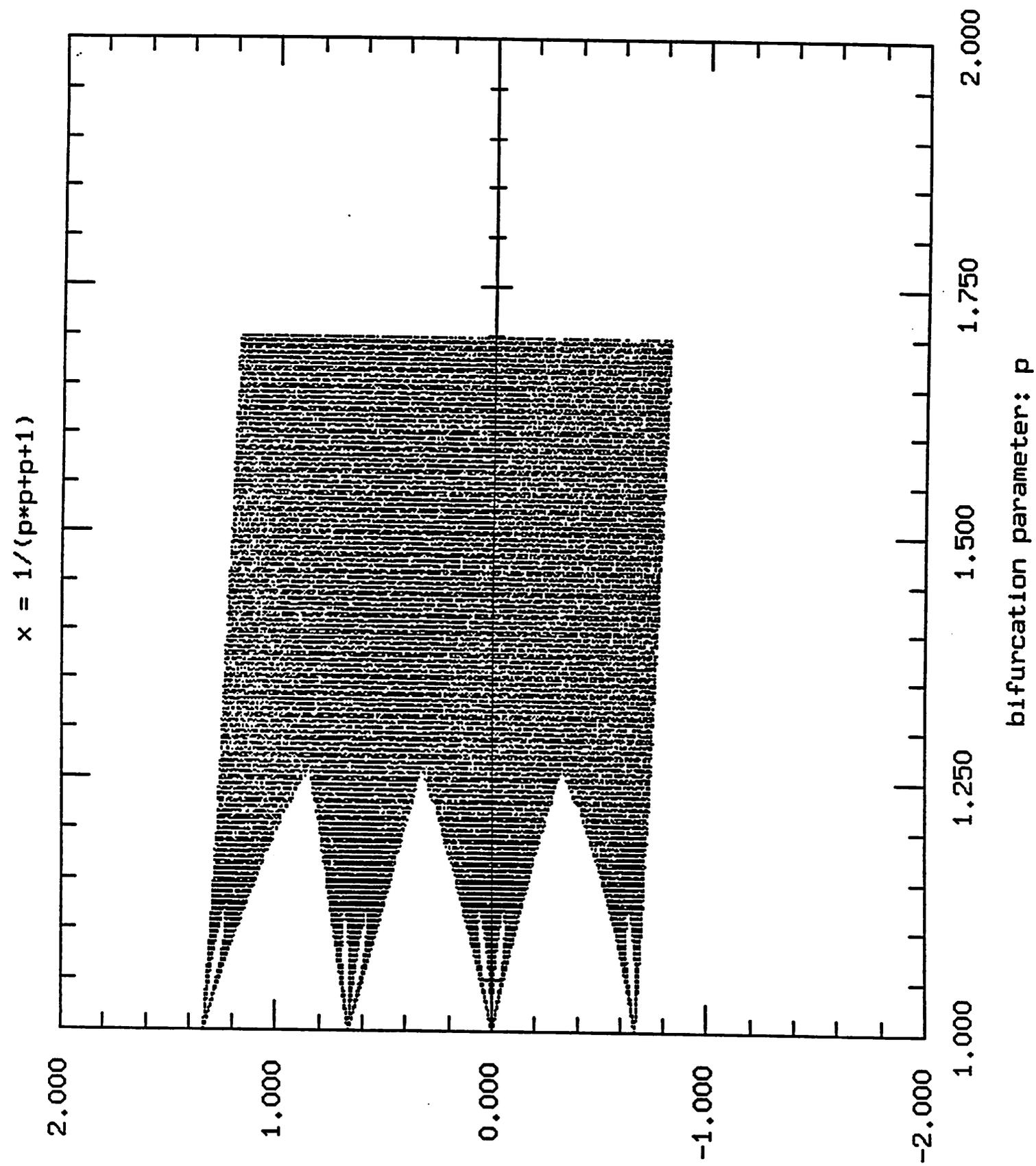


Figure 10

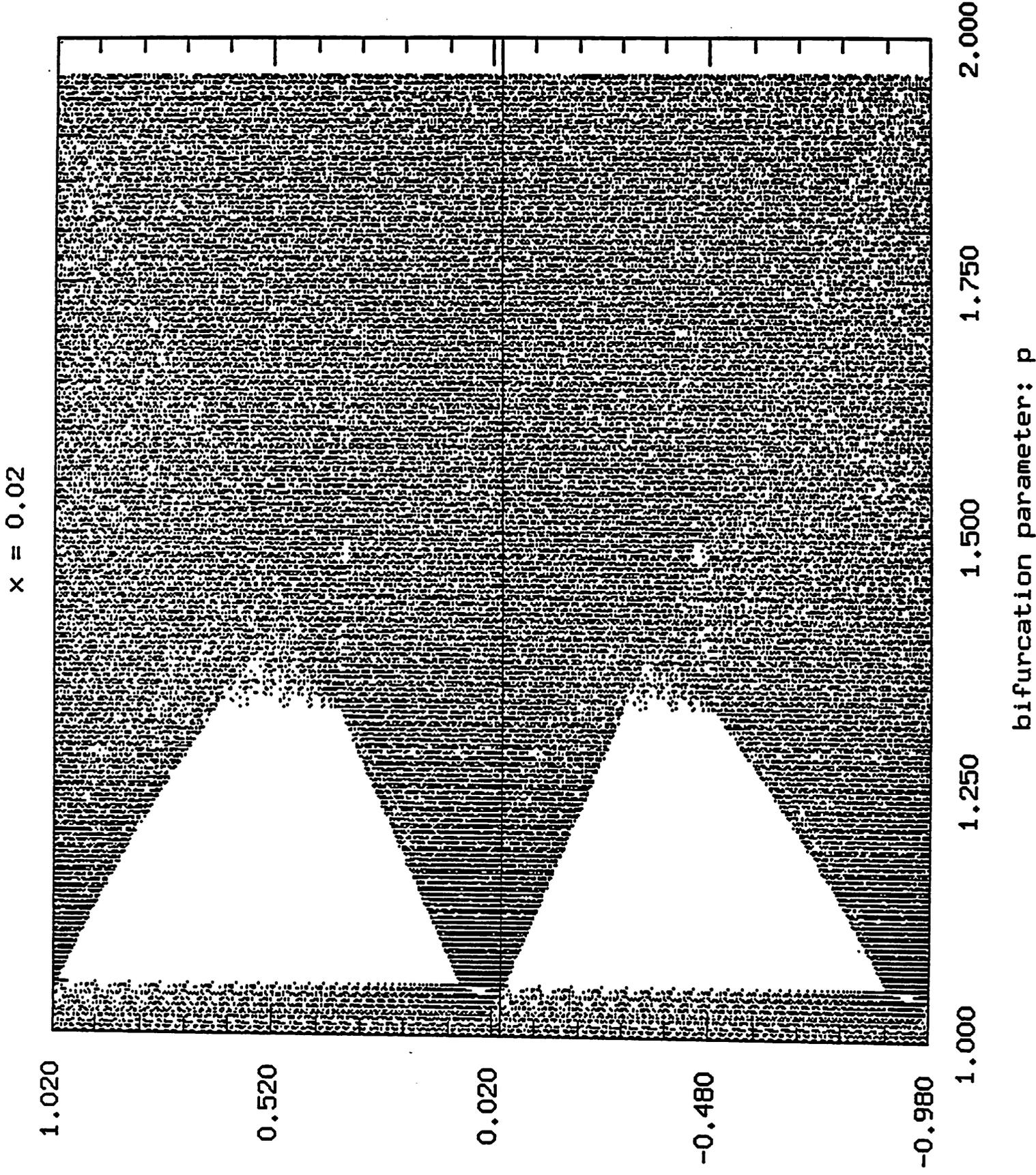


Figure 11

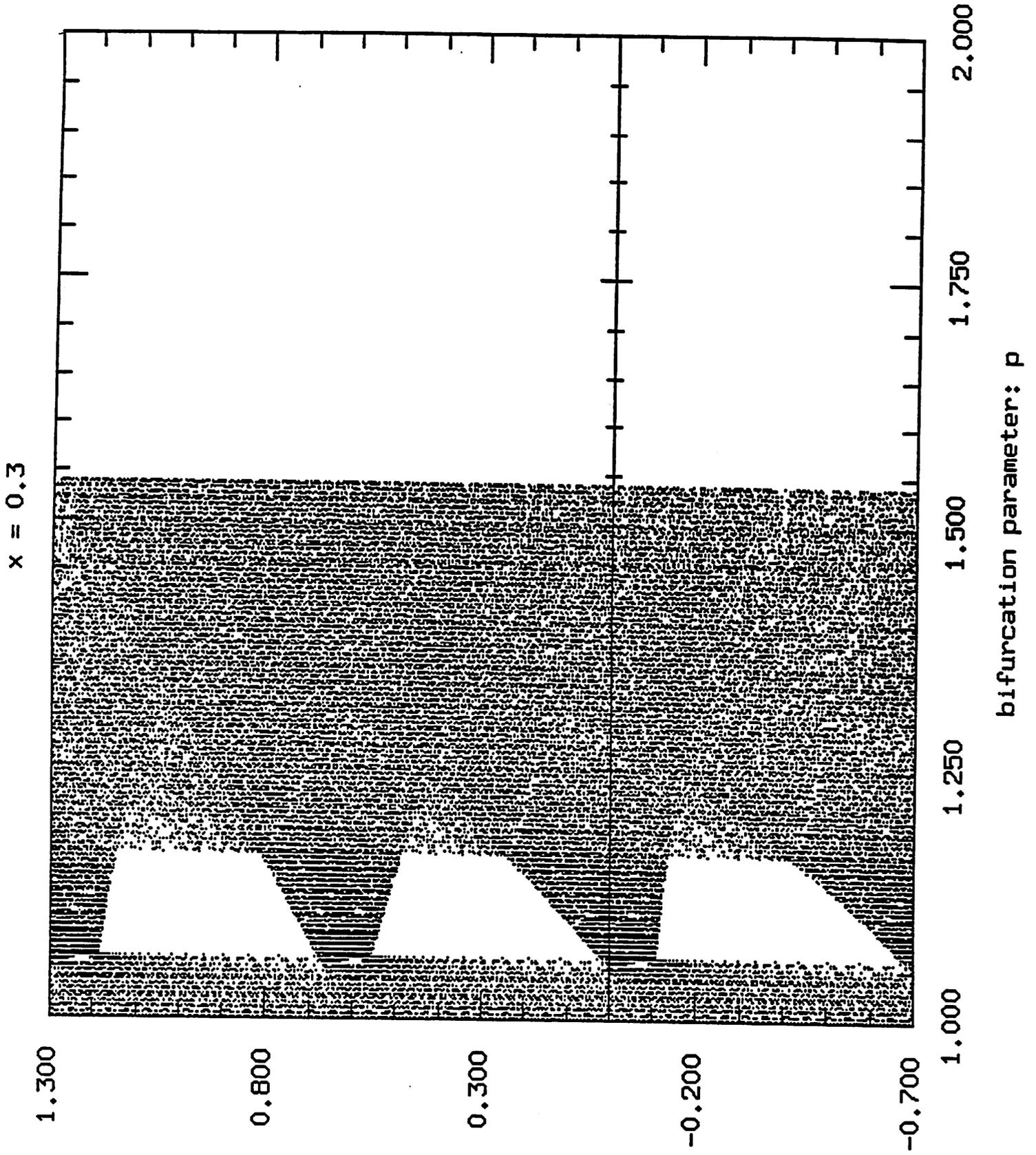


Figure 12

$x = 0.3$

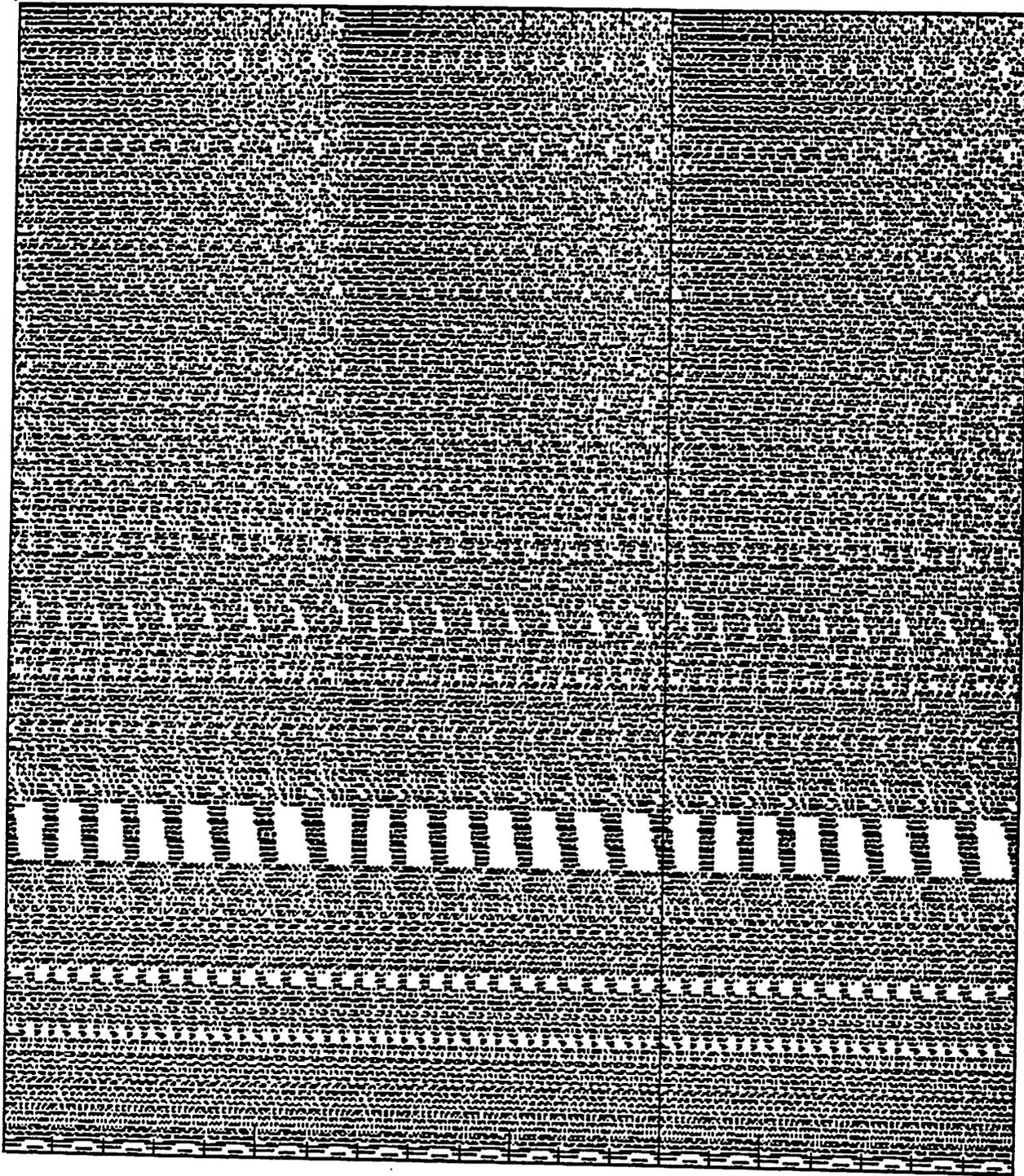
1.300

0.800

0.300

-0.200

-0.700



1.000

1.012

1.025

1.038

1.050

bifurcation parameter: p

Figure 13